

An Inverse Source Problem for Semilinear Stochastic Hyperbolic Equations*

Qi Lü[†] and Yu Wang[‡]

Abstract

This paper investigates an inverse source problem for general semilinear stochastic hyperbolic equations. Motivated by the challenges arising from both randomness and nonlinearity, we develop an iterative regularization method with a global error bound that combines Carleman estimate with fixed-point iteration. The error estimate uses partial lateral Cauchy data and an arbitrary admissible initial guess; it retains a regularization bias. We establish a new Carleman estimate for stochastic hyperbolic equations and prove a weighted energy error bound for the proposed method. Furthermore, we design an efficient numerical algorithm that avoids solving backward stochastic partial differential equations and is robust to randomness in both the model and the data. Numerical experiments are provided to demonstrate the effectiveness of the method.

Keywords. Semilinear stochastic hyperbolic equations, inverse source problem, numerical methods, regularization, Carleman estimate.

1 Introduction

Let $T > 0$, and let $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$ be a complete filtered probability space, where $\mathbf{F} = \{\mathcal{F}_t\}_{t \geq 0}$ is the natural filtration generated by a one-dimensional standard Brownian motion $\{W(t)\}_{t \geq 0}$, augmented by all $\mathbb{P}$ -null sets in $\mathcal{F}$. Write $\mathbb{F}$ for the progressive σ -field with respect to $\mathbf{F}$. Let $\mathcal{H}$ be a Banach space. Denote by $L^2_{\mathbb{F}}(0, T; \mathcal{H})$ the Banach space consisting of all $\mathcal{H}$ -valued $\mathbf{F}$ -adapted processes $X(\cdot)$ satisfying $\mathbb{E}(|X(\cdot)|^2_{L^2(0, T; \mathcal{H})}) < \infty$, equipped with the canonical norm; by $L^\infty_{\mathbb{F}}(0, T; \mathcal{H})$ the Banach space of all $\mathcal{H}$ -valued $\mathbf{F}$ -adapted essentially bounded processes; and by $L^2_{\mathbb{F}}(\Omega; C([0, T]; \mathcal{H}))$ the Banach space of all $\mathcal{H}$ -valued $\mathbf{F}$ -adapted continuous processes $X(\cdot)$ for which $\mathbb{E}(|X(\cdot)|^2_{C([0, T]; \mathcal{H})}) < \infty$ (similarly, one can define $L^2_{\mathbb{F}}(\Omega; H^1(0, T; \mathcal{H}))$).

Let G be a bounded domain in $\mathbb{R}^n$ with a C^2 boundary $\Gamma = \partial G$. Let $(b^{jk})_{1 \leq j, k \leq n} \in C^1(G; \mathbb{R}^n)$ satisfy that $b^{jk} = b^{kj}$ for all $j, k = 1, \dots, n$ and for some constant $s_0 > 0$,

$$\sum_{j, k=1}^n b^{jk} \xi_j \xi_k \geq s_0 |\xi|^2, \text{ for all } (x, \xi) \in G \times \mathbb{R}^n.$$

Denote by $\nu(x) = (\nu^1(x), \nu^2(x), \dots, \nu^n(x))$ the unit outward normal vector of Γ at $x \in \Gamma$. Put

$$\Gamma_0 \triangleq \left\{ x \in \Gamma \mid \sum_{j, k=1}^n b^{jk}(x) \psi_{x_j}(x) \nu^k(x) > 0 \right\}. \quad (1.1)$$

*This work is supported by the NSF of China under grants 12401589 and 12025105, and by the Fundamental Research Funds for the Central Universities 2682024CX013.

[†]School of Mathematics, Sichuan University, Chengdu, P. R. China. Email: lu@scu.edu.cn.

[‡]School of Mathematics, Southwest Jiaotong University, Chengdu, P. R. China. Email: yuwangmath@163.com.

where ψ is given in Condition 1.2. We use the notation $y_{x_i} = y_{x_i}(x) \equiv \frac{\partial y(x)}{\partial x_i}$, where $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and x_i is the i -th coordinate.

Let $Q = (0, T) \times G$, $\Sigma = (0, T) \times \Gamma$ and $\Sigma_0 = (0, T) \times \Gamma_0$. Consider the following semilinear stochastic hyperbolic equation:

$$\begin{cases} du_t - \sum_{j,k=1}^n (b^{jk} u_{x_j})_{x_k} dt = F(u, u_t, \nabla u) dt + a u dW(t) & \text{in } Q, \\ u_t(0) = 0 & \text{in } G, \\ u(0) = u_0 & \text{in } G, \end{cases} \quad (1.2)$$

where $a \in L^\infty_{\mathbb{F}}(0, T; L^\infty(G))$ and $F : [0, T] \times \Omega \times G \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$ satisfying the following assumptions:

Condition 1.1. (1) For each $(\eta, \rho, \zeta) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n$, $F(\cdot, \cdot, \cdot, \eta, \rho, \zeta) : [0, T] \times \Omega \times G$ is an $L^2(G)$ -valued $\mathbf{F}$ -adapted processes;

(2) There exists a constant L such that for a.e. $(t, \omega, x) \in [0, T] \times \Omega \times G$ and any $(\eta_i, \varrho_i, \zeta_i) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n$ ($i = 1, 2$),

$$|F(t, \omega, x, \eta_1, \varrho_1, \zeta_1) - F(t, \omega, x, \eta_2, \varrho_2, \zeta_2)| \leq L(|\eta_1 - \eta_2| + |\varrho_1 - \varrho_2| + |\zeta_1 - \zeta_2|_{\mathbb{R}^n});$$

(3) $|F(\cdot, \cdot, \cdot, 0, 0, 0)| \in L^2_{\mathbb{F}}(0, T; L^2(G))$.

Stochastic hyperbolic equations play a fundamental role in modeling the vibrations of strings and membranes under random force, as well as wave propagation in random media (e.g., [5, 6, 12]). Their inverse problems have significant practical applications. A notable example is the international Solar and Heliospheric Observatory (SOHO) project [13], which seeks to reconstruct the Sun's internal structure by analyzing surface motion measurements [6]. This problem can be mathematically formulated as a series of inverse problems for stochastic hyperbolic equations, one of which is stated below:

Problem 1. Determine the source function u_0 from the lateral Cauchy data $u|_{\Sigma}$ and $\frac{\partial u}{\partial \nu}|_{\Sigma_0}$.

To study Problem 1, we introduce a technical assumption on $(b^{jk})_{1 \leq j, k \leq n}$:

Condition 1.2. There exists a positive function $\psi(\cdot) \in C^2(\overline{G})$ satisfying the following:

(i) For some constant $\mu_0 > 0$ and any $(x, \xi) \in \overline{G} \times \mathbb{R}^n$, it holds that

$$\sum_{j,k=1}^n \left\{ \sum_{j',k'=1}^n \left[2b^{jk'} (b^{j'k} \psi_{x_{j'}})_{x_{k'}} - b^{jk} b^{j'k'} \psi_{x_{j'}} \right] \right\} \xi_j \xi_k \geq \mu_0 \sum_{j,k=1}^n b^{jk} \xi_j \xi_k.$$

(ii) There is no critical point of $\psi(\cdot)$ in $\overline{G}$, i.e.,

$$\min_{x \in \overline{G}} |\nabla \psi(x)| > 0.$$

Remark 1.1. Condition 1.2 ensures the existence of a suitable weight function for establishing the necessary Carleman estimate for linear stochastic hyperbolic equations. This condition represents a pseudoconvexity requirement (cf. [14, XXVIII]). In particular, when the matrix $(b^{jk}(\cdot))_{1 \leq j, k \leq n}$

coincides with the $n \times n$ identity matrix I_n , the function $\varphi(x) = |x - x_0|^2$ with $x_0 \in \mathbb{R}^n \setminus \overline{G}$ satisfies Condition 1.2 with the constant $\mu_0 = 4$.

Observe that if $\psi(\cdot)$ satisfies Condition 1.2, then for any constants $a \geq 1$ and $b \in \mathbb{R}$, the rescaled function $\tilde{\psi} = a\psi + b$ also satisfies this condition with μ_0 replaced by $a\mu_0$. Consequently, we may select ψ , μ_0 , and positive constants $c_0 > 0$, $c_1 > 0$ such that the following strengthened condition holds:

Condition 1.3.

$$\begin{cases} \frac{1}{4} \sum_{j,k=1}^n b^{jk}(x) \psi_{x_j}(x) \psi_{x_k}(x) \geq R_1^2 \triangleq \max_{x \in \overline{G}} \psi(x), & \forall x \in \overline{G}, \\ c_1 < \frac{R_1}{\sqrt{2}T}, & \mu_0 - 4c_1 - c_0 > 0. \end{cases}$$

Inverse problems for deterministic hyperbolic equations have been extensively studied (see, e.g., [3, 4, 16, 20, 22, 37] and references therein). In contrast, inverse problems for stochastic hyperbolic equations remain far from being well studied. The reason is that, compared with inverse problems for deterministic hyperbolic equations, the presence of randomness and uncertainty in both the model and the data introduces significant additional challenges in the study of inverse problems for stochastic hyperbolic equations. Those intrinsic difficulties caused by randomness and uncertainty make it difficult to directly extend methods developed for deterministic hyperbolic equations to the stochastic setting. As a result, the explicit determination and reconstruction of these stochastic sources and potentials remain largely unexplored, representing a critical gap in current research.

To our best knowledge, only a few works concern the determination of these sources and potentials. In recent years, substantial progress has been made in determining the *expectation* and *variance* of *random sources* for random time-harmonic acoustic, electromagnetic, and elastic wave equations; see, e.g., [2, 11, 23–27]. In this paper, we are concerned with the explicit determination and reconstruction of the *source function* for semilinear stochastic hyperbolic equations, which is different from the problem of determining the expectation and variance of random sources.

On the other hand, in [29, 30, 41, 42], it is proved that the unknown source function was determined from partial boundary measurements for semilinear stochastic hyperbolic equations. Nevertheless, those works considered the case of homogeneous Dirichlet boundary conditions and did not provide a method for reconstructing the source function. Recently, in [9] the inverse Cauchy problem for linear stochastic hyperbolic equations was addressed, and numerical examples were provided using a kernel-based learning method, with *full* boundary measurements. For other studies on inverse problems for stochastic wave equations, see also [38], where a class of inverse problems for nonlinear stochastic wave equations is investigated using optimal control techniques, determining part of the boundary shape and the reflection coefficient.

Classical approaches to nonlinear inverse problems primarily rely on optimization techniques, which are local in nature and work well when a good initial guess is available. However, solving nonlinear problems without a reliable initial guess remains a challenging and important issue. One widely recognized global strategy is convexification, which employs a suitable Carleman weight function to convexify the functional, with the justification provided by Carleman estimates. Since its introduction in [18], the convexification method has seen various adaptations [17, 19]. Nevertheless, a major weak point of convexification is its high computational cost [34]. To address this, and inspired by [7, 34], we propose a new method that combines fixed-point iteration with suitable Carleman estimates to obtain a global regularized error estimate for nonlinear inverse problems. Here, “global” means that our approach does not require an initial guess close to the true solution.

Our approach to solving [Problem 1](#) is to construct a map $\mathfrak{G}(u_n) = u_{n+1} = \arg \min J_n(u; u_n)$ via a Tikhonov functional (see [\(3.2\)](#)), which defines a regularized iteration; a fixed point need not equal the true solution. Starting from an arbitrary initial guess u_0 , we recursively define u_n for $n \in \mathbb{N}$. The main result of this paper, [Theorem 3.2](#), gives a geometric iteration-error term and a Hölder-type data-noise term in a weighted energy norm, together with a nonvanishing regularization bias for fixed κ . It does not imply exact recovery or uniqueness for arbitrary observation times.

Finally, we present a numerical [Algorithm 1](#) for solving the inverse [Problem 1](#). Compared with deterministic inverse problems, the stochastic counterparts introduce additional challenges: the lateral Cauchy data (see [Remark 4.2](#)) and the gradients involved in optimization (see [Remark 4.4](#)) are themselves stochastic processes, making classical deterministic optimization methods such as the conjugate gradient method unstable or inefficient (see [Figure 1](#)). To address this, we employ the adaptive moment estimation (Adam) optimizer [\[15\]](#) for stochastic optimization and utilize automatic differentiation [\[35\]](#) to efficiently compute gradients, thereby avoiding the need to solve backward stochastic partial differential equations (see, e.g., [\[8, 40\]](#)). The numerical behavior and its model and implementation qualifications are discussed in [Section 4](#).

This work is not a straightforward extension of previous studies on inverse problems for semilinear deterministic hyperbolic equations [\[7, 34\]](#) or linear stochastic hyperbolic equations [\[9\]](#). Below, we discuss the distinctions and connections with these works in detail.

Comparison with [\[7, 34\]](#). In this work, we study a semilinear stochastic hyperbolic equation with general coefficients, using only partial boundary observation data. Unlike deterministic inverse problems, the stochastic nature of our setting introduces significant challenges, particularly due to the fact that Brownian sample paths are almost surely non-differentiable in time. To overcome these difficulties, we develop:

- A new Carleman estimate tailored to stochastic hyperbolic equations (see [Theorem 2.1](#));
- A modified Tikhonov functional (see [\(3.2\)](#)) designed to handle the irregularity of stochastic processes.

Furthermore, stochastic setting poses additional structural challenges, as the solution spaces typically lack compact embedding properties. Consequently, the techniques employed in [\[7, 34\]](#) for deterministic problems are not directly applicable here.

Comparison with [\[9\]](#). While [\[9\]](#) investigates the inverse Cauchy problem for linear stochastic hyperbolic equations, our work focuses on the semilinear case. The introduction of nonlinearity necessitates fundamentally different approaches at both theoretical and computational levels:

- Unlike the direct regularization method employed in [\[9\]](#), we develop an iterative regularization scheme based on fixed-point iteration for [\(1.2\)](#).
- Our Carleman estimate (see [Remark 2.1](#)) differs from that in [\[9\]](#), as it must account for nonlinear effects in proving iterative convergence.
- While [\[9\]](#) utilizes mesh-free kernel-based learning, we propose [Algorithm 1](#) based on discretization of the Tikhonov functional [\(4.11\)](#).
- Whereas [\[9\]](#) needs measurements of both u and $\frac{\partial u}{\partial \nu}$ on the entire lateral boundary, our approach solves [Problem 1](#) without requiring the boundary data for $\frac{\partial u}{\partial \nu}$ on $(0, T) \times (\Gamma \setminus \Gamma_0)$.

In a word, our new method offers several advantages:

1. It is specifically designed for stochastic PDEs, effectively handling randomness and circumventing the challenges of solving backward stochastic PDEs.
2. It does not require a good initial guess.
3. It exhibits a fast convergence rate.
4. It enables the reconstruction of the source function from partial boundary measurement.
5. It does not require the special structure on the nonlinearity F .

We end this section by highlighting an important point.

Remark 1.2. Although we assume in [Condition 1.1](#) (2) that $F(t, \omega, x, \cdot, \cdot, \cdot)$ is Lipschitz continuous, this may not hold in practice. A cutoff argument applies when F is locally Lipschitz in its state variables, with Lipschitz and growth bounds uniform in (t, ω, x) on each bounded state set. One must also know a constant M in advance such that the true solution u^* of equation (1.2) satisfies $|u^*|_{L^\infty_{\mathbb{F}}(\Omega, W^{1,\infty}(0,T;W^{1,\infty}(G)))} \leq M$. These are additional hypotheses; a cutoff alone cannot regularize a non-Lipschitz singularity. Then we construct a C^1 function $\chi : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$ satisfying $\chi = 1$ in $\{(\eta, \varrho, \zeta) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \mid |\eta| + |\varrho| + |\zeta|_{\mathbb{R}^n} \leq M\}$ and $\chi = 0$ in $\{(\eta, \varrho, \zeta) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \mid |\eta| + |\varrho| + |\zeta|_{\mathbb{R}^n} > 2M\}$. We then solve [Problem 1](#) for equation (1.2) with F replaced by χF . [Example 4.2](#) uses an already globally Lipschitz truncated exponential nonlinearity.

