Iranian Journal of Numerical Analysis and Optimization
Vol. 14, No. 4, 2024, pp 1037–1068
https://doi.org/10.22067/ijnao.2024.86883.1392
https://ijnao.um.ac.ir/

Research Article

A space-time least-squares support vector regression scheme for inverse source problem of the time-fractional wave equation

A. Mohammadi and A. Tari Marzabad*

Abstract

The inverse problems in various fields of applied sciences and industrial design are concerned with the estimation of parameters that cannot be directly measured. In this work, we present a novel numerical approach for addressing the fractional inverse source problem by a machine learning algorithm and considering the ideas behind the spectral methods. The introduced algorithm utilizes a space-time Galerkin type of least-squares support vector regression to approximate the unknown source in a finite-dimensional space, providing a stable and efficient solution. With the proposed machine learning method, we overcome the limitations of classical

*Corresponding author

Received 15 February 2024; revised 8 June 2024; accepted 11 June 2024

Abumoslem Mohammadi
Department of Mathematics, Shahed University, Tehran, Iran. e-mail: abumoslem.mohammadi@shahed.ac.ir

Abolfazl Tari Marzabad
Department of Mathematics, Shahed University, Tehran, Iran. e-mail: tari@shahed.ac.ir

How to cite this article
Mohammadi, A. and Tari Marzabad, A., A space-time least-squares support vector regression scheme for inverse source problem of the time-fractional wave equation. *Iran. J. Numer. Anal. Optim.*, 2024; 14(4): 1037-1068. https://doi.org/10.22067/ijnao.2024.86883.1392

numerical methods and offer a promising alternative for tackling inverse source problems while avoiding overfitting by carefully selecting regularization parameters. To validate the effectiveness of our approach and illustrate an exponential convergence, we present some test problems along with the corresponding numerical results. The proposed method's superior accuracy compared to the existing methods is also illustrated.

AMS subject classifications (2020): Primary 35R11; Secondary 65M70, 65M22.

Keywords: Machine learning; Support vector machines; Inverse Source problem; Time fractional wave equation; Space-time Galerkin.

1 Introduction

Inverse problems are a significant area of research in various fields, including physics, engineering, geophysics, medical imaging, and many others. The fundamental nature of an inverse problem lies in determining the cause or source of an observed effect or measurement.

In the process of designing industrial system components, there are often instances where the desired future state of the system is known as a set of observations, and the aim is to calculate the underlying causes that result from those observations. As an illustration, consider certain diffusion processes where the final temperature, together with the initial and boundary conditions, is known. However, the value of the diffusion coefficient is yet to be determined. Likewise, in other scenarios, information regarding the vibrations of a rod or the voltage at a specific future time is provided, necessitating the determination of the associated source term. These two situations can be classified as the coefficient inverse problem and the inverse source problem, respectively.

Within the medical domain, the coefficient inverse problem plays a notable role in drug delivery applications. In this context, the proper selection of ingredients is vital to achieve a targeted final density of medicine within a predetermined time frame within the blood vessels. Inverse source problems appear in diverse domains, including crack analysis in nondestructive material

evaluation [32], solving inverse heat equations to determine boundary conditions on inaccessible surfaces of a scramjet combustor [6], cell detection [7], models of light propagation in optical tomography across various scales [2], as well as determining absorption and diffusion coefficients in photo-acoustic imaging [27]. Additional instances include the retrieval of the adsorption-desorption source density rate, the exploration of elasticity problems [36], and the investigation of functionally graded materials [11].

Two important types of inverse source problems concern the wave and diffusion equations. In this work, we focus on the inverse damped wave problems and present a novel numerical simulation method based on Galerkin-type least-squares support vector regression (LS-SVR), which is a supervised machine learning algorithm.

The numerical solution of inverse problems that involve time-fractional wave equations has attracted considerable attention due to its wide range of applications in various scientific and engineering domains [3, 16, 25, 28]. Table 1 provides a short literature review of the recent works, focusing on the applications, methods, and key references in the field of inverse problems.