The rest of this paper is organized as follows. In [Section 2](#), we establish a Carleman estimate for a linear stochastic hyperbolic equation. In [Section 3](#), we propose an iterative regularization scheme to solve [Problem 1](#) and establish its convergence properties. Finally, in [Section 4](#) we present the numerical algorithm for [Problem 1](#) along with numerical experiments.

2 Carleman estimate for a linear stochastic hyperbolic equation

In this section, we establish a global Carleman estimate for the following linear stochastic hyperbolic equation:

$$\begin{cases} dz_t - \sum_{j,k=1}^n (b^{jk} z_{x_j})_{x_k} dt = \Upsilon dt + az dW(t) & \text{in } Q, \\ z = 0 & \text{on } \Sigma, \\ \frac{\partial z}{\partial \nu} = 0 & \text{on } \Sigma_0. \end{cases} \quad (2.1)$$

Theorem 2.1. Assume [Conditions 1.2](#) and [1.3](#) hold, $b^{jk} \in C^3(\overline{G})$ and $\psi \in C^4(\overline{G})$. These smoothness hypotheses ensure that the multiplier Ψ used below has the derivatives required by the weighted identity. Then there exist positive constants C and λ_0 such that for all $\lambda \geq \lambda_0$, the solution z to (2.1) with initial condition $z_t(0) = 0$ satisfies the following Carleman estimate:

$$\mathbb{E} \int_Q \theta^2 (\lambda z_t^2 + \lambda |\nabla z|^2 + \lambda^3 z^2) dx dt \leq C \mathbb{E} \int_G \theta^2 (\lambda |\nabla z|^2 + \lambda^3 |z|^2) |_{t=T} dx + C \mathbb{E} \int_Q \theta^2 |\Upsilon|^2 dx dt, \quad (2.2)$$

where the weight function θ is defined by

$$\theta = e^\ell, \quad \ell(t, x) = \lambda(\psi(x) - c_1 t^2). \quad (2.3)$$

Remark 2.1. We emphasize that [Theorem 2.1](#) employs a different Carleman estimate than [\[9\]](#), where the weight function θ is chosen with $\theta_t(0) = 0$ to ensure iteration convergence ([Theorem 3.2](#)). Furthermore, since our assumption on the measurement noise is given by [\(3.6\)](#), it is sufficient here to derive the Carleman estimate under homogeneous boundary conditions.

Here and in what follows, we denote by C a generic positive constant that depends only on the domain G , time horizon T , and coefficients $(b^{jk}(\cdot))_{1 \leq j,k \leq n}$ and a , the fixed weight parameters and their required derivative bounds, and the Lipschitz constant of F , whose value may change from line to line. The absorption threshold in [Theorem 3.2](#) also depends on κ .

To prove [Theorem 2.1](#), we recall the following known result.

Lemma 2.2. *[31, Lemma 10.18] Let $\phi \in C^2((0, T) \times \mathbb{R}^n)$, $b^{jk} = b^{kj} \in C^2(\mathbb{R}^n)$ for $j, k = 1, 2, \dots, n$, and $\ell \in C^3((0, T) \times \mathbb{R}^n)$, $\Psi \in C^2((0, T) \times \mathbb{R}^n)$. Let φ be an $H^2(G)$ -valued, $\mathbf{F}$ -adapted process such that φ_t is an $L^2(G)$ -valued Itô process. Set $\theta = e^\ell$ and $w = \theta\varphi$. Then, for a.e. $x \in G$ and $\mathbb{P}$ -a.s. $\omega \in \Omega$,*

$$\begin{aligned}
& \theta \left(-2\phi \ell_t w_t + 2 \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k} + \Psi w \right) \left[\phi d\varphi_t - \sum_{j,k=1}^n (b^{jk} \varphi_{x_j})_{x_k} dt \right] \\
& + \sum_{j,k=1}^n \left[\sum_{j',k'=1}^n \left(2b^{jk} b^{j'k'} \ell_{x_{j'}} w_{x_j} w_{x_{k'}} - b^{jk} b^{j'k'} \ell_{x_j} w_{x_{j'}} w_{x_{k'}} \right) - 2\phi b^{jk} \ell_t w_{x_j} w_t + \phi b^{jk} \ell_{x_j} w_t^2 \right. \\
& \quad \left. + \Psi b^{jk} w_{x_j} w - \left(\mathcal{A} \ell_{x_j} + \frac{\Psi_{x_j}}{2} \right) b^{jk} w^2 \right]_{x_k} dt \\
& + d \left\{ \phi \sum_{j,k=1}^n b^{jk} \ell_t w_{x_j} w_{x_k} - 2\phi \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k} w_t + \phi^2 \ell_t w_t^2 - \phi \Psi w_t w + \left[\phi \mathcal{A} \ell_t + \frac{1}{2} (\phi \Psi)_t \right] w^2 \right\} \\
& = \left[(\phi^2 \ell_t)_t + \sum_{j,k=1}^n (\phi b^{jk} \ell_{x_j})_{x_k} - \phi \Psi \right] w_t^2 dt - 2 \sum_{j,k=1}^n \left[(\phi b^{jk} \ell_{x_k})_t + b^{jk} (\phi \ell_t)_{x_k} \right] w_{x_j} w_t dt \\
& + \sum_{j,k=1}^n c^{jk} w_{x_j} w_{x_k} dt + \mathcal{B} w^2 dt + \left(-2\phi \ell_t w_t + 2 \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k} + \Psi w \right)^2 dt + \phi^2 \theta^2 \ell_t (d\varphi_t)^2, \quad (2.4)
\end{aligned}$$

where

$$\begin{cases} \mathcal{A} \triangleq \phi(\ell_t^2 - \ell_{tt}) - \sum_{j,k=1}^n (b^{jk} \ell_{x_j} \ell_{x_k} - b_{x_k}^{jk} \ell_{x_j} - b^{jk} \ell_{x_j x_k}) - \Psi, \\ \mathcal{B} \triangleq \mathcal{A} \Psi + (\phi \mathcal{A} \ell_t)_t - \sum_{j,k=1}^n (\mathcal{A} b^{jk} \ell_{x_j})_{x_k} + \frac{1}{2} \left[(\phi \Psi)_{tt} - \sum_{j,k=1}^n (b^{jk} \Psi_{x_j})_{x_k} \right], \\ c^{jk} \triangleq (\phi b^{jk} \ell_t)_t + \sum_{j',k'=1}^n \left[2b^{jk'} (b^{j'k} \ell_{x_{j'}})_{x_{k'}} - (b^{jk} b^{j'k'} \ell_{x_{j'}})_{x_{k'}} \right] + \Psi b^{jk}. \end{cases} \quad (2.5)$$

Throughout this paper, for $\lambda \in \mathbb{R}$ and $p \in \mathbb{N}$, we denote by $O(\lambda^p)$ any function of order λ^p for large λ .

Proof of [Theorem 2.1](#). Choose

$$\varphi = z, \quad \phi \equiv 1, \quad \Psi = \ell_{tt} + \sum_{j,k=1}^n (b^{jk} \ell_{x_j})_{x_k} - c_0 \lambda, \quad (2.6)$$

in [Theorem 2.2](#). We now estimate the terms in [\(2.4\)](#) one by one.

We first note that

$$\ell_t = -2c_1\lambda t, \quad \ell_{tt} = -2c_1\lambda, \quad \ell_{x_j} = \lambda\psi_{x_j}, \quad \ell_{tx_j} = 0, \quad j = 1, \dots, n. \quad (2.7)$$

From [\(2.6\)](#) and [\(2.7\)](#), we derive that

$$\left[(\phi^2 \ell_t)_t + \sum_{j,k=1}^n (\phi b^{jk} \ell_{x_j})_{x_k} - \phi \Psi \right] w_t^2 = c_0 \lambda w_t^2, \quad (2.8)$$

and

$$-2 \sum_{j,k=1}^n [(\phi b^{jk} \ell_{x_k})_t + b^{jk} (\phi \ell_t)_{x_k}] w_{x_j} w_t = 0. \quad (2.9)$$

Furthermore, combining [\(2.6\)](#) and [\(2.7\)](#) and [Condition 1.2](#) yields

$$\begin{aligned} \sum_{j,k=1}^n c^{jk} w_{x_j} w_{x_k} &= \sum_{j,k=1}^n \left\{ (b^{jk} \ell_t)_t + \sum_{j',k'=1}^n [2b^{jk'} (b^{j'k} \ell_{x_{j'}})_{x_{k'}} - (b^{jk} b^{j'k'} \ell_{x_{j'}})_{x_{k'}}] + \Psi b^{jk} \right\} w_{x_j} w_{x_k} \\ &= \sum_{j,k=1}^n \left\{ b^{jk} \ell_{tt} + \sum_{j',k'=1}^n [2b^{jk'} (b^{j'k} \ell_{x_{j'}})_{x_{k'}} - b_{x_{k'}}^{jk} b^{j'k'} \ell_{x_{j'}} - b^{jk} (b^{j'k'} \ell_{x_{j'}})_{x_{k'}}] \right. \\ &\quad \left. + [\ell_{tt} + \sum_{j',k'=1}^n (b^{j'k'} \ell_{x_{j'}})_{x_{k'}} - c_0 \lambda] b^{jk} \right\} w_{x_j} w_{x_k} \\ &= \sum_{j,k=1}^n \left\{ 2b^{jk} \ell_{tt} + \sum_{j',k'=1}^n [2b^{jk'} (b^{j'k} \ell_{x_{j'}})_{x_{k'}} - b_{x_{k'}}^{jk} b^{j'k'} \ell_{x_{j'}}] - c_0 \lambda b^{jk} \right\} w_{x_j} w_{x_k} \\ &= \sum_{j,k=1}^n \left\{ -4c_1 \lambda b^{jk} + \lambda \sum_{j',k'=1}^n [2b^{jk'} (b^{j'k} \psi_{x_{j'}})_{x_{k'}} - b_{x_{k'}}^{jk} b^{j'k'} \psi_{x_{j'}}] - c_0 \lambda b^{jk} \right\} w_{x_j} w_{x_k} \\ &\geq \lambda(\mu_0 - 4c_1 - c_0) \sum_{j,k=1}^n b^{jk} w_{x_j} w_{x_k}. \end{aligned} \quad (2.10)$$

From [\(2.7\)](#), we obtain

$$\mathcal{A} = \lambda^2 \left(4c_1^2 t^2 - \sum_{j,k=1}^n b^{jk} \psi_{x_j} \psi_{x_k} \right) + \mathcal{O}(\lambda) \quad (2.11)$$

$$(\mathcal{A} \ell_t)_t = -24c_1^3 \lambda^3 t^2 + 2c_1 \lambda^3 \sum_{j,k=1}^n b^{jk} \psi_{x_j} \psi_{x_k} + \mathcal{O}(\lambda^2),$$

and

$$\begin{aligned} - \sum_{j,k=1}^n (\mathcal{A} b^{jk} \ell_{x_j})_{x_k} &= -\lambda \sum_{j,k=1}^n [\mathcal{A} (b^{jk} \psi_{x_j})_{x_k} + \mathcal{A}_{x_k} b^{jk} \psi_{x_j}] \\ &= \lambda^3 \sum_{j,k=1}^n \left\{ -4c_1^2 t^2 (b^{jk} \psi_{x_j})_{x_k} + \sum_{j',k'=1}^n (b^{j'k'} \psi_{x_{j'}} \psi_{x_{k'}}) (b^{jk} \psi_{x_j})_{x_k} \right\} \end{aligned}$$

$$+ \sum_{j',k'=1}^n (b^{j'k'} \psi_{x_{j'}} \psi_{x_{k'}})_{x_k} b^{jk} \psi_{x_j} \} + \mathcal{O}(\lambda^2).$$

Combining this with (2.5)–(2.7) yields

$$\begin{aligned} \mathcal{B} &= (4c_1 + c_0) \lambda^3 \sum_{j,k=1}^n b^{jk} \psi_{x_j} \psi_{x_k} + \lambda^3 \sum_{j,k=1}^n \sum_{j',k'=1}^n b^{jk} \psi_{x_j} (b^{j'k'} \psi_{x_{j'}} \psi_{x_{k'}})_{x_k} \\ &\quad - 4(8c_1^3 + c_0 c_1^2) \lambda^3 t^2 + O(\lambda^2). \end{aligned} \quad (2.12)$$

From Condition 1.2, we get that

$$\begin{aligned} &\sum_{j,k=1}^n \sum_{j',k'=1}^n b^{jk} \psi_{x_j} (b^{j'k'} \psi_{x_{j'}} \psi_{x_{k'}})_{x_k} \\ &= \sum_{j,k=1}^n \sum_{j',k'=1}^n (b^{jk} b_{x_k}^{j'k'} \psi_{x_j} \psi_{x_{j'}} \psi_{x_{k'}} + 2b^{jk} b^{j'k'} \psi_{x_j} \psi_{x_k} \psi_{x_{j'}} \psi_{x_{k'}}) \\ &= \sum_{j,k=1}^n \sum_{j',k'=1}^n (b^{jk} b_{x_k}^{j'k'} \psi_{x_j} \psi_{x_{j'}} \psi_{x_{k'}} - 2b^{j'k'} b_{x_{j'}}^{jk} \psi_{x_j} \psi_{x_k} \psi_{x_{k'}} + 2b^{j'k'} (b^{jk} \psi_{x_k})_{x_{j'}} \psi_{x_j} \psi_{x_{k'}}) \\ &= \sum_{j,k=1}^n \left\{ \sum_{j',k'=1}^n \left[2b^{jk'} (b^{j'k} \psi_{x_{j'}})_{x_{k'}} - b_{x_{k'}}^{jk} b^{j'k'} \psi_{x_{j'}} \right] \right\} \psi_{x_j} \psi_{x_k} \\ &\geq \mu_0 \sum_{j,k=1}^n b^{jk} \psi_{x_j} \psi_{x_k}. \end{aligned} \quad (2.13)$$

Under Condition 1.3, we establish from (2.12) and (2.13) the lower bound

$$\begin{aligned} \mathcal{B} &\geq (4c_1 + c_0 + \mu_0) \lambda^3 \sum_{j,k=1}^n b^{jk} \psi_{x_j} \psi_{x_k} - 4(8c_1^3 + c_0 c_1^2) \lambda^3 t^2 + O(\lambda^2) \\ &\geq 4(4c_1 + c_0 + \mu_0) \lambda^3 R_1^2 - 4(8c_1^3 + c_0 c_1^2) \lambda^3 T^2 + O(\lambda^2) \\ &\geq (4c_1 + c_0) \lambda^3 (4R_1^2 - 8c_1^2 T^2) + O(\lambda^2). \end{aligned}$$

By Condition 1.3, there exists a positive constant λ_1 such that for all $\lambda \geq \lambda_1$, we have

$$\mathcal{B} w^2 \geq C \lambda^3 w^2. \quad (2.14)$$

Since $\ell_t(0) = 0$ and $w_t(0) = \theta(0)z_t(0) + \theta_t(0)z(0) = 0$, applying (2.6), (2.7) and (2.11) and Conditions 1.2 and 1.3, the Cauchy-Schwarz inequality and Young's inequality yields for any $\varepsilon > 0$:

$$\begin{aligned} &\mathbb{E} \int_0^T \int_G d \left\{ \phi \sum_{j,k=1}^n b^{jk} \ell_t w_{x_j} w_{x_k} - 2\phi \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k} w_t + \phi^2 \ell_t w_t^2 - \phi \Psi w_t w + \left[\phi \mathcal{A} \ell_t + \frac{1}{2} (\phi \Psi)_t \right] w^2 \right\} \\ &= \mathbb{E} \int_G \left(\sum_{j,k=1}^n b^{jk} \ell_t w_{x_j} w_{x_k} - 2 \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k} w_t + \ell_t w_t^2 - \Psi w_t w + \mathcal{A} \ell_t w^2 \right) \Big|_{t=T} dx \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E} \int_G \left[-2c_1 T \lambda \left(\sum_{j,k=1}^n b^{jk} w_{x_j} w_{x_k} + w_t^2 \right) - 2 \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k} w_t - \Psi w_t w \right. \\
&\quad \left. - 2c_1 T \lambda^3 \left(4c_1^2 t^2 - \sum_{j,k=1}^n b^{jk} \psi_{x_j} \psi_{x_k} \right) w^2 + \mathcal{O}(\lambda^2) w^2 \right] \Big|_{t=T} dx \\
&\leq \mathbb{E} \int_G \left[-2c_1 T \lambda (s_0 |\nabla w|^2 + w_t^2) + 2\lambda \left| \sum_{j,k=1}^n b^{jk} \psi_{x_j} w_{x_k} \right| |w_t| + |\Psi w_t w| \right. \\
&\quad \left. C \lambda^3 w^2 + \mathcal{O}(\lambda^2) w^2 \right] \Big|_{t=T} dx \\
&\leq \mathbb{E} \int_G \left[-2c_1 T \lambda (s_0 |\nabla w|^2 + w_t^2) + \varepsilon \lambda w_t^2 + C(\varepsilon) \lambda \left| \sum_{j,k=1}^n b^{jk} \psi_{x_j} w_{x_k} \right|^2 + C(\varepsilon) \lambda w^2 \right. \\
&\quad \left. C \lambda^3 w^2 + \mathcal{O}(\lambda^2) w^2 \right] \Big|_{t=T} dx \\
&\leq \mathbb{E} \int_G \left[-2c_1 T \lambda (s_0 |\nabla w|^2 + w_t^2) + \varepsilon \lambda |w_t|^2 + C(\varepsilon) \lambda (|\nabla w|^2 + w^2) \right. \\
&\quad \left. C \lambda^3 w^2 + \mathcal{O}(\lambda^2) w^2 \right] \Big|_{t=T} dx.
\end{aligned}$$

Choosing $\varepsilon < 2c_1 T$ and noting $c_1 T < R_1$, there exists $\lambda_2 > 0$ such that for all $\lambda \geq \lambda_2$, we have

$$\begin{aligned}
&\mathbb{E} \int_0^T \int_G d \left\{ \phi \sum_{j,k=1}^n b^{jk} \ell_t w_{x_j} w_{x_k} - 2\phi \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k} w_t + \phi^2 \ell_t w_t^2 - \phi \Psi w_t w + \left[\phi \mathcal{A} \ell_t + \frac{1}{2} (\phi \Psi)_t \right] w^2 \right\} \\
&\leq C \mathbb{E} \int_G (\lambda^3 |w(T)|^2 + \lambda |\nabla w(T)|^2) dx. \tag{2.15}
\end{aligned}$$

Thanks to the divergence theorem, and noting that $w = \theta z$ and $w = z = 0$ on Σ , we have

$$\begin{aligned}
&\mathbb{E} \int_Q \sum_{j,k=1}^n \left[\sum_{j',k'=1}^n \left(2b^{jk} b^{j'k'} \ell_{x_{j'}} w_{x_j} w_{x_{k'}} - b^{jk} b^{j'k'} \ell_{x_j} w_{x_{j'}} w_{x_{k'}} \right) - 2\phi b^{jk} \ell_t w_{x_j} w_t + \phi b^{jk} \ell_{x_j} w_t^2 \right. \\
&\quad \left. + \Psi b^{jk} w_{x_j} w - \left(\mathcal{A} \ell_{x_j} + \frac{\Psi_{x_j}}{2} \right) b^{jk} w^2 \right]_{x_k} dt \\
&= \mathbb{E} \int_\Sigma \sum_{j,k=1}^n \sum_{j',k'=1}^n \left(2b^{jk} b^{j'k'} \ell_{x_{j'}} w_{x_j} w_{x_{k'}} - b^{jk} b^{j'k'} \ell_{x_j} w_{x_{j'}} w_{x_{k'}} \right) \nu_k d\Gamma dt \\
&= \lambda \mathbb{E} \int_\Sigma \sum_{j,k=1}^n \sum_{j',k'=1}^n \left(2b^{jk} b^{j'k'} \psi_{x_{j'}} w_{x_j} w_{x_{k'}} - b^{jk} b^{j'k'} \psi_{x_j} w_{x_{j'}} w_{x_{k'}} \right) \nu_k d\Gamma dt. \tag{2.16}
\end{aligned}$$

Since $w = \theta z$, thanks to (2.3), there exists a positive constant C such that

$$\frac{1}{C} \theta^2 [\lambda (|\nabla z|^2 + |z_t|^2) + \lambda^3 |z|^2] \leq \lambda (|\nabla w|^2 + |w_t|^2) + \lambda^3 |w|^2 \leq C \theta^2 [\lambda (|\nabla z|^2 + |z_t|^2) + \lambda^3 |z|^2]. \tag{2.17}$$

Integrating (2.4) over Q , taking mathematical expectation of the integral, and noting (2.8)–(2.10)

and (2.14)–(2.17), for all $\lambda \geq \max\{\lambda_1, \lambda_2\}$, we get that

$$\begin{aligned}
& \mathbb{E} \int_Q \theta \left\{ \left(-2\ell_t w_t + 2 \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k} + \Psi w \right) \left[dz_t - \sum_{j,k=1}^n (b^{jk} \varphi_{x_j})_{x_k} dt \right] \right\} dx \\
& + \lambda \mathbb{E} \int_{\Sigma} \sum_{j,k=1}^n \sum_{j',k'=1}^n (2b^{jk} b^{j'k'} \psi_{x_{j'}} w_{x_j} w_{x_{k'}} - b^{jk} b^{j'k'} \psi_{x_j} w_{x_{j'}} w_{x_{k'}}) \nu_k d\Gamma dt \\
& + C \mathbb{E} \int_G \theta^2 (\lambda |\nabla z|^2 + \lambda^3 |z|^2) |_{t=T} dx - \mathbb{E} \int_Q \theta^2 \ell_t |dz_t|^2 dx \\
& \geq C \mathbb{E} \int_Q \theta^2 \left[(\lambda |z_t|^2 + \lambda |\nabla z|^2) + \lambda^3 |z|^2 \right] dx dt \\
& + \mathbb{E} \int_Q \left(-2\ell_t w_t + 2 \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k} + \Psi w \right)^2 dx dt.
\end{aligned} \tag{2.18}$$

Recalling that $w = \theta z$ and $z = 0$ on Σ , we find that

$$\nabla w = \nabla \theta z + \theta \nabla z = \theta \nabla z = \theta \frac{\partial z}{\partial \nu} \nu, \quad \text{on } \Sigma. \tag{2.19}$$

Noting that $\frac{\partial z}{\partial \nu} = 0$ on Σ_0 and (1.1) and (2.19), we find that

$$\begin{aligned}
& \mathbb{E} \int_{\Sigma} \sum_{j,k=1}^n \sum_{j',k'=1}^n \left(2b^{jk} b^{j'k'} \psi_{x_{j'}} w_{x_j} w_{x_{k'}} - b^{jk} b^{j'k'} \psi_{x_j} w_{x_{j'}} w_{x_{k'}} \right) \nu^k d\Gamma dt \\
& = \mathbb{E} \int_{\Sigma} \sum_{j,k=1}^n \sum_{j',k'=1}^n \left(2b^{jk} b^{j'k'} \psi_{x_{j'}} \nu^j \nu^{k'} - b^{jk} b^{j'k'} \psi_{x_j} \nu^{j'} \nu^{k'} \right) \nu^k \theta^2 \left| \frac{\partial z}{\partial \nu} \right|^2 d\Gamma dt \\
& = \mathbb{E} \int_{\Sigma} \sum_{j,k=1}^n \sum_{j',k'=1}^n \left(2b^{j'k'} b^{jk} \psi_{x_j} \nu^{j'} \nu^k \nu^{k'} - b^{jk} b^{j'k'} \psi_{x_j} \nu^{j'} \nu^{k'} \nu^k \right) \theta^2 \left| \frac{\partial z}{\partial \nu} \right|^2 d\Gamma dt \\
& = \mathbb{E} \int_{\Sigma} \left(\sum_{j',k'=1}^n b^{j'k'} \nu^{j'} \nu^{k'} \right) \left(\sum_{j,k=1}^n b^{jk} \psi_{x_j} \nu^k \right) \theta^2 \left| \frac{\partial z}{\partial \nu} \right|^2 d\Gamma dt \\
& \leq \mathbb{E} \int_{\Sigma_0} \left(\sum_{j',k'=1}^n b^{j'k'} \nu^{j'} \nu^{k'} \right) \left(\sum_{j,k=1}^n b^{jk} \psi_{x_j} \nu^k \right) \theta^2 \left| \frac{\partial z}{\partial \nu} \right|^2 d\Gamma dt \\
& = 0.
\end{aligned} \tag{2.20}$$

From (2.1) and (2.7), we infer that

$$\begin{aligned}
-\mathbb{E} \int_Q \theta^2 \ell_t |dz_t|^2 dx &= -\mathbb{E} \int_Q (-2c_1 \lambda t) \theta^2 |az|^2 dx dt \\
&\leq 2c_1 \lambda T \mathbb{E} \int_Q \theta^2 |az|^2 dx dt \\
&\leq 2c_1 \lambda T |a|_{L_{\mathbb{F}}^{\infty}(0,T;L^{\infty}(G))}^2 \mathbb{E} \int_Q \theta^2 |z|^2 dx dt,
\end{aligned} \tag{2.21}$$

and

$$\begin{aligned}
& \mathbb{E} \int_Q \theta \left\{ \left(-2\ell_t w_t + 2 \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k} + \Psi w \right) \left[dz_t - \sum_{j,k=1}^n (b^{jk} z_{x_j})_{x_k} dt \right] \right\} dx \\
& \leq \mathbb{E} \int_Q \theta^2 |\Upsilon|^2 dx dt + \mathbb{E} \int_Q \left(-2\ell_t w_t + 2 \sum_{j,k=1}^n b^{jk} \ell_{x_j} w_{x_k} + \Psi w \right)^2 dx dt.
\end{aligned} \tag{2.22}$$

Combining (2.18) and (2.20)–(2.22), choosing $\lambda_0 \geq \max\{C, \lambda_1, \lambda_2\}$, we obtain the desired Carleman estimate (2.2). $\square$

3 Convergence analysis for source determination

Denote $\mathcal{H}^p = L_{\mathbb{F}}^2(0, T; H^p(G)) \cap L_{\mathbb{F}}^2(\Omega; H^1(0, T; H^1(G)))$ for $p \in \mathbb{N}$. Define an operator $\mathcal{P} : \mathcal{H}^2 \rightarrow L_{\mathbb{F}}^2(0, T; L^2(G))$ by as follows:

$$\begin{aligned}
(\mathcal{P}\varphi)(t, x) & \triangleq \varphi_t(t, x) - \int_0^t \sum_{j,k=1}^n (b^{jk}(s, x) \varphi_{x_j}(s, x))_{x_k} ds \\
& \quad - \int_0^t a(s, x) \varphi(s, x) dW(s), \quad \mathbb{P}\text{-a.s.}, \quad \forall \varphi \in \mathcal{H}, \quad (t, x) \in Q.
\end{aligned}$$

The admissible set is given by

$$\mathcal{U} \triangleq \left\{ \varphi \in \mathcal{H}^2 \mid \mathcal{P}\varphi \in L_{\mathbb{F}}^2(\Omega; H^1(0, T; L^2(G))), \quad \varphi_t(0) = 0, \quad \varphi|_{\Sigma} = f, \quad \frac{\partial \varphi}{\partial \nu} \Big|_{\Sigma_0} = g \right\}. \tag{3.1}$$

For any solution u to (1.2), we have:

$$\mathcal{P}u = \int_0^t \mathcal{F}(u) ds, \quad \text{where } \mathcal{F}(u) \triangleq F(u, u_t, \nabla u).$$

Under Condition 1.1, a solution $u \in \mathcal{H}^2$ satisfies $\mathbf{IF}(u) \in L_{\mathbb{F}}^2(\Omega; H^1(0, T; L^2(G)))$. We assume that the exact solution u^* has this strong regularity. For exact data it belongs to $\mathcal{U}$; for the noisy data in (3.6), $u^* + \mathcal{E} \in \mathcal{U}$, so the admissible set is nonempty.

Let θ be the weight function defined in (2.3). We introduce the following weighted Sobolev spaces:

$$\begin{aligned}
& L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G))) \\
& \triangleq \left\{ \varphi \in L_{\mathbb{F}}^2(\Omega; H^1(0, T; L^2(G))) \mid \mathbb{E} \int_0^T \int_G \theta^2 |\varphi|^2 dx dt + \mathbb{E} \int_0^T \int_G \theta^2 |\varphi_t|^2 dx dt < \infty \right\}
\end{aligned}$$

and

$$L_{\mathbb{F}}^{2,w}(0, T; H^p(G)) \triangleq \left\{ \varphi \in L_{\mathbb{F}}^2(0, T; H^p(G)) \mid \sum_{|\alpha|=0}^p \mathbb{E} \int_0^T \int_G \theta^2 |D_{\alpha} \varphi|^2 dx dt < \infty \right\},$$

where $\alpha = (\alpha_1, \dots, \alpha_n)$ is a multi-index with $|\alpha| = \sum_{i=1}^n \alpha_i$ and $D_{\alpha} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$ denotes the weak derivative operator. Let

$$\mathcal{H}_{\lambda, c_1}^p \triangleq L_{\mathbb{F}}^{2,w}(0, T; H^p(G)) \cap L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; H^1(G))),$$

where λ, c_1 are the parameters from (2.3). Each of these spaces, when equipped with its canonical

norm, forms a Banach space. In $\mathcal{H}_{\lambda,c}^p$ the weight is $e^{\lambda(\psi-ct^2)}$; in particular $c = 0$ denotes the purely spatial weight. The strong regularization space $\mathcal{H}_{\lambda,0}^2$ includes ∇v_t . For the iteration error we use the distinct energy norm

$$|v|_{\mathcal{E}_{\lambda,c_1}}^2 := \mathbb{E} \int_Q \theta^2 (|v_t|^2 + |\nabla v|^2 + |v|^2) dx dt.$$

No mixed derivative ∇v_t is included in this energy norm.

Let the integral operator $\mathbf{I}$ be defined by

$$\mathbf{I}f(t) = \int_0^t f(s) ds \quad \text{for } t \in [0, T].$$

Let us construct a sequence $\{u_n\}_{n \in \mathbb{N}} \in \mathcal{U}$ by induction.

1. Select an arbitrary initial solution $u_0 \in \mathcal{U}$.
2. Assume we have constructed a sequence $u_n \in \mathcal{U}$ for $n \geq 0$.
3. Define u_{n+1} as the minimizer of the Tikhonov regularization functional:

$$J_n(\varphi) = |\mathcal{P}\varphi - \mathbf{I}\mathcal{F}(u_n)|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))}^2 + \kappa |\varphi|_{\mathcal{H}_{\lambda,0}^2}^2, \quad (3.2)$$

where $\varphi \in \mathcal{U}$ and $\kappa \in (0, 1)$ is the regularization parameter.