In recent years, support vector machines (SVM), support vector regression (SVR), and LS-SVR have emerged as powerful and versatile tools for solving differential and integral equations [18, 22, 23, 35], among others. These machine learning methods offer advantages such as high accuracy, robustness, and the ability to handle large datasets.

In this research article, our primary objective is to present a space-time Galerkin method that implements LS-SVR as a numerical solution for the inverse time-fractional wave equations. In the proposed algorithms, we aim to improve the efficiency and effectiveness of solving inverse problems in the context of time-fractional wave models. Through extensive numerical experiments and comparative analyses, we will demonstrate the superior performance of these methods in terms of accuracy, stability, and computational efficiency.

The time-fractional wave equations concerning the inverse source problems arise in reconstructing the source functions. They have found many other applications in engineering [3, 13, 25, 28, 17]. Here we focus on such equations. We denote the space domain by $\Lambda \subset \mathbb{R}^d$, which is assumed to be

Table 1: A literature review for the numerical methods for inverse problems. Let ICs and BCs stand for initial and boundary conditions.

inverse type	ref	method	applications
all types	[37]	*review work*	vehicle–bridge interaction dynamics
n-pixel image	[20]	machine learning	image reconstruction
source	[1]	deep learning	seismic inverse problem
ICs, BCs	[4]	adjoint method	heat transfer problem
source	[16]	finite element method	wave source reconstruction
source	[19]	real-time reconstruction	locations, and magnitudes of wave sources

a bounded Lipschitz domain, the time domain by $I = [0, T]$, the space-time domain by $\Omega = \Lambda \times I$, and the boundary of the space domain by $\partial\Lambda = \Gamma$. We consider the following inverse wave equation with damping:

$$tial_t^2 u + \tilde{g}(\partial_t^\alpha u) - \nabla \cdot (a\nabla u) + cu + K * u = f + \tilde{s}, \qquad \text{for all } (x,t) \in \Omega, \quad (1)$$

with the unknown source f. Here the operator $*$ denotes the convolution integral given by

$$(k * u)(t) = \int_0^t K(t-s)u(s)ds.$$

In this problem, $u,\ a,\ c$, and $\tilde{s}$ are spatiotemporal functions, $K = K(t)$, and $f = f(x)$. The damping term $\tilde{g}$ may be a nonlinear monotonic function, and the kernel K is assumed as a smooth function [12]. The linear and nonlinear cases of $\tilde{g}(x) = x$ or x^2 are discussed exclusively for the integer order $\alpha = 1$ in [29].

The problem (1) is equipped with the boundary and initial conditions

$$u(x,0) = h_0(x), \quad \text{for } x \in \Lambda, \quad (2)$$

$$u_t(x,0) = h_1(x), \quad \text{for } x \in \Lambda, \quad (3)$$

$$u|_\Gamma = 0, \quad \text{for } t \in I. \quad (4)$$

In this paper, the final observation is assumed as

$$u(x,T) = \tilde{\psi}_T(x). \quad (5)$$

The goal is to recover an unknown spatially distributed source $f = f(x) \in L^2(\Lambda)$, as well as the velocity $u \in H^1(\Omega)$.

Wave propagation in some complex media has special non-local and memory-dependent properties that are better captured in modeling with fractional derivatives. This approach allows for the incorporation of memory effects into wave equations, which leads to more accurate models [17]. For the inverse source problem (1)–(5), the uniqueness of the solution is discussed in [29].

Stability issues and challenges in inverse problems arise due to their ill-posed nature, sensitivity to noise, computational complexity, and uncertainties in the models. However, by employing regularization techniques, Bayesian inference, data filtering and preprocessing, careful model selection and validation, and proper regularization parameter selection, it is possible to mitigate these issues and obtain stable and reliable solutions to inverse problems. The proposed machine learning method in this work utilizes a Tikhonov regularization to overcome the stability issues. These aspects are discussed in the numerical results.