We have the following result on the existence and uniqueness of the solution to the minimization problem.

Proposition 3.1. *For each fixed regularization parameter $\kappa \in (0, 1)$ and iteration index $n \geq 0$, the Tikhonov regularization functional J_n defined in (3.2) admits a unique minimizer u_{n+1} in the admissible set $\mathcal{U}$.*

Proof. We begin by defining the weighted solution space:

$$\mathcal{U}_0^w \triangleq \left\{ \varphi \in \mathcal{H}^2 \mid \mathcal{P}\varphi \in L_{\mathbb{F}}^2(\Omega; H^1(0, T; L^2(G))), \quad \varphi_t(0) = 0, \quad \varphi|_{\Sigma} = 0, \quad \frac{\partial \varphi}{\partial \nu} \Big|_{\Sigma_0} = 0 \right\}.$$

For fixed $\kappa \in (0, 1)$, we endow $\mathcal{U}_0^w$ with the inner product:

$$\langle p, q \rangle_{\mathcal{U}_0^w} \triangleq \langle \mathcal{P}p, \mathcal{P}q \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))} + \kappa \langle p, q \rangle_{\mathcal{H}_{\lambda,0}^2}, \quad \forall p, q \in \mathcal{U}_0^w. \quad (3.3)$$

The space $\mathcal{U}_0^w$ is complete for this graph norm. Indeed, a Cauchy sequence converges to v in $\mathcal{H}^2$ and its images under $\mathcal{P}$ converge in $L^2(\Omega; H^1(0, T; L^2(G)))$. The defining deterministic integral and Itô integral are continuous from $\mathcal{H}^2$ into $L^2(\Omega; C([0, T]; L^2(G)))$, by Cauchy–Schwarz and the Itô maximal inequality. Their limits identify the graph limit as $\mathcal{P}v$ and give convergence of v_t in this continuous-time space. Thus $v_t(0) = 0$ is preserved. The spatial trace theorem preserves both homogeneous boundary traces, and adaptedness is closed under L^2 convergence. Hence completion does not enlarge the admissible class.

Given a reference function $R \in \mathcal{U}$, we consider the shifted variable $v = \varphi - R$ and define the functional:

$$\overline{\mathcal{J}}_n(v) = |\mathcal{P}v + \mathcal{P}R - \mathbf{I}\mathcal{F}(u_n)|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))}^2 + \kappa |v + R|_{\mathcal{H}_{\lambda,0}^2}^2, \quad \forall v \in \mathcal{U}_0^w.$$

The equivalence between minimizing $\mathcal{J}_n$ and $\overline{\mathcal{J}}_n$ is clear: v_{n+1} minimizes $\overline{\mathcal{J}}_n$ if and only if $u_{n+1} = v_{n+1} + R$ minimizes $\mathcal{J}_n$.

The variational principle yields the Euler-Lagrange equation for the minimizer v_{n+1} :

$$\langle \mathcal{P}v_{n+1} + \mathcal{P}R - \mathbf{IF}(u_n), \mathcal{P}\rho \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))} + \kappa \langle v_{n+1} + R, \rho \rangle_{\mathcal{H}_{\lambda,0}^2} = 0, \quad \forall \rho \in \mathcal{U}_0^w. \quad (3.4)$$

To verify that v_{n+1} is the minimizer of $\overline{\mathcal{J}}_n$, consider for any $\rho \in \mathcal{U}_0^w$:

$$\begin{aligned} & \overline{\mathcal{J}}_n(\rho) - \overline{\mathcal{J}}_n(v_{n+1}) \\ &= |\mathcal{P}\rho + \mathcal{P}R - \mathbf{IF}(u_n)|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))}^2 + \kappa |\rho + R|_{\mathcal{H}_{\lambda,0}^2}^2 \\ & \quad - |\mathcal{P}v_{n+1} + \mathcal{P}R - \mathbf{IF}(u_n)|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))}^2 - \kappa |v_{n+1} + R|_{\mathcal{H}_{\lambda,0}^2}^2 \\ &= |\mathcal{P}v_{n+1} + \mathcal{P}R - \mathbf{IF}(u_n) + \mathcal{P}(\rho - v_{n+1})|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))}^2 \\ & \quad + \kappa |v_{n+1} + R + (\rho - v_{n+1})|_{\mathcal{H}_{\lambda,0}^2}^2 - |\mathcal{P}v_{n+1} + \mathcal{P}R - \mathbf{IF}(u_n)|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))}^2 - \kappa |v_{n+1} + R|_{\mathcal{H}_{\lambda,0}^2}^2 \\ &= 2\langle \mathcal{P}v_{n+1} + \mathcal{P}R - \mathbf{IF}(u_n), \mathcal{P}(\rho - v_{n+1}) \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))} \\ & \quad + 2\kappa \langle v_{n+1} + R, \rho - v_{n+1} \rangle_{\mathcal{H}_{\lambda,0}^2} + |\mathcal{P}(\rho - v_{n+1})|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))}^2 + \kappa |\rho - v_{n+1}|_{\mathcal{H}_{\lambda,0}^2}^2 \\ &= |\mathcal{P}(\rho - v_{n+1})|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))}^2 + \kappa |\rho - v_{n+1}|_{\mathcal{H}_{\lambda,0}^2}^2 \\ &\geq 0, \end{aligned}$$

where we used (3.4) to eliminate the first two terms.

Using (3.3), we can rewrite (3.4) as:

$$\langle v_{n+1}, \rho \rangle_{\mathcal{U}_0^w} = \langle \mathbf{IF}(u_n) - \mathcal{P}R, \mathcal{P}\rho \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))} - \kappa \langle R, \rho \rangle_{\mathcal{H}_{\lambda,0}^2}, \quad \forall \rho \in \mathcal{U}_0^w. \quad (3.5)$$

The right-hand side of (3.5) defines a bounded linear functional on $\mathcal{U}_0^w$ since

$$\begin{aligned} & \langle \mathbf{IF}(u_n) - \mathcal{P}R, \mathcal{P}\rho \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))} - \kappa \langle R, \rho \rangle_{\mathcal{H}_{\lambda,0}^2} \\ & \leq (|\mathcal{P}R|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))} + C|R|_{\mathcal{H}_{\lambda,0}^2} + C|u_n|_{\mathcal{E}_{\lambda,c_1}} + |\mathbf{IF}(0)|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))})|\rho|_{\mathcal{U}_0^w}. \end{aligned}$$

Thanks to Riesz representation theorem, there exists a unique $w_{n+1} \in \mathcal{U}_0^w$ satisfying (3.5).

Assume that $\tilde{w}_{n+1}$ is another minimizer of $\overline{\mathcal{J}}_n$. Then $\tilde{w}_{n+1}$ satisfies (3.5). For all $\rho \in \mathcal{U}_0^w$, we obtain

$$\langle w_{n+1} - \tilde{w}_{n+1}, \rho \rangle_{\mathcal{U}_0^w} = 0,$$

which implies that $w_{n+1} = \tilde{w}_{n+1}$. Hence, the minimizer of $\overline{\mathcal{J}}_n$ is unique. $\square$

Let u^* denote the exact solution to the semilinear stochastic hyperbolic system (1.2) with the exact boundary data $u^*|_{\Sigma} = f^*$ and $\frac{\partial u}{\partial \nu}\Big|_{\Sigma_0} = g^*$. Assume that there is a function $\mathcal{E} \in \mathcal{H}^2$ satisfying

$$|\mathcal{P}\mathcal{E}|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))} + |\mathcal{E}|_{\mathcal{H}^2} \leq \delta, \quad \mathcal{E}_t(0) = 0, \quad f|_{\Sigma} = f^* + \mathcal{E}|_{\Sigma}, \quad g = g^* + \frac{\partial \mathcal{E}}{\partial \nu}\Big|_{\Sigma_0}, \quad (3.6)$$

where $\delta > 0$ quantifies the noise level in the measurements.

Remark 3.1. Here, we assume that the boundary measurement data f and g can be used to construct a function $\mathcal{E}$, so that the error is considered only in terms of $\mathcal{E}$, as in [20, 34]. This is a key point that allows us to work with partial boundary measurements. However, our numerical experiments show that the method converges effectively even for more general noise, as in (4.9). At present, we do not know how to obtain convergence results for semilinear stochastic hyperbolic

equations under the assumption $|f - f^*| + |g - g^*| < \delta$ in a suitable norm, as has been achieved for linear stochastic hyperbolic equations in [9].

We have the following regularized error estimate.

Theorem 3.2. *Assume the smoothness hypotheses of Theorem 2.1, $u^* \in \mathcal{H}^2$, and (3.6). Let the regularization parameter $\kappa \in (0, 1)$ be fixed, and let c_1 satisfy Condition 1.3. Fix $\gamma \in (0, 1)$. There exist positive constants C and $\tilde{\lambda}_0 > C$ such that for all $\lambda \geq \tilde{\lambda}_0$, $\delta < e^{-\lambda R_1^2/\gamma}$ and $n \in \mathbb{N}$, we have*

$$|u_{n+1} - u^*|_{\mathcal{E}_{\lambda, c_1}}^2 \leq \left(\frac{C}{\lambda}\right)^{n+1} |u_0 - u^*|_{\mathcal{E}_{\lambda, c_1}}^2 + C(\delta^{2-2\gamma} + \kappa |u^*|_{\mathcal{H}_{\lambda, 0}^2}^2),$$

where u_{n+1} is the minimizer of the functional (3.2).

Proof. Since u_{n+1} minimizes (3.2), the Euler-Lagrange equation (3.4) gives

$$\langle \mathcal{P}u_{n+1} - \mathbf{IF}(u_n), \mathcal{P}\rho \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))} + \kappa \langle u_{n+1}, \rho \rangle_{\mathcal{H}_{\lambda, 0}^2} = 0, \quad \forall \rho \in \mathcal{U}_0^w.$$

For the exact solution u^* satisfying $\mathcal{P}u^* - \mathbf{IF}(u^*) = 0$, we have

$$\langle \mathcal{P}u^* - \mathbf{IF}(u^*), \mathcal{P}\rho \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))} + \kappa \langle u^*, \rho \rangle_{\mathcal{H}_{\lambda, 0}^2} = \kappa \langle u^*, \rho \rangle_{\mathcal{H}_{\lambda, 0}^2}, \quad \forall \rho \in \mathcal{U}_0^w.$$

Subtracting the above two equations yields that

$$\langle \mathcal{P}(u_{n+1} - u^*) - (\mathbf{IF}(u_n) - \mathbf{IF}(u^*)), \mathcal{P}\rho \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))} + \kappa \langle u_{n+1} - u^*, \rho \rangle_{\mathcal{H}_{\lambda, 0}^2} = -\kappa \langle u^*, \rho \rangle_{\mathcal{H}_{\lambda, 0}^2}. \quad (3.7)$$

Letting $z = u_{n+1} - u^* - \mathcal{E}$, by (3.6), we get $z \in \mathcal{U}_0^w$. From (3.7), we arrive at

$$\langle \mathcal{P}(z + \mathcal{E}) - (\mathbf{IF}(u_n) - \mathbf{IF}(u^*)), \mathcal{P}z \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))} + \kappa \langle z + \mathcal{E}, z \rangle_{\mathcal{H}_{\lambda, 0}^2} = -\kappa \langle u^*, z \rangle_{\mathcal{H}_{\lambda, 0}^2}.$$

Hence, we obtain

$$\begin{aligned} & |\mathcal{P}z|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))}^2 + \kappa |z|_{\mathcal{H}_{\lambda, 0}^2}^2 \\ &= \langle (\mathbf{IF}(u_n) - \mathbf{IF}(u^*)), \mathcal{P}z \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))} - \kappa \langle u^*, z \rangle_{\mathcal{H}_{\lambda, 0}^2} \\ &\quad - \langle \mathcal{P}\mathcal{E}, \mathcal{P}z \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))} - \kappa \langle \mathcal{E}, z \rangle_{\mathcal{H}_{\lambda, 0}^2}. \end{aligned} \quad (3.8)$$

Next, we estimate the right-hand side of (3.8) one by one. From Cauchy-Schwarz inequality and Condition 1.1, for $\varepsilon > 0$, we get that

$$\begin{aligned} & \langle (\mathbf{IF}(u_n) - \mathbf{IF}(u^*)), \mathcal{P}z \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))} - \kappa \langle u^*, z \rangle_{\mathcal{H}_{\lambda, 0}^2} - \kappa \langle \mathcal{E}, z \rangle_{\mathcal{H}_{\lambda, 0}^2} \\ &\leq C(\varepsilon) |\mathbf{IF}(u_n) - \mathbf{IF}(u^*)|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))}^2 + \varepsilon |\mathcal{P}z|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))}^2 + \frac{\kappa}{\varepsilon} |u^*|_{\mathcal{H}_{\lambda, 0}^2}^2 + \frac{\kappa}{\varepsilon} |\mathcal{E}|_{\mathcal{H}_{\lambda, 0}^2}^2 \\ &\quad + 2\varepsilon \kappa |z|_{\mathcal{H}_{\lambda, 0}^2}^2 \\ &\leq C(\varepsilon) |u_n - u^*|_{\mathcal{E}_{\lambda, c_1}}^2 + \varepsilon |\mathcal{P}z|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))}^2 + \frac{\kappa}{\varepsilon} |u^*|_{\mathcal{H}_{\lambda, 0}^2}^2 + \frac{\kappa}{\varepsilon} |\mathcal{E}|_{\mathcal{H}_{\lambda, 0}^2}^2 + 2\varepsilon \kappa |z|_{\mathcal{H}_{\lambda, 0}^2}^2 \end{aligned} \quad (3.9)$$

and that

$$-\langle \mathcal{P}\mathcal{E}, \mathcal{P}z \rangle_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))} \leq C(\varepsilon) |\mathcal{P}\mathcal{E}|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))}^2 + \varepsilon |\mathcal{P}z|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))}^2. \quad (3.10)$$

Combining (3.8)–(3.10), choosing ε sufficiently small, we conclude that

$$|\mathcal{P}z|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))}^2 + \kappa |z|_{\mathcal{H}_{\lambda, 0}^2}^2$$

$$\leq C(|u_n - u^*|_{\mathcal{E}_{\lambda, c_1}}^2 + \kappa|u^*|_{\mathcal{H}_{\lambda, 0}}^2 + |\mathcal{PE}|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))}^2 + |\mathcal{E}|_{\mathcal{H}_{\lambda, 0}}^2). \quad (3.11)$$

Apply [Theorem 2.1](#) to z , for $\lambda \geq \lambda_0$, we have

$$|\mathcal{P}z|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))}^2 \geq c \mathbb{E} \int_Q \theta^2 (\lambda z_t^2 + \lambda |\nabla z|^2 + \lambda^3 z^2) dx dt - C \mathbb{E} \int_G \theta^2 (\lambda |\nabla z|^2 + \lambda^3 |z|^2) |_{t=T} dx. \quad (3.12)$$

From [\(2.3\)](#) and trace theorem of Sobolev space, we obtain that

$$\begin{aligned} & \mathbb{E} \int_G \theta^2 (\lambda |\nabla z|^2 + \lambda^3 |z|^2) |_{t=T} dx \\ & \leq \lambda^3 \mathbb{E} \int_G e^{2\lambda(\psi(x) - c_1 T^2)} (|\nabla z(T)|^2 + |z(T)|^2) dx \\ & \leq C \lambda^5 e^{-2c_1 \lambda T^2} \mathbb{E} \int_Q e^{2\lambda\psi(x)} (|z|^2 + |\nabla z|^2 + |z_t|^2 + |\nabla^2 z|^2 + |\nabla z_t|^2) dx dt \\ & = C \lambda^5 e^{-2c_1 \lambda T^2} |z|_{\mathcal{H}_{\lambda, 0}}^2. \end{aligned} \quad (3.13)$$

Combining [\(3.11\)](#)–[\(3.13\)](#), for $\lambda \geq \lambda_0$, we arrive at

$$\begin{aligned} & c \mathbb{E} \int_Q \theta^2 (\lambda z_t^2 + \lambda |\nabla z|^2 + \lambda^3 z^2) dx dt + \kappa |z|_{\mathcal{H}_{\lambda, 0}}^2 - C \lambda^5 e^{-2c_1 \lambda T^2} |z|_{\mathcal{H}_{\lambda, 0}}^2 \\ & \leq C(|u_n - u^*|_{\mathcal{E}_{\lambda, c_1}}^2 + \kappa|u^*|_{\mathcal{H}_{\lambda, 0}}^2 + |\mathcal{PE}|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))}^2 + |\mathcal{E}|_{\mathcal{H}_{\lambda, 0}}^2). \end{aligned}$$

By selecting λ_1 sufficiently large such that $C \lambda^5 e^{-2c_1 \lambda T^2} \leq \kappa/2$ for every $\lambda \geq \lambda_1$ and recalling that $z = u_{n+1} - u^* - \mathcal{E}$, we deduce that for all $\lambda \geq \max\{\lambda_0, \lambda_1\}$,

$$\lambda |u_{n+1} - u^*|_{\mathcal{E}_{\lambda, c_1}}^2 \leq C(|u_n - u^*|_{\mathcal{E}_{\lambda, c_1}}^2 + \kappa|u^*|_{\mathcal{H}_{\lambda, 0}}^2 + |\mathcal{PE}|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0, T; L^2(G)))}^2 + \lambda |\mathcal{E}|_{\mathcal{H}_{\lambda, 0}}^2).$$

Next, setting $\delta < e^{-\lambda R_1^2/\gamma}$ for $x \in G$, we obtain the bound $e^{2\lambda\psi(x)} \leq e^{2\lambda R_1^2} \leq \delta^{-2\gamma}$. By [\(3.6\)](#), and choosing $\tilde{\lambda}_0 \geq \max\{C, \lambda_0, \lambda_1\}$, we find that for all $\lambda \geq \tilde{\lambda}_0$ and $\delta < e^{-\lambda R_1^2/\gamma}$, the following inequality holds:

$$|u_{n+1} - u^*|_{\mathcal{E}_{\lambda, c_1}}^2 \leq \left(\frac{C}{\lambda}\right) |u_n - u^*|_{\mathcal{E}_{\lambda, c_1}}^2 + C(\kappa|u^*|_{\mathcal{H}_{\lambda, 0}}^2 + \delta^{2-2\gamma}). \quad (3.14)$$

Choose $\tilde{\lambda}_0$ larger if necessary so that $q = C/\lambda \leq 1/2$. Iterating [\(3.14\)](#) and summing the geometric series gives the stated bound, with a factor at most $1/(1 - q) \leq 2$ in its error term. $\square$

4 Numerical experiments for Problem 1

In this section, we present some numerical tests to illustrate the performance of the proposed method. For simplicity, let the spatial domain be $G = (0, 1) \times (0, 1.5)$, the final time $T = 1$ and $b^{jk} = \delta_{jk}$ for $j, k = 1, 2$. Consider the following semilinear stochastic hyperbolic equation:

$$\begin{cases} du_t - \Delta u dt = F(u, u_t, \nabla u) dt + a u dW(t) & \text{in } Q, \\ u_t(0) = 0 & \text{in } G, \\ u(0) = u_0 & \text{in } G, \end{cases} \quad (4.1)$$

where $a \in L_{\mathbb{F}}^\infty(0, T; L^\infty(G))$.

We solve the forward problem by employing the finite difference method in space and the explicit

Euler-Maruyama method (see [10, 21, 28]) in time. Afterwards, we solve the inverse [Problem 1](#) by iteratively minimizing the Tikhonov functional.

Before presenting the numerical results, we first introduce the weight function θ in (2.3) as follows:

$$\theta(t, x, y) = e^{\lambda(\psi(x) - 0.25t^2)}, \quad \psi(x, y) = (x - 0.5)^2 + (y + 0.5)^2 - 4.175. \quad (4.2)$$

These weight parameters are empirical numerical choices; they are not asserted to satisfy the positivity and sufficient parameter bounds in [Conditions 1.2](#) and [1.3](#) or the large- λ threshold. The sign of the normal derivative still determines the observation boundary, so the observation boundary Γ_0 derived from (1.1) consists of:

$$\Gamma_0 = \{(1, y) \mid y \in [0, 1.5]\} \cup \{(x, 1.5) \mid x \in [0, 1]\} \cup \{(0, y) \mid y \in [0, 1.5]\}.$$

We shall employ the lateral Cauchy data $u|_{(0,T) \times \Gamma}$ and $\frac{\partial u}{\partial \nu} \Big|_{\Sigma_0}$ to reconstruct the unknown initial source function u_0 .

Remark 4.1. It is worth noting that, unlike in the deterministic case of [34], solving [Problem 1](#) does not require the boundary data for $\frac{\partial u}{\partial \nu}$ on the lower boundary $\{(x, 0) \mid x \in [0, 1]\}$, which is typically more difficult to obtain. We also employ a rectangular domain instead of a square one in [34], which usually poses difficulties in reconstructing the initial condition.

Now we present the computational algorithm employed in our numerical experiments. Let $J_{n,\ell}(u; u_n)$ denote the Tikhonov functional defined in (3.2), where the subscript ℓ indicates the dependence on sample paths ω^ℓ through the Itô integration term in operator $\mathcal{P}$. Let $\mathcal{A}_j$ be the affine assembly map for boundary-data sample j , defined below, and let q denote its free tensor. The Wiener increments indexed by ℓ are fixed throughout each outer sample. At update t use $j_t = 1 + ((t - 1) \bmod N_B)$, where N_B is the number of boundary-data samples, and differentiate $J_{n,\ell}(\mathcal{A}_{j_t} q; u_\ell^{n-1})$ with respect to q . The numerical inverse problem is approximated through the following iterative optimization procedure ([Algorithm 1](#)), where $\mathbf{g}_t^2$ represents the element-wise square of $\mathbf{g}_t$:

For fixed Wiener increments, boundary sample, and previous iterate, the discrete functional is a deterministic convex quadratic in the free tensor. Its normal equations and deterministic conjugate gradient method are therefore well defined. In the available implementation, the boundary sample is cycled at each update, so the affine assembly and objective change with j_t . We use Adam for this sample-cycling procedure and the differently scaled gradient components; the archived optimizer comparison does not establish an obstruction to CG for a fixed quadratic objective.

The Adam algorithm is specifically designed to address stochastic optimization problems. Its key innovation lies in the incorporation of a momentum mechanism:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \mathbf{g}_t.$$

This approach offers two significant benefits:

1. **Noise Reduction:** By computing an exponentially weighted average of past gradients, the algorithm effectively smooths out stochastic fluctuations in the gradient estimates.
2. **Stable Convergence:** The momentum term helps maintain consistent update directions, leading to more reliable convergence behavior compared to standard stochastic gradient methods.

On the other hand, our computations reveal a significant disparity in the magnitudes of different gradient components, due to the exponential form of the weight function θ defined in (4.2). This issue reduces the convergence rate of the CGM (see [Figure 1](#)). However, Adam mitigates the problem by adjusting the learning rate adaptively based on the historical gradient information for

Algorithm 1 The algorithm for solving the inverse problem

- 1: Choose maximum iteration N_I and the number of sample paths N_S .
 - 2: **for** $\ell = 1$ to N_S **do**
 - 3: Choose a free initial tensor q_ℓ^0 and set $u_\ell^0 = \mathcal{A}_1 q_\ell^0$.
 - 4: **for** $n = 1$ to N_I **do**
 - 5: Choose the maximum number of update steps N_U , the learning rate α , and the parameters β_1, β_2 , and ε .
 - 6: Initialize $q_{(0)} = q_\ell^{n-1}$, $v_0 = 0$, $m_0 = 0$.
 - 7: **for** $t = 1$ to N_U **do**
 - 8: Set $j_t = 1 + ((t - 1) \bmod N_B)$ and $\mathbf{g}_t = \nabla_q J_{n,\ell}(\mathcal{A}_{j_t} q_{(t-1)}; u_\ell^{n-1})$.
 - 9: Set $m_t = \beta_1 m_{t-1} + (1 - \beta_1) \mathbf{g}_t$ and $v_t = \beta_2 v_{t-1} + (1 - \beta_2) \mathbf{g}_t^2$.
 - 10: Set $\hat{m}_t = m_t / (1 - \beta_1^t)$ and $\hat{v}_t = v_t / (1 - \beta_2^t)$.
 - 11: Set $q_{(t)} = q_{(t-1)} - \alpha \hat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon)$.
 - 12: **end for**
 - 13: Set $q_\ell^n = q_{(N_U)}$ and $u_\ell^n = \mathcal{A}_1 q_\ell^n$ as the computed iterate; finite updates do not guarantee an exact minimizer.
 - 14: **end for**
 - 15: **end for**
 - 16: Set the compute solution $u_c = \frac{1}{N_S} \sum_{\ell=1}^{N_S} u_\ell^{N_I}$.
 - 17: **Output:** The reconstructed initial source function $u_0 = u_c(0, x, y)$ for $(x, y) \in G$.
-

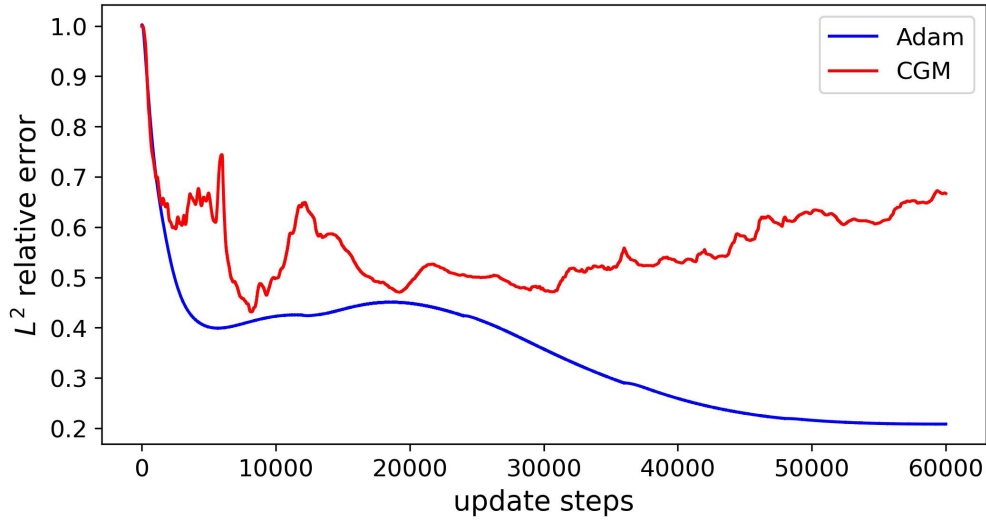


Figure 1: We adopt the setting of [Example 4.4](#) with $N_S = 1$. The figure records the implemented Adam and CGM runs with cycling boundary samples. It is an empirical comparison of those implementations, not a claim that CG fails for a fixed quadratic objective.

each parameter. More precisely, the learning rate is given by

$$\gamma_t = \frac{\gamma}{\sqrt{\hat{v}_t} + \varepsilon},$$

where $v_t = \beta_2 v_{t-1} + (1 - \beta_2) \mathbf{g}_t^2$ and $\hat{v}_t = v_t / (1 - \beta_2^t)$.

The next challenging task is to compute the gradient $\nabla_u J_{n,\ell}(u_{(t-1)}; u_\ell^{n-1})$. For stochastic problems, using the adjoint method typically requires solving an adjoint system, that is, a backward stochastic differential equation (see, e.g., [8]). Numerically, the evaluation of the correction term in a backward stochastic differential equation is very difficult because, to date, there is no effective method to compute the conditional expectation of a general random variable [32]. To overcome this issue, we employ the technique of automatic differentiation to compute the gradient.

One of the fundamental reasons behind the significant advances in machine learning is the underlying computational infrastructure centered on automatic differentiation. The basic idea of automatic differentiation is to decompose a function into a series of simple operations and then use the chain rule to compute derivatives. Automatic differentiation is an efficient method for evaluating the gradient of the complicated functional, with reverse-mode cost comparable to a small constant number of objective evaluations; the absolute cost still depends on the tensor size. Modern open-source machine learning libraries such as PyTorch [35] have already incorporated this method, and it has been successfully applied in areas such as computational fluid dynamics [1].

In this paper, the objective changes when the boundary-data sample is cycled (see Remark 4.4); after fixing all data it is deterministic. Furthermore, its parameter dimension is extremely high (approximately 80,000 parameters in our numerical examples, see (4.12)). Hence, automatic differentiation is employed to efficiently compute the gradient of $J_{n,\ell}(u; u_\ell^{n-1})$ for each sample path. This approach avoids the need for an explicit derivative calculation or the solution of a backward stochastic partial differential equation. Our subsequent numerical experiments demonstrate the efficiency and accuracy of this approach.

4.1 Generating the simulated data

Introduce the uniform mesh in $(0, T) \times G$ as

$$\mathfrak{M} = \{(t_k, x_i, y_j) \mid t_k = k\tau, x_i = ih_x, y_j = jh_y, k \in \llbracket 0, M \rrbracket, i \in \llbracket 0, N_x \rrbracket, j \in \llbracket 0, N_y \rrbracket\},$$

where $h_x = 1/N_x$, $h_y = 3/(2N_y)$ and $\tau = 1/M$. Here, we use the notation $\llbracket a, b \rrbracket = [a, b] \cap \mathbb{N}$. We choose $M = 65$ in Examples 4.1–4.2 and $M = 60$ in Examples 4.3–4.4, with $N_x = 32$ and $N_y = 48$. These values are chosen such that Courant–Friedrichs–Lewy condition $\tau(1/h_x^2 + 1/h_y^2)^{1/2} \simeq 0.696 < 1$ for $M = 65$ (and $\simeq 0.754 < 1$ for $M = 60$) is satisfied. Thus, the computed solution obtained by our scheme is stable.