This paper introduces a space-time Galerkin LS-SVR approach with the orthogonal polynomial kernel that offers a unified framework for discretization, implementation, and regularization. It also demonstrates rapid convergence in both the spatial and temporal dimensions for solving the inverse time-fractional wave equation (1)–(4). The proposed method effectively addresses stability issues and achieves exponential convergence. During the training process, the solution is sought within a finite-dimensional space using simple orthogonal polynomial functions. Polynomial kernels in LS-SVR offer advantages in inverse source problems due to their ability to capture nonlinear relationships, computational efficiency, simplicity, interpretability, robustness to noise, and flexibility in model customization.

The main contribution of this work is to propose a space-time LS-SVR method, utilizing an orthogonal polynomial kernel for solving the inverse source problem of the time-fractional wave equation.

The paper is structured as follows. In Section 2, a concise overview of fractional derivatives, SVM for classification and regression, as well as approximation by orthogonal kernels is presented. Section 3 focuses on intro-

ducing our LS-SVR numerical method for simulating damped wave equations with memory. In Section 4, some test problems are provided to demonstrate the efficacy of the method. Finally, the paper concludes with some closing remarks.

2 Preliminaries

In this section, we present some preliminaries required in the next sections. The following subsections include the basic definitions of fractional calculus, function approximation with orthogonal polynomials, weighted residual methods, and SVM for classification and regression.

2.1 Modeling with fractional derivatives

Fractional derivatives provide a flexible tool for modeling certain phenomena in science and engineering offering greater compatibility with the experimental results compared to classical derivatives. It is now well established that the fractional diffusion and wave equations exhibit more realistic results and support various applications ranging from population dynamics to mathematical finance (e.g., as Black–Scholes equations). Due to the lack of analytical methods for fractional partial differential equations, developing numerical techniques is an ongoing research field.

Let $n-1 \leq \alpha < n,\ n \in \mathbb{N}$. The Caputo and Riemann–Liouville definitions for fractional derivative of order $\alpha > 0$, denoted by ${}_a^c D_t^{(\alpha)} f$ and ${}_a^{RL} D_t^{(\alpha)} f$, are, respectively, given by [8, 24]

$$ {}_a^c D_t^{(\alpha)} f(t) = \frac{1}{\Gamma(n-\alpha)} \int_a^t \frac{f^{(n)}(\tau)}{(t-\tau)^{\alpha-n+1}} d\tau, \tag{6} $$

$$ {}_a^{RL} D_t^{(\alpha)} f(t) = \frac{1}{\Gamma(n-\alpha)} \frac{\partial}{\partial t} \int_a^t \frac{f(\tau)}{(t-\tau)^{\alpha-n+1}} d\tau, \tag{7} $$

where $n-1 \leq \alpha < n$ and $n \in \mathbb{N}$.

The right Riemann–Liouville derivative of order α is given by

$$ {}_{t}^{RL}D_{T}^{(\alpha)} f(t) = \frac{(-1)^n}{\Gamma(n-\alpha)} \frac{\partial}{\partial t} \int_t^T \frac{f(\tau)}{(\tau - t)^{\alpha - n + 1}} d\tau. \tag{8} $$

2.2 LS-SVR and orthogonal kernels

Machine learning techniques, such as SVM and SVR, have gained widespread recognition within the research community and industry for their ability to classify data and predict patterns in large datasets. SVM addresses classification tasks, while SVR focuses on regression problems. Both algorithms formulate an optimization problem with inequality constraints to achieve their objectives. Similarly, the least squares variants, LS-SVM, and LS-SVR, convert these inequalities into equalities, simplifying the quadratic programming into a system of linear equations. This transformation allows for efficient training and easier solution computation. For more details, see, for example, [5, 9].