Let $u_{k,i,j} = u(t_k, x_i, y_j)$ for $k \in \llbracket 0, M \rrbracket$, $i \in \llbracket 0, N_x \rrbracket$, $j \in \llbracket 0, N_y \rrbracket$. The equation (4.1) is solved by centred difference approximation for the Laplacian operator and explicit Euler-Maruyama method for the time in the examples below, whose nonlinearities do not depend on u_t . For a velocity-dependent nonlinearity, the forward difference in (4.4) would make (4.3) implicit; that general case is not asserted to be an explicit scheme here. For $k \in \llbracket 1, M - 1 \rrbracket$, $i \in \llbracket 1, N_x - 1 \rrbracket$, $j \in \llbracket 1, N_y - 1 \rrbracket$, we solve the following equation:

$$[du_t]_{k,i,j} = [u_{xx}]_{k,i,j}\tau + [u_{yy}]_{k,i,j}\tau + F(t_k, x_i, y_j, u_{k,i,j}, [u_t]_{k,i,j}, [\nabla u]_{k,i,j})\tau + a(t_k, x_i, y_j)u_{k,i,j}\Delta W_k, \quad (4.3)$$

where the increments $\Delta W_k = W_{k+1} - W_k = \sqrt{\tau} \xi_k$, with independent $\xi_k \sim N(0, 1)$, and

$$\begin{aligned} [du_t]_{k,i,j} &= \frac{u_{k+1,i,j} - 2u_{k,i,j} + u_{k-1,i,j}}{\tau}, & [u_{xx}]_{k,i,j} &= \frac{u_{k,i+1,j} - 2u_{k,i,j} + u_{k,i-1,j}}{h_x^2} \\ [u_{yy}]_{k,i,j} &= \frac{u_{k,i,j+1} - 2u_{k,i,j} + u_{k,i,j-1}}{h_y^2}, & [\nabla u]_{k,i,j} &= \left(\frac{u_{k,i+1,j} - u_{k,i,j}}{h_x}, \frac{u_{k,i,j+1} - u_{k,i,j}}{h_y} \right)^T, \\ [u_t]_{k,i,j} &= \frac{u_{k+1,i,j} - u_{k,i,j}}{\tau}. \end{aligned} \quad (4.4)$$

The initial value condition $u(0) = u_0$ and $u_t(0) = 0$ is given by

$$u_{0,i,j} = u_0(x_i, y_j), \quad u_{1,i,j} = u_{0,i,j}, \quad i \in \llbracket 0, N_x \rrbracket, \quad j \in \llbracket 0, N_y \rrbracket. \quad (4.5)$$

The boundary condition $u|_{(0,T) \times \Gamma} = f^*$ is given by

$$u_{k,i,j} = \begin{cases} f^*(t_k, 1, y_j), & k \in \llbracket 0, M \rrbracket, \quad i = N_x, \quad j \in \llbracket 0, N_y \rrbracket, \\ f^*(t_k, 0, y_j), & k \in \llbracket 0, M \rrbracket, \quad i = 0, \quad j \in \llbracket 0, N_y \rrbracket, \\ f^*(t_k, x_i, 0), & k \in \llbracket 0, M \rrbracket, \quad i \in \llbracket 0, N_x \rrbracket, \quad j = 0, \\ f^*(t_k, x_i, 3/2), & k \in \llbracket 0, M \rrbracket, \quad i \in \llbracket 0, N_x \rrbracket, \quad j = N_y. \end{cases} \quad (4.6)$$

After solving (4.3), the boundary data for $\frac{\partial u}{\partial \nu} = g^*$ on $(0, T) \times \Gamma_0$ is computed using

$$g^*(t_k, x_i, y_j) = \begin{cases} \frac{u_{k,i,j} - u_{k,i-1,j}}{h_x}, & k \in \llbracket 0, M \rrbracket, \quad i = N_x, \quad j \in \llbracket 0, N_y \rrbracket, \\ \frac{u_{k,i,j} - u_{k,i+1,j}}{h_x}, & k \in \llbracket 0, M \rrbracket, \quad i = 0, \quad j \in \llbracket 0, N_y \rrbracket, \\ \frac{u_{k,i,j} - u_{k,i,j-1}}{h_y}, & k \in \llbracket 0, M \rrbracket, \quad i \in \llbracket 0, N_x \rrbracket, \quad j = N_y. \end{cases} \quad (4.7)$$

Recalling Remark 4.1, we do not need to compute the boundary data for $\frac{\partial u}{\partial \nu}$ on the lower boundary $\{(x, 0) \mid x \in [0, 1]\}$.

Remark 4.2. Note that f^* and g^* actually depend on the sample point ω . In the sequel, a total of 8 sample paths for the forward equations were generated for each test.

4.2 Specifying the Tikhonov functional

To solve the inverse problem, we employ the Tikhonov regularization approach using the functional defined in (3.2). The numerical implementation requires the following key components:

1. Discretization Scheme: The optimization variable is discretized as a three-dimensional tensor:

$$(u_{k,i,j})_{k \in \llbracket 0, M \rrbracket, i \in \llbracket 0, N_x \rrbracket, j \in \llbracket 0, N_y \rrbracket}, \quad (4.8)$$

representing the solution across temporal and spatial dimensions.

2. Noise Model: We introduce 10% multiplicative noise to the lateral Cauchy data:

$$\begin{aligned} f(t, x, y) &= f^*(t, x, y)(1 + \delta \xi(t, x, y)), \\ g(t, x, y) &= g^*(t, x, y)(1 + \delta \xi(t, x, y)), \end{aligned} \quad (4.9)$$

where $\xi(t, x, y)$ is a random variable uniformly distributed in $[-1, 1]$, which models uniformly dis-

tributed measurement noise.

3. Discrete Assembly Space: For each measured-data sample b , define the affine map $\mathcal{A}_b$ as follows. Its free entries have the indices in (4.12). Copy their first time layer to $k = 0$. On the two x boundaries and their adjacent strips ($i = 0, 1, N_x - 1, N_x$), insert the measured Dirichlet values and the one-sided Neumann reconstruction from (4.6) and (4.7), over the entire y range. For $2 \leq i \leq N_x - 2$, insert the y -boundary values at $j = 0, N_y$ and the upper Neumann strip at $j = N_y - 1$. This gives the x strips priority at intersections; no lower Neumann strip is imposed. The assembly space used for that data sample is

$$\mathcal{U}_D^{(b)} = \{\mathcal{A}_b q : q \text{ is a free tensor with the indices in (4.12)}\}.$$

The equality $u_{1,i,j} = u_{0,i,j}$ is imposed only on the free spatial nodes. Boundary and adjacent-strip values retain their separately measured time layers. Thus this space does not impose every noisy face constraint at intersections, or zero initial velocity on measured strips; those constraints are generally incompatible for independently perturbed data. Below $\mathcal{U}_D$ denotes the assembly space for the current data sample.

Remark 4.3. Note that f^* and g^* depend on the sample point ω of the forward equation (4.3). In step 8 of Algorithm 1, inspired by stochastic gradient descent [36], we compute the gradient corresponding to the Cauchy data f^* and g^* for each sample ω of (4.3) in turn. For brevity, we do not distinguish between the f^* and g^* generated by different sample paths here.

To compute the Tikhonov functional (3.2), we first discretize the operator $\mathcal{P}$ appearing in (3.2). Observe that if φ satisfies

$$d\varphi_t - \Delta\varphi dt = \Phi dt + a\varphi dW(t),$$

then it follows that

$$\begin{aligned} & |\mathcal{P}\varphi - \mathbf{LF}(u_n)|_{L_{\mathbb{F}}^{2,w}(\Omega; H^1(0,T; L^2(G)))}^2 \\ &= \mathbb{E} \int_0^T \int_G \theta^2 \left[\int_0^t (\Phi - \mathcal{F}(u_n)) ds \right]^2 dx dt + \mathbb{E} \int_0^T \int_G \theta^2 |\Phi - \mathcal{F}(u_n)|^2 dx dt. \end{aligned}$$

Motivated by this observation, we define a new Tikhonov functional

$$\hat{J}_n(\varphi; u_n) = \mathbb{E} \int_0^T \int_G \theta^2 |\Phi - \mathcal{F}(u_n)|^2 dx dt + \kappa |\varphi|_{\mathcal{H}_{\lambda,0}^2}^2. \quad (4.10)$$

By the arguments in the proof of Theorem 3.2, the minimizer of $\hat{J}_n(\varphi; u_n)$ also satisfies the convergence result in Theorem 3.2. Hence, in the sequel, we consider the discrete form of (4.10) for $u = (u_{k,i,j}) \in \mathcal{U}_D$:

$$\begin{aligned} J_{n,\ell}(u; u_\ell^{n-1}) &= \sum_{k=1}^{M-1} \sum_{i=1}^{N_x-1} \sum_{j=1}^{N_y-1} \left\{ \theta_{k,i,j}^2 (\Phi_{k,i,j} - [\mathcal{F}(u_\ell^{n-1})]_{k,i,j})^2 \right. \\ &\quad + \kappa [\theta_0]_{k,i,j}^2 (u_{k,i,j}^2 + [u_t]_{k,i,j}^2 + [\nabla u]_{k,i,j}^2 + [u_{xx}]_{k,i,j}^2 + [u_{yy}]_{k,i,j}^2 \\ &\quad \left. + [u_{xy}]_{k,i,j}^2 + [u_{tx}]_{k,i,j}^2 + [u_{ty}]_{k,i,j}^2) \right\} \tau h_x h_y, \end{aligned} \quad (4.11)$$

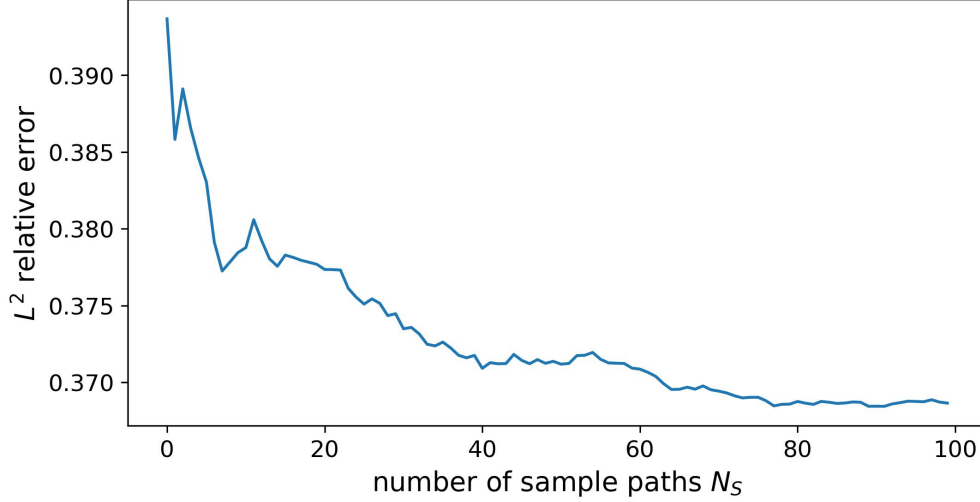


Figure 2: We adopt the setting of [Example 4.1](#) with $N_U = 3000$, $N_I = 3$, and $N_S = 100$. The figure shows how the L^2 relative error $\frac{|u_0 - u_0^*|_{L^2(G)}}{|u_0^*|_{L^2(G)}}$ decreases with the number of sample paths N_S . In subsequent numerical tests, we set $N_S = 30$.

where

$$\begin{aligned}
\theta_{k,i,j} &= e^{\lambda(\psi(x_i, y_j) - c_1 t_k^2)}, & [\theta_0]_{k,i,j} &= e^{\lambda\psi(x_i, y_j)}, \\
\Phi_{k,i,j} &= \{[du_t]_{k,i,j} - ([u_{xx}]_{k,i,j} + [u_{yy}]_{k,i,j})\tau - a_{k,i,j}u_{k,i,j}\Delta W_k\}/\tau, \\
[\mathcal{F}(u_\ell^{n-1})]_{k,i,j} &= F(t_k, x_i, y_j, [u_\ell^{n-1}]_{k,i,j}, [\partial_t u_\ell^{n-1}]_{k,i,j}, [\nabla u_\ell^{n-1}]_{k,i,j}), \\
[u_{xy}]_{k,i,j} &= \frac{u_{k,i+1,j+1} - u_{k,i+1,j} - u_{k,i,j+1} + u_{k,i,j}}{h_x h_y}, \\
[u_{tx}]_{k,i,j} &= \frac{u_{k+1,i+1,j} - u_{k+1,i,j} - u_{k,i+1,j} + u_{k,i,j}}{h_x \tau}, \\
[u_{ty}]_{k,i,j} &= \frac{u_{k+1,i,j+1} - u_{k+1,i,j} - u_{k,i,j+1} + u_{k,i,j}}{h_y \tau}.
\end{aligned}$$

Other notations are defined in [\(4.4\)](#).

Using the assembly map above at each update, we optimize the functional [\(4.11\)](#) with respect to the tensor

$$(u_{k,i,j})_{k \in \llbracket 1, M \rrbracket, i \in \llbracket 2, N_x - 2 \rrbracket, j \in \llbracket 1, N_y - 2 \rrbracket}. \quad (4.12)$$

At free spatial nodes only, the unknown initial layer is tied to the first free time layer by $u_{0,i,j} = u_{1,i,j}$; it is not fixed to the unknown true source. The extra free row $j = 1$ reflects the absence of lower-boundary Neumann data. We employ automatic differentiation techniques [\[35\]](#) to directly compute the gradient of the discrete functional [\(4.11\)](#) with respect to the tensor [\(4.12\)](#). The recorded implementation performs $N_U = 12000$ optimization updates in approximately 15 seconds on an Apple M3 Pro; this runtime does not certify attainment of the exact minimum.

Remark 4.4. The Wiener increments ΔW_k are sampled once for each outer sample and then held fixed. Boundary-data samples are cycled during the inner updates. Conditional on all sampled quantities and the previous iterate, each evaluated objective is a deterministic quadratic. In

Algorithm 1, we perform $N_S = 30$ samples (see Figure 2).

4.3 Numerical examples

Remark 4.5 (Scope of the archived numerical illustrations). The forward formulas below have been checked against the saved arrays; the optimization runs have not been repeated and their historical source versions are not fully recorded. The saved data for Examples 4.1–4.2 correspond to $a = 1$. In Examples 4.3–4.4, the forward data use $a = xyt^2$ in the enlarged physical domain, which becomes $(x + 2)(y + 3)t^2$ after restriction and translation. In the available training entry points, the case index is not passed to the loss function, so the inverse residual uses its defaults $F_{\text{inv}}(u, p) = \sqrt{1 + u^2} + |p|$ and $a_{\text{inv}} = 1$. The truncated-exponential and xyt^2 branches exist in the code but are not selected by those calls. Since the historical training versions were not frozen, the archived results in Examples 4.2–4.4 are numerical illustrations with unresolved historical model matching, not validation of the stated same-model inverse problem. The pathwise optimization also does not establish equivalence with the adapted variational minimization in Theorem 3.2.

For all numerical experiments presented in this work, we employ the following set of optimized hyperparameter: $\lambda = 0.2$, $c_1 = 0.25$, $\kappa = 10^{-4}$, $N_S = 30$, $N_I = 5$, $N_U = 12000$, $\alpha = 0.01$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\varepsilon = 10^{-8}$.

For the numerical experiments, the free entries are initialized to zero and assembled with the first boundary-data sample. Only the free first layer is copied to $t = 0$; measured boundary strips keep their own time layers. The full initial iterate belongs to this assembly space and is generally nonzero.

Write $\bar{u}^n = N_S^{-1} \sum_{\ell=1}^{N_S} u_\ell^n$, and let $\|\cdot\|_h$ denote the discrete spatial L^2 norm on the uniform grid. The final reported relative error is the error of the averaged reconstruction,

$$E_{\text{final}} = \frac{\|\bar{u}^{N_I}(0) - u_0^*\|_h}{\|u_0^*\|_h}.$$

By contrast, panels (d) average the individual relative errors over samples at each update step and show their minimum–maximum range; their mean is $E_{\text{paths}} = N_S^{-1} \sum_{\ell=1}^{N_S} \|u_\ell(0) - u_0^*\|_h / \|u_0^*\|_h$. These two statistics need not agree. The archived panels (e) use the signed diagnostic

$$D_n = \max_{i,j} [\bar{u}^{n+1}(0, x_i, y_j) - \bar{u}^n(0, x_i, y_j)],$$

computed after averaging the samples and without taking absolute values. This is not a norm of the consecutive difference.

Example 4.1. Let $F(u, u_t, \nabla u) = \sqrt{1 + u^2} + |\nabla u|$, $a = 1$ and the true source function

$$u_0^* = \begin{cases} 10e^{\frac{r(x,y)^2}{r(x,y)^2-1}}, & r(x, y) < 1, \\ 0, & r(x, y) \geq 1, \end{cases}$$

where

$$r(x, y) = 16 \left[\left(x - \frac{1}{2} \right)^2 + \frac{1}{2}(y - 1)^2 \right].$$

The boundary conditions for the forward equation are given by $f^* = u_0^*|_\Gamma$. The lateral Cauchy data, namely f and g , are obtained from (4.6), (4.7) and (4.9). The numerical result is presented in Figure 3.

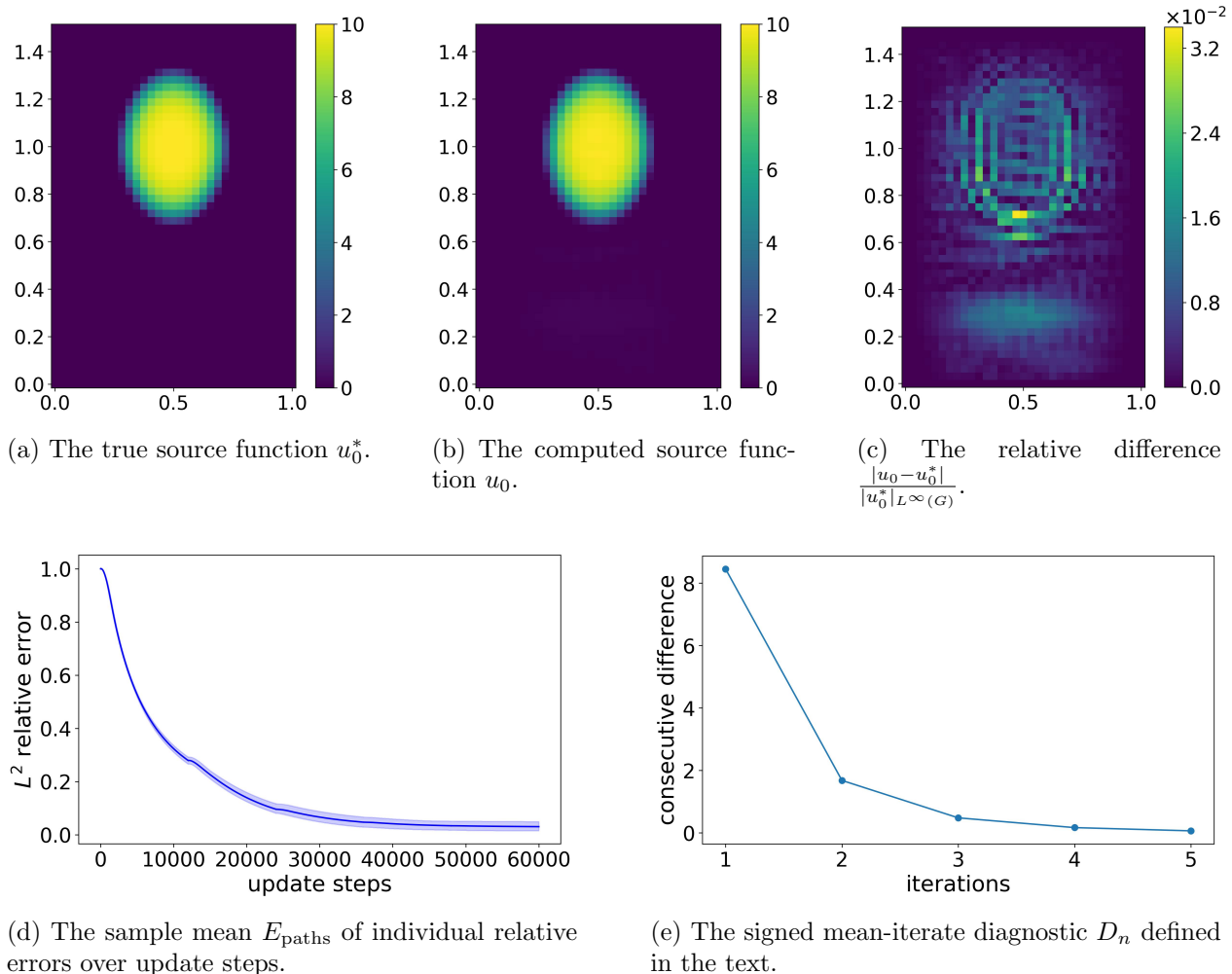


Figure 3: The numerical results for [Example 4.1](#). The computed source function u_0 is very close to the true source function u_0^* . The relative difference is small, and the L^2 relative error is 2.1%.

From [Figures 3a](#) and [3b](#), it is evident that the source function u_0^* is very well recovered, and from [Figure 3c](#) one can observe that the relatively large differences occur along the boundary of the initial condition. The final L^2 relative error is 2.1%. From [Figure 3d](#), it can be seen that the L^2 relative error remains largely unchanged for different sample paths. Panel (e) records a decrease of the signed diagnostic D_n . Since this quantity is not a norm and permits cancellation across samples, its decrease alone does not establish stabilization or validate [Theorem 3.2](#).

Example 4.2. Let $F(u, u_t, \nabla u) = \min\{e^u + |\nabla u|, 10\}$, $a = 1$ and the true source function

$$u_0^* = \sin(2\pi(x + y)) + \sin(4\pi(x - y))$$

The boundary conditions for the forward equation are given by $f^* = u_0^*|_{\Gamma}$. The lateral Cauchy data f and g are obtained from [\(4.6\)](#), [\(4.7\)](#) and [\(4.9\)](#). The numerical result is presented in [Figure 4](#).

This nonlinearity is globally Lipschitz: $|F(u, p) - F(v, q)| \leq 10|u - v| + |p - q|$. The truncation caps the exponential derivative. [Figures 4a](#) and [4b](#) show that the source function u_0^* is accurately recovered, with a final L^2 relative error of 6%. The forward nonlinearity lies in the globally Lipschitz class of [Condition 1.1](#); the scope remark above explains why this archived reconstruction does not

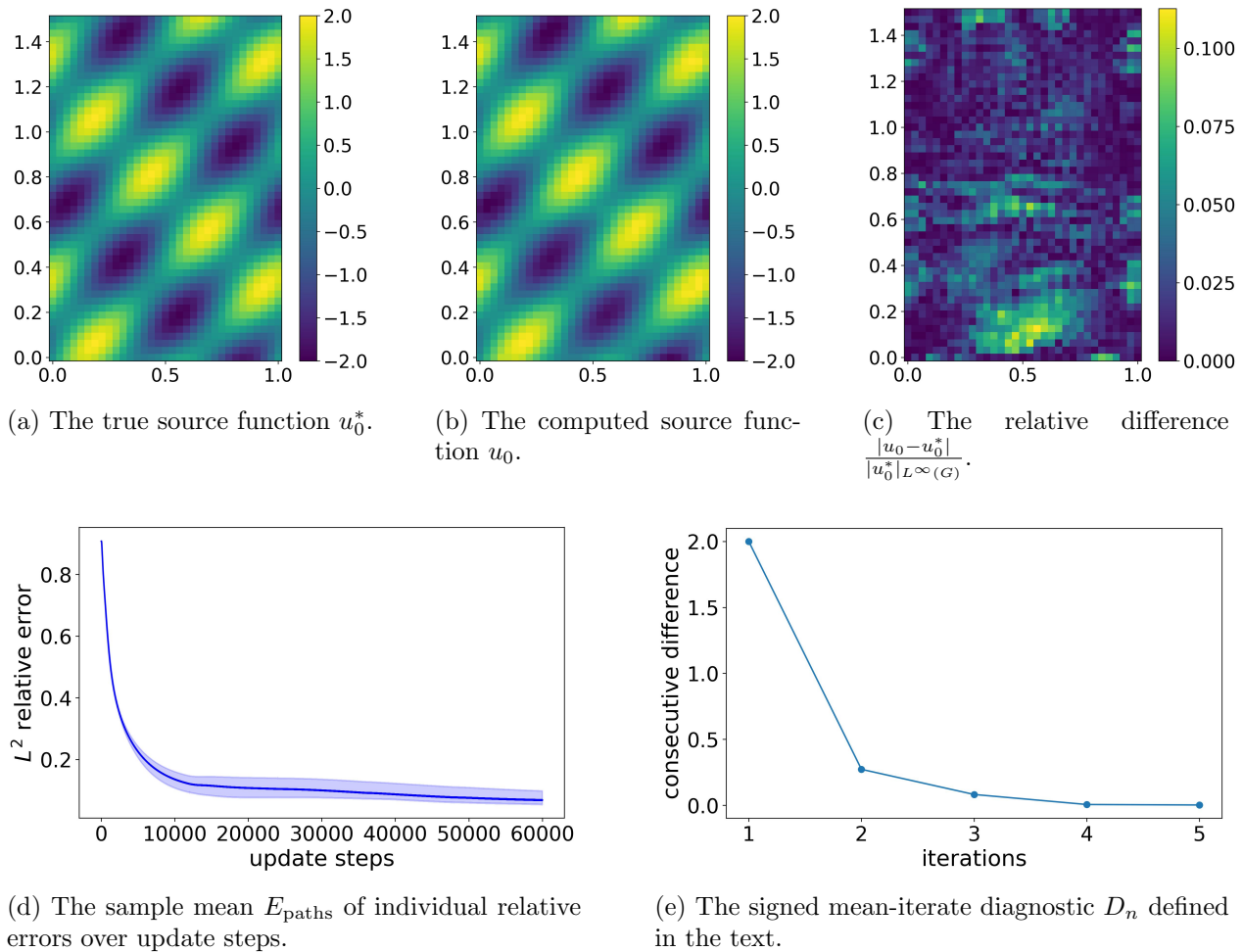


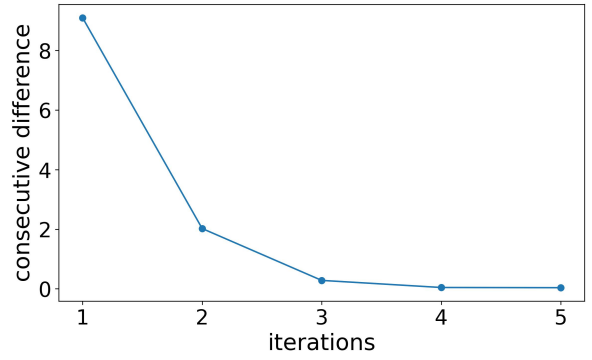
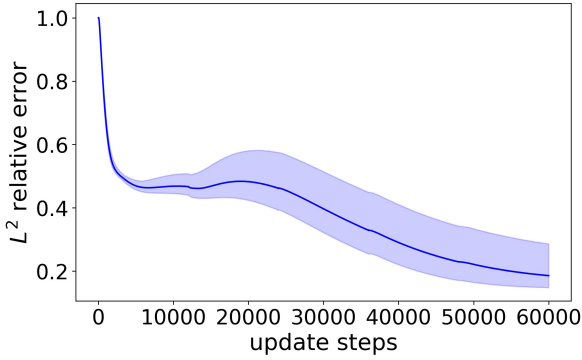
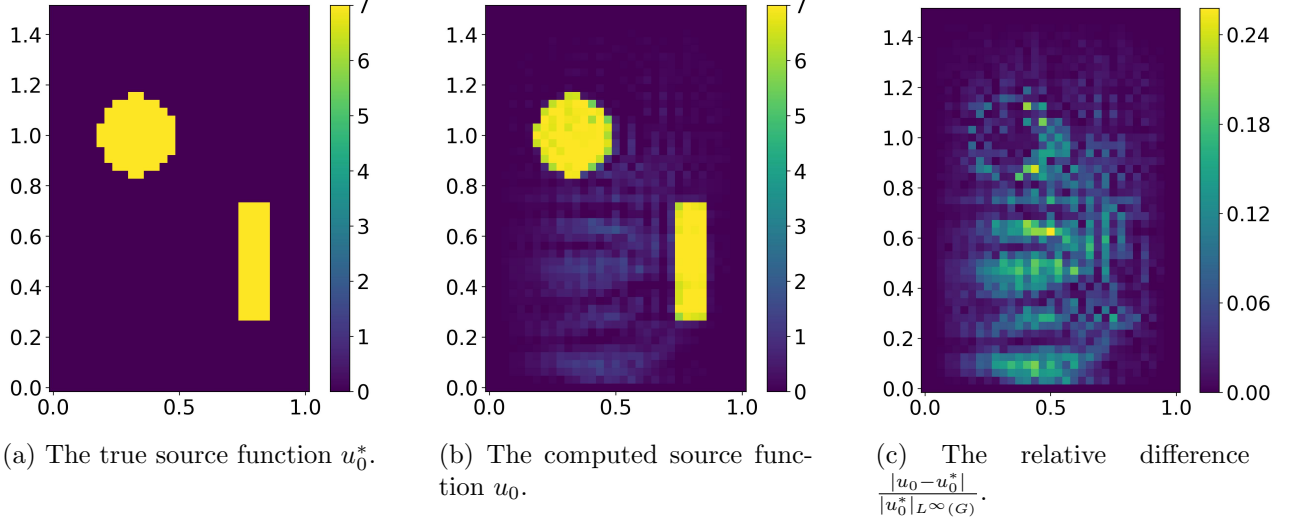
Figure 4: The numerical results for [Example 4.2](#). The computed source function u_0 approximates the true source function u_0^* well in this archived run, whose forward data use the truncated exponential nonlinearity F . The relative difference is small, and the L^2 relative error is 6%.

verify the corresponding same-model theorem. Moreover, as illustrated in [Figure 4d](#), the L^2 relative error stays largely consistent across different sample paths.

Example 4.3. Let $F(u, u_t, \nabla u) = \sqrt{1 + u^2} + |\nabla u|$, $a = xyt^2$ in the physical coordinates of the forward domain and

$$\tilde{u}_0^* = \begin{cases} 7, & \left(x - \frac{7}{3}\right)^2 + (y - 4)^2 < \frac{1}{40} \text{ or } \max \left\{ 16 \left|x - \frac{14}{5}\right|, 4 \left|y - \frac{7}{2}\right| \right\} < 1 \\ 0, & \text{others.} \end{cases}$$

Letting $\tilde{G} = (0, 5) \times (0, 7.5)$, $\tilde{N}_x = 160$ and $\tilde{N}_y = 240$, we solve [\(4.3\)](#) in $\tilde{G}$. The boundary conditions for the forward equation are given by $\tilde{f}^* = \tilde{u}_0^*|_{\partial\tilde{G}}$. The true source function is given by $\tilde{u}_0^*$ restricted to $[2, 3] \times [3, 4.5]$ and translated by $(x, y) \mapsto (x - 2, y - 3)$. The lateral Cauchy data f and g are also obtained from the restricted true solution function $\tilde{u}^*$. In this case, both f and g are stochastic processes as well, which further complicates the problem. The archived numerical demonstration, with the model-mismatch qualification stated above, is presented in [Figure 5](#).



(d) The sample mean E_{paths} of individual relative errors over update steps.

(e) The signed mean-iterate diagnostic D_n defined in the text.

Figure 5: The numerical results for [Example 4.3](#). The computed source function u_0 closely approximates the true source function u_0^* . Despite both f and g being stochastic processes, which further complicates the problem, the relative difference remains minimal and the L^2 relative error is 17.9%.

[Figures 5a](#) and [5b](#) show that the source function u_0^* is accurately recovered, with a final L^2 relative error of 17.9%. Since both f and g are stochastic processes in this case, additional randomness is introduced into the problem. As can be seen from [Figure 5d](#), although the L^2 relative errors vary for different sample paths, the averaged error remains stable.