For a known data set (x_i, y_i), $i = 1, \ldots, N$, where the real numbers y_i's are the target value for the independent data $x_i \in R^{n_d}$, LS-SVR finds the wights w_i's and the bias $b \in \mathbb{R}$ in $y(x) = w^T \phi(x) + b$ with known functions ϕ_i, by solving the following optimization problem:

$$ \begin{aligned} &\min \ \frac{1}{2} w^T w + \frac{\gamma}{2} e^T e \\ &\text{s.t.} \quad y_i = w^T \phi(x_i) + b + e_i, \quad i = 1, \ldots, N. \end{aligned} \tag{9} $$

Here $\gamma \in \mathbb{R}^+$ denotes the tuning parameter $, w = [w_1, \ldots, w_M]^T, \phi = [\phi_1, \ldots, \phi_M]^T$. The hyperparameter γ controls the trade-off between the model complexity and the accuracy of the regression. It is used to penalize the deviation of the predictions from the actual target values. A smaller value of γ allows for more complex models that are able to capture fine details and noise in the data but may suffer from overfitting. In our work, in which an inverse fractional partial differential equation is considered without noise, we follow the same values as the recent publications [18, 22].

Utilizing the Lagrangian, which combines the objective function and a linear combination of the constraints using the dual variables α_i, this quadratic programming problem is transformed into a dual problem given by an equiva-

lent linear system. The following theorem establishes the equivalence between the primal problem, represented as a quadratic programming formulation (9), and a linear system [18, 21, 33, 34].

Theorem 1. An equivalent form of the quadratic programming (9) is given by

$$\begin{bmatrix} W + \frac{1}{\gamma} I_N & \mathbf{1}_N \\ \mathbf{1}_N^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ b \end{bmatrix} = \begin{bmatrix} \boldsymbol{y} \\ 0 \end{bmatrix}, \tag{10}$$

in which W is a positive definite matrix with the entries $W_{i,j} = \phi^T(x_i)\phi(x_j)$, I_N is the identity matrix, $\mathbf{1}_N^T = [1, \ldots, 1] \in \mathbb{R}^N$, $\boldsymbol{y} = [y_1, \ldots, y_N]^T$, and the dual variables $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_N]^T$. Then with the kernel $K(x, x_j) = \phi^T(x)\phi(x_j)$, we have

$$y(x) = \sum_{j=1}^{N} \alpha_j K(x, x_j) + b. \tag{11}$$

The kernel function computes similarities between pairs of data points in the input space. The purpose of the kernel in SVM is to transform the data into a higher-dimensional feature space, where it becomes easier to separate or make accurate predictions. The Mercer theorem states that a symmetric positive definite kernel function is necessary and sufficient for SVM to guarantee the existence of a corresponding feature space mapping [15].

To use orthogonal polynomials as kernels in SVR, we can define the kernel function as follows: $K(x, x') = \phi^T(x) \cdot \phi(x')$, where $\phi(x)$ is the feature map that transforms the input space into the space of orthogonal polynomials. We denote the set of polynomials with a degree less than or equal to M on the unit interval $[0, 1]$ by $\mathcal{P}_M$. A basis for this set is given by Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ with $\alpha, \beta > -1$, which are orthogonal on $[-1, 1]$, [10, 31]

$$\int_{-1}^{1} P_n^{(\alpha,\beta)}(x) P_m^{(\alpha,\beta)}(x) \chi^{(\alpha,\beta)}(x) dx = 0, \quad n \neq m, \tag{12}$$

with $\chi^{(\alpha,\beta)}(x) = (1 - x)^\alpha (1 + x)^\beta$. These functions result in fewer computations due to the sparsity of the involved matrices in the variational formulation of the method. The endpoint values are also given by

$$P_n^{(\alpha,\beta)}(-1) = (-1)^n \binom{n+\beta}{n}, P_n^{(\alpha,\beta)}(1) = \binom{n+\alpha}{n}, \tag{13}$$

which are used in handling boundary conditions for an efficient formulation of the proposed method [30].

A special case when $\alpha = 0, \beta = 0$, the Legendre polynomials are easily obtained by the recurrence formula

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x), \quad n \geq 1, \tag{14}$$

starting with $P_0(x) = 1, P_1(x) = x$. By (13), they have the boundary values $P_n(-1) = (-1)^n, P_n(1) = 1$.

We use the shifted Legendre polynomials that are orthogonal on the desired domain $[0, T]$ as

$$L_n(t) = P_n(\frac{2t}{T} - 1), \quad n = 0, 1, 2, \ldots, \tag{15}$$

with

$$\int_I L_m(t)L_n(t)dt = \frac{T}{2n+1}\delta_{m,n}. \tag{16}$$

3 LS-SVR for the inverse time-fractional wave equation

In this section, we first recall the Sobolev spaces and the specific spaces required in our formulation and the associated norms [14, 26, 31]. Then, we present the variational formulation and a space-time Galerkin LS-SVR for (1)–(5).

3.1 Functional spaces

The Lebesgue spaces L^p, $p \geq 1$ on the domain Λ, denoted as $L^p(\Lambda)$, refers to the space of measurable functions u with pth power Lebesgue integrable on Λ, with the inner product. For $p = 2$, it is a Hilbert space with the inner product $(u, v) = \int_\Lambda uvd\Lambda$ endowed with the norm $\|u\|_2 = (u, u)^{1/2}$. On the other hand, the Sobolev spaces $W^{k,p}(\Lambda)$ consist of functions with weak derivatives up to order k that are also in $L^p(\Lambda)$. It can be mathematically

expressed as

$$W^{k,p}(\Lambda) = \{u : D^s u \in L^p(\Lambda), |s| \leq k\},$$

where $D^s u$ represents the weak derivative of u with respect to the multi-index s. The associated Sobolev norm $\|u\|_{k,p}$ is given by $\|u\|_{k,p} = (\sum_{i=0}^{k} \|u^{(i)}\|_p^p)^{1/p}$. For the special case, where $p = 2$, $W^{k,2}(\Lambda)$ is a Hilbert space, denoted by $H^k(\Lambda)$. These functional spaces are defined on the spatial domain Λ. Also, these spaces on the temporal domain I are defined similarly. For a functional space X, the $x - t$ spaces for $s > 0$, are defined as

$$H^s(I; X) = \{v : \|v(\cdot, t)\|_X \in H^s(I)\},$$
$$H_0^s(I; X) = \{v : \|v(\cdot, t)\|_X \in H_0^s(I)\},$$

with the subscript zero as the Sobolev space with compact support.

Also, we define the Banach space $B^s(\Omega) = H^s(I; L^2(\Lambda)) \cap L^2(I; H_0^1(\Lambda))$ equipped with the norm

$$\|v\|_{B^s(\Omega)} = \|v\|_{H^s(I;L^2(\Lambda))} + \|v\|_{L^2(I;H_0^1(\Lambda))}.$$

Also, $C_0^\infty(I)$ refers to the smooth functions on I with compact support with the closure in terms of $\|\cdot\|_2$, $H_0^2(I)$.

Note that we first use a change of variable for the damped wave equation (1)–(5) as

$$v(x, t) = u(x, t) + h_0(x) + th_1(x),$$

to make the initial conditions homogeneous. So we have

$$\partial_t^2 u + g(\partial_t^\alpha u) - \nabla \cdot (a\nabla u) + \beta u + K * u = f + s, \qquad \text{for all } (x, t) \in \Omega, \quad (17)$$

with the initial conditions

$$u(x, 0) = 0, u_t(x, 0) = 0, \qquad \text{for all } x \in \Lambda, \quad (18)$$

the BCs

$$u|_\Gamma = 0, \qquad \text{for all } t \in I, \quad (19)$$

and the final condition

$$u(x,T) = \psi_T(x) := \tilde{\psi}_T(x) - h_0(x) - th_1(x). \tag{20}$$

Now, we introduce the variational formulation of the problem (17)–(20), an easy-to-compute associated bilinear form as well as a space-time Galerkin LS-SVR for the numerical simulation of the problem.

3.2 LS-SVR with variational formulation of the inverse time-fractional wave equation

Here, we first present the LS-SVR for $\mathcal{L}u = f$ with a linear differential operator $\mathcal{L}$ and with some appropriate homogeneous boundary conditions. The approximate solution in dual variables by using a kernel function is assumed as

$$u(x) = \sum_{j=0}^{M} \alpha_j K(x, x_j), \tag{21}$$

with a kernel $K(\cdot,\cdot)$ in which the weights are given by the following problem:

$$\begin{aligned} &\min \ \frac{1}{2} w^T w + \frac{\gamma}{2} e^T e \\ &\text{s.t. } (w^T \mathcal{L}\phi(x), \psi_i) - (f(x), \psi_i) = e_i, \quad i = 0, \ldots, M, \end{aligned} \tag{22}$$

with the assumption that ϕ_i's satisfy the homogeneous boundary conditions so do u (21).

The quadratic programming (22) may be written as an equivalent linear system. This is given by the following theorem.

Theorem 2. The dual variables α_i's in (21) are given by

$$(W + \frac{1}{\gamma} I_M)\boldsymbol{\alpha} = \boldsymbol{b}, \tag{23}$$

with

$$\begin{aligned} &b_i = (f, \psi_i), \quad i = 0, \ldots, M, \\ &W_{k,i} = \sum_{j=0}^{M} (\mathcal{L}\phi_j, \psi_k)(\mathcal{L}\phi_j, \psi_i), \qquad k, i = 0, \ldots, M. \end{aligned} \tag{24}$$

Proof. We consider the Lagrangian as

$$L = \frac{1}{2}w^T w + \frac{\gamma}{2}e^T e - \sum_{i=0}^{M} \alpha_i((w^T \mathcal{L}\phi, \psi_i) - (f, \psi_i) - e_i).$$

Then, we have the following optimality conditions:

$$\frac{\partial L}{\partial w_k} = 0 \to w_k = \sum_{i=0}^{M} \alpha_i(\mathcal{L}\phi_k, \psi_i), \tag{25}$$

$$\frac{\partial L}{\partial e_k} = 0 \to \gamma e_k + \alpha_k = 0, \tag{26}$$

$$\frac{\partial L}{\partial \alpha_k} = 0 \to (w^T \mathcal{L}\phi, \psi_k) - (f, \psi_k) - e_k = 0, \tag{27}$$

for $k = 0, \ldots, M$. So, substituting e_k from (26) into (27), we have

$$\sum_{j=0}^{M} w_j(\mathcal{L}\phi_j, \psi_k) + \frac{1}{\gamma}\alpha_k = (f, \psi_k).$$

Now, by (25), we get

$$\sum_{j=0}^{M}\sum_{i=0}^{M} \alpha_i(\mathcal{L}\phi_j, \psi_i)(\mathcal{L}\phi_j, \psi_k) + \frac{1}{\gamma}\alpha_k = (f, \psi_k).$$

Using the Kronecker delta, this can be written as

$$\sum_{i=0}^{M} \alpha_i(\sum_{j=0}^{M}(\mathcal{L}\phi_j, \psi_k)(\mathcal{L}\phi_j, \psi_i) + \frac{1}{\gamma}\alpha_k\delta_{ki}) = (f, \psi_k),$$

which is the desired result. □

Note that the system (23) has a simpler form than (10) for known datasets since the constant function is involved in the basis functions.

Now, we present an algorithm for the numerical simulation of the fractional inverse source problem in the LS-SVR using orthogonal Legendre polynomials as the kernel. For the approximation in space $[0, \pi]$, and time interval $[0, T]$, we use combinations of $\phi_m(x) = P_m(\frac{2}{\pi}x - 1)$ and $\psi_n(t) = P_n(2t/T - 1)$, respectively, which satisfy the boundary conditions.

We consider the weak formulation of (17)–(20), on the space-time domain $\Omega = \Lambda \times I$:

$$(\partial_t^2 u, v)_\Omega + (g(\partial_t^\alpha u), v)_\Omega - (\nabla \cdot (a\nabla u), v)_\Omega + c(u, v)_\Omega + (K * u, v)_\Omega = (f, v)_\Omega + (s, v)_\Omega, \tag{28}$$

for all $v \in B^{\alpha/2}(\Omega)$. For $f \in L^2(\Omega)$, the problem is to find $u \in B^{\alpha/2}(\Omega)$ such that

$$\mathcal{A}(u, v) = \mathcal{F}(v), \quad \text{for all } v \in B^{\alpha/2}(\Omega), \tag{29}$$

in which the bilinear form $\mathcal{A}$ is a manipulation of (28) and the the functional $\mathcal{F}$ are given by

$$\begin{aligned} \mathcal{A}(u, v) &= -(\partial_t u, \partial_t v)_\Omega + (g(\partial_t^\alpha u), v)_\Omega + ((a\nabla u), \nabla v)_\Omega + c(u, v)_\Omega + (K * u, v)_\Omega, \\ \mathcal{F}(v) &= (f, v)_\Omega + (s, v)_\Omega. \end{aligned} \tag{30}$$

We later consider a flexible margin with a tolerance for errors in the LS-SVR for the constraints (29). The bilinear map in the weak form characterizes the interactions between the unknown function and the derivatives using suitable test functions. The variational formulation, based on the bilinear form, leads to an algebraic system of equations enabling us to solve using numerical methods. On the other hand, the Lax–Milgram lemma ensures the existence and uniqueness of the solution as well as the continuous dependence of the solution on the given data when $\mathcal{A}$ in (30) is a continuous, bounded, and weakly coercive function, that is,

$$\begin{aligned} &\sup_{\|v\|=1} |\mathcal{A}(u, v)| \geq c\|u\| \quad \text{for all } u, \\ &\sup_{\|u\|=1} |\mathcal{A}(u, v)| > 0, \quad \text{for all } v \neq 0. \end{aligned}$$

Note that the term $(\tilde{g}(\partial_t^\alpha u), v)_\Omega$ may be written in terms of the derivative of order $\alpha/2$ by using integration by parts, the definitions of the left and right Riemann–Liouville derivatives and some algebraic manipulations. We do this by following the work of [14, Lemmas 2.1 and 2.6].

Lemma 1. Let $s \in (0, 1), u(t) \in H^s(I), v(t) \in C_0^\infty(I)$. Then, the inner product of a fractional derivative with a function changes the order as

$$({}^{RL}\partial_t^\alpha u, v)_I = (u, {}_T\partial_t^\alpha v)_I. \tag{31}$$

Proof. By (7) and integration by parts, we have

$$
\begin{aligned}
({}^{RL}\partial_t^{\alpha} u, v)_I &= \int_0^T v\, {}^{RL}\partial_t^{\alpha} u dt \\
&= \frac{1}{\Gamma(1-\alpha)} \int_0^T \frac{d}{dt} \int_0^t \frac{u(x,\tau)}{(t-\tau)^{\alpha}} d\tau\, v(x,t) dt \\
&= \frac{v(x,t)}{\Gamma(1-\alpha)} \int_0^t \frac{u(x,\tau)}{(t-\tau)^{\alpha}} d\tau \Big|_0^T - \frac{1}{\Gamma(1-\alpha)} \int_0^T \int_0^t \frac{u(x,\tau)}{(t-\tau)^{\alpha}} d\tau \partial_t v dt \\
&= -\frac{1}{\Gamma(1-\alpha)} \int_0^T \int_0^t \frac{u(x,\tau)}{(t-\tau)^{\alpha}} d\tau \partial_t v dt \\
&= -\frac{1}{\Gamma(1-\alpha)} \int_0^T \int_\tau^T \frac{\partial_t v}{(t-\tau)^{\alpha}} dt\, u(x,\tau) d\tau \\
&= -\frac{1}{\Gamma(1-\alpha)} \int_0^T (\frac{d}{d\tau} \int_\tau^T \frac{v}{(t-\tau)^{\alpha}} dt)\, u(x,\tau) d\tau