Example 4.4. Let $F(u, u_t, \nabla u) = \sqrt{1 + u^2} + |\nabla u|$, $a = xyt^2$ in the physical coordinates of the forward domain and

$$\tilde{u}_0^* = \begin{cases} 7, & \left(x - \frac{5}{2}\right)^2 + \left(y - \frac{7}{2}\right)^2 < \frac{1}{8} \text{ and } \max \left\{ \left|x - \frac{5}{2}\right|, \left|y - \frac{7}{2}\right| \right\} > \frac{1}{5} \\ 0, & \text{others.} \end{cases}$$

Letting $\tilde{G} = (0, 5) \times (0, 7.5)$, $\tilde{N}_x = 160$ and $\tilde{N}_y = 240$, we solve [\(4.3\)](#) in $\tilde{G}$. The boundary conditions for the forward equation are given by $\tilde{f}^* = \tilde{u}_0^*|_{\partial\tilde{G}}$. The true source function is given by $\tilde{u}_0^*$ restricted to $[2, 3] \times [3, 4.5]$ and translated by $(x, y) \mapsto (x - 2, y - 3)$. The lateral Cauchy data f and g are also

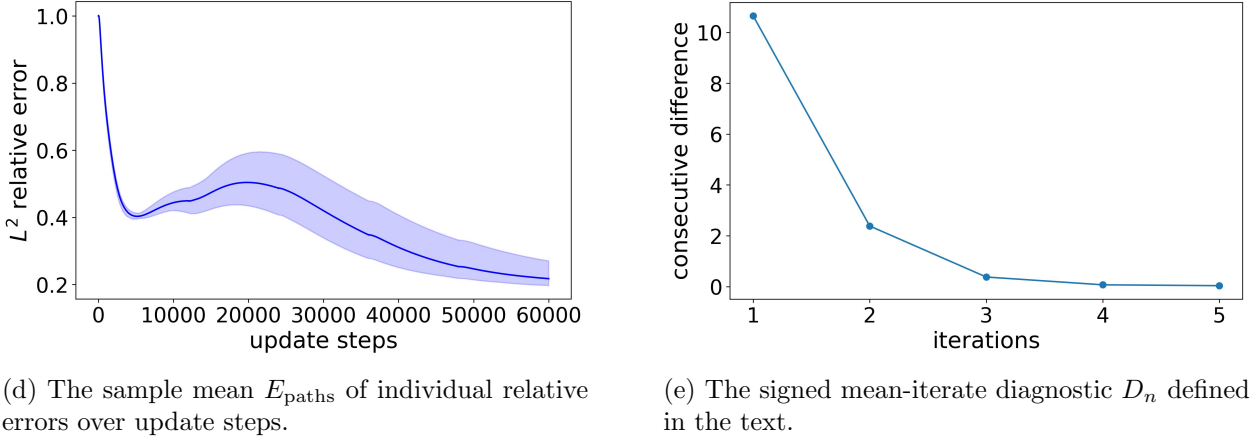
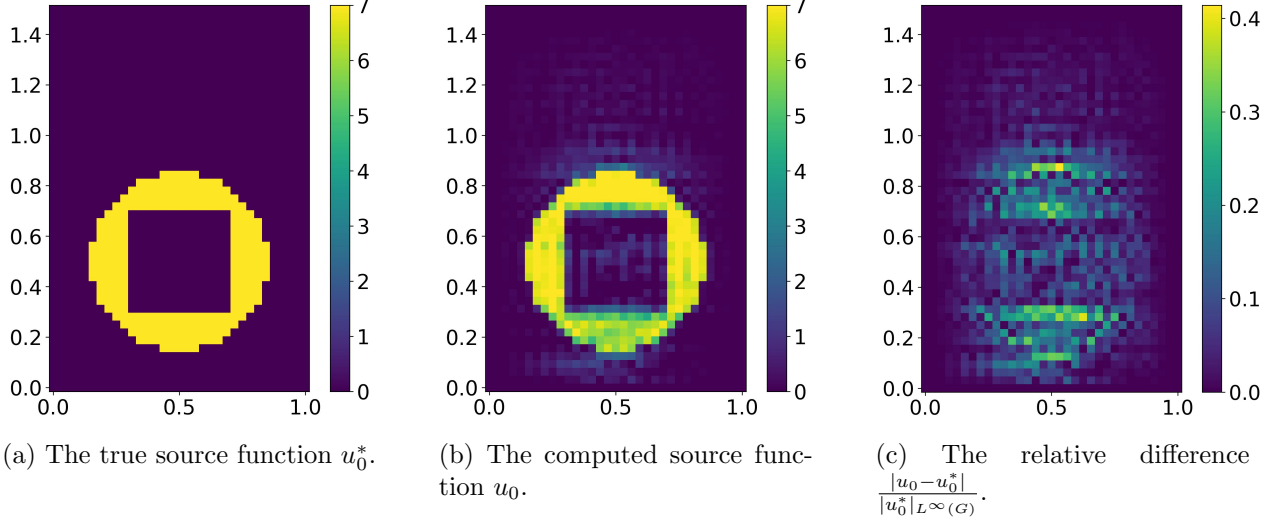


Figure 6: The numerical results for [Example 4.4](#). The computed source function u_0 closely approximates the true source function u_0^* . In this case, both f and g are stochastic processes, which further complicates the problem. The relative difference remains minimal and the L^2 relative error is 20.7%.

obtained from the restricted true solution function $\tilde{u}^*$. In this case, both f and g are stochastic processes as well, which further complicate the problem. The archived numerical demonstration, with the model-mismatch qualification stated above, is presented in [Figure 6](#).

[Figures 6a](#) and [6b](#) show that the source function u_0^* is accurately recovered, with a final L^2 relative error of 20.7%. Although the stochastic nature of f and g introduces additional randomness, and the absence of normal derivative information on the lower boundary (with final time $T = 1$) makes the reconstruction in the lower portion less accurate than in the upper part, the overall recovery remains robust. Moreover, as seen in [Figure 6c](#), the regions with relatively larger error occur at the boundary of the source function, while errors elsewhere are very small. Finally, [Figure 6d](#) indicates that although the L^2 relative errors vary across different sample paths, the averaged error remains stable.

5 Concluding remarks

In this paper, we have proposed a new approach for solving inverse source problems for semi-linear stochastic hyperbolic equations. Our method combines fixed-point iteration and Carleman estimates to obtain a regularized error bound from any admissible initial guess. To address the challenges arose by the stochastic nature of the equations, such as the non-differentiability of Brownian motion sample paths and the lack of compactness in the solution space, we redesigned the Tikhonov functional and developed a new Carleman estimate suitable for stochastic hyperbolic equations.

The theoretical error bound in [Theorem 3.2](#) contains a Hölder-type noise term and a regularization bias. With fixed κ , it does not assert convergence to the exact source. The archived numerical illustrations show reconstruction behavior with partial boundary observations, subject to the model and implementation qualifications in the numerical section. [Figure 1](#) compares the recorded optimizers under the implemented sample cycling. It gives no general conclusion about conjugate gradient for the fixed deterministic quadratic objective.

The results of this paper have potential applications in various fields where wave phenomena are affected by randomness and uncertainty perturbations. For instance, in ocean wave forecasting, one can reconstruct the wave field from ship motion and sea surface observation data; see [\[39\]](#). In astrophysics and helioseismology, as motivated by the SOHO project [\[13\]](#), one can reconstruct the internal structure of stars or planets from surface observations under random perturbations. These examples illustrate the potential applicability of our method for inverse source problems in stochastic hyperbolic equations across various scientific and engineering disciplines.

Future work may include extending the approach to more general classes of stochastic partial differential equations and improving computational efficiency.

References

- [1] M. G. ALHASHIM, K. HAUSKNECHT AND M. P. BRENNER, *Control of flow behavior in complex fluids using automatic differentiation*, Proc. Natl. Acad. Sci., 122 (2025).
- [2] G. BAO, C. CHEN, AND P. LI, *Inverse random source scattering for elastic waves*, SIAM J. Numer. Anal., 55 (2017), pp. 2616–2643.
- [3] M. I. BELISHEV, *Recent progress in the boundary control method*, Inverse Problems, 23 (2007), pp. R1–R67.
- [4] M. BELLASSOUED AND M. YAMAMOTO, *Carleman estimates and applications to inverse problems for hyperbolic systems*, Springer, Tokyo, 2017.
- [5] E. M. CABANA, *The vibrating string forced by white noise*, Z. Wahrscheinlichkeitstheor. Verw. Geb., 15 (1970), pp. 111–130.
- [6] R. DALANG, D. KHOSHNEVISAN, C. MUELLER, D. NUALART, AND Y. XIAO, *A Minicourse on Stochastic Partial Differential Equations*, Springer Berlin Heidelberg, 2009.
- [7] T. D. DANG, L. H. NGUYEN AND H. T. T. VU, *Determining initial conditions for nonlinear hyperbolic equations with time dimensional reduction and the carleman contraction principle*, Inverse Problems, 40 (2024), p. 125021.
- [8] F. DOU AND W. DU, *Determination of the solution of a stochastic parabolic equation by the terminal value*, Inverse Problems, 38 (2022), pp. Paper No. 075010, 24.

- [9] F. DOU AND P. LÜ, *An inverse Cauchy problem of a stochastic hyperbolic equation*, Inverse Problems, 41 (2025), pp. Paper No. 045003, 28.
- [10] M. ERRAOUI, Y. OUKNINE, AND D. NUALART, *Hyperbolic stochastic partial differential equations with additive fractional Brownian sheet*, Stoch. Dyn., 3 (2003), pp. 121–139.
- [11] X. FENG, M. ZHAO, P. LI AND X. WANG, *An inverse source problem for the stochastic wave equation*, Inverse Problems Imaging 16 (2022), pp. 397–415.
- [12] T. FUNAKI, *Random motion of strings and related stochastic evolution equations*, Nagoya Math. J., 89 (1983), pp. 129–193.
- [13] A. HELLEMANS, *Soho probes sun’s interior by tuning in to its vibrations*, Science, 272 (1996), pp. 1264–1265.
- [14] L. HÖRMANDER, *The analysis of linear partial differential operators. IV*, Springer-Verlag, Berlin, 2009.
- [15] D. P. KINGMA AND J. BA, *Adam: A method for stochastic optimization*, in 3rd International Conference on Learning Representations (ICLR), 2015.
- [16] M. V. KLIBANOV, *Carleman estimates for global uniqueness, stability and numerical methods for coefficient inverse problems*, J. Inverse Ill-Posed Probl., 21 (2013), pp. 477–560.
- [17] ———, *Carleman weight functions for solving ill-posed Cauchy problems for quasilinear PDEs*, Inverse Problems, 31 (2015), pp. 125007, 20.
- [18] M. V. KLIBANOV AND O. V. IOUSSOUPOVA, *Uniform strict convexity of a cost functional for three-dimensional inverse scattering problem*, SIAM J. Math. Anal., 26 (1995), pp. 147–179.
- [19] M. V. KLIBANOV, N. A. KOSHEV, J. LI AND A. G. YAGOLA, *Numerical solution of an ill-posed Cauchy problem for a quasilinear parabolic equation using a Carleman weight function*, J. Inverse Ill-Posed Probl., 24 (2016), pp. 761–776.
- [20] M. V. KLIBANOV AND J. LI, *Inverse problems and Carleman estimates. Global uniqueness, global convergence and experimental data*, vol. 63 of Inverse Ill-Posed Probl. Ser., Berlin: De Gruyter, 2021.
- [21] P. E. KLOEDEN AND E. PLATEN, *Numerical solution of stochastic differential equations*, vol. 23 of Applications of Mathematics (New York), Springer-Verlag, Berlin, 1992.
- [22] M. LASSAS, M. SALO AND G. UHLMANN, *Wave phenomena*, in Handbook of mathematical methods in imaging. In 3 volumes, New York, NY: Springer, 2015, pp. 1205–1252.
- [23] J. LI, T. HELIN, AND P. LI, *Inverse random source problems for time-harmonic acoustic and elastic waves*, Commun. Partial Differ. Equations, 45 (2020), pp. 1335–1380.
- [24] J. LI AND P. LI, *Inverse elastic scattering for a random source*, SIAM J. Math. Anal., 51 (2019), pp. 4570–4603.
- [25] J. LI, H. LIU, AND S. MA, *Determining a random Schrödinger operator: both potential and source are random*, Commun. Math. Phys., 381 (2021), pp. 527–556.

- [26] P. LI AND Y. LIANG, *Stability for inverse source problems of the stochastic Helmholtz equation with a white noise*, SIAM J. Appl. Math., 84 (2024), pp. 687–709.
- [27] P. LI AND X. WANG, *Inverse random source scattering for the Helmholtz equation with attenuation*, SIAM J. Appl. Math., 81 (2021), pp. 485–506.
- [28] G. J. LORD, C. E. POWELL, AND T. SHARDLOW, *An introduction to computational stochastic PDEs*, Cambridge University Press, New York, 2014.
- [29] Q. LÜ, *Observability estimate and state observation problems for stochastic hyperbolic equations*, Inverse Problems, 29 (2013), p. 095011.
- [30] Q. LÜ AND X. ZHANG, *Global uniqueness for an inverse stochastic hyperbolic problem with three unknowns*, Comm. Pure Appl. Math., 68 (2015), pp. 948–963.
- [31] Q. LÜ AND X. ZHANG, *Mathematical Control Theory for Stochastic Partial Differential Equations*, Springer, Cham, 2021.
- [32] Q. LÜ, P. WANG, Y. WANG AND X. ZHANG, *Numerics for stochastic distributed parameter control systems: a finite transposition method*, in Numerical Control: Part A, E. Trélat and E. Zuazua, eds., vol. 23 of Handbook of Numerical Analysis, Elsevier, 2022, pp. 201–232.
- [33] L. H. NGUYEN, *An inverse space-dependent source problem for hyperbolic equations and the Lipschitz-like convergence of the quasi-reversibility method*, Inverse Problems, 35 (2019), pp. 035007, 28.
- [34] L. H. NGUYEN AND M. V. KLIBANOV, *Carleman estimates and the contraction principle for an inverse source problem for nonlinear hyperbolic equations*, Inverse Problems, 38 (2022), pp. Paper No. 035009, 19.
- [35] A. PASZKE, S. GROSS, S. CHINTALA, G. CHANAN, E. YANG, Z. DEVITO, Z. LIN, A. DESMAISON, L. ANTIGA AND A. LERER, *Automatic differentiation in pytorch*, in NIPS 2017 Workshop on Autodiff, 2017.
- [36] H. ROBBINS AND S. MONRO, *A stochastic approximation method*, Ann. Math. Statistics, 22 (1951), pp. 400–407.
- [37] W. W. SYMES, *The seismic reflection inverse problem*, Inverse Problems, 25 (2009), pp. 123008, 39.
- [38] B. A. TON, *An inverse problem for a nonlinear stochastic wave equation*, Stochastic Anal. Appl., 22 (2004), pp. 19–42.
- [39] G. WANG AND Y. PAN, *Phase-resolved ocean wave forecast with ensemble-based data assimilation*, J. Fluid Mech., 918 (2021), pp. Paper No. A19, 27.
- [40] Z. WANG, W. ZHU, B. WU AND B. HU, *A backward problem for stochastic Kuramoto-Sivashinsky equation: conditional stability and numerical solution*, J. Math. Anal. Appl., 543 (2025), pp. Paper No. 128988, 18.
- [41] G. YUAN, *Determination of two kinds of sources simultaneously for a stochastic wave equation*, Inverse Problems, 31 (2015), pp. 085003, 13.
- [42] X. ZHANG, *Carleman and observability estimates for stochastic wave equations*, SIAM J. Math. Anal., 40 (2008), pp. 851–868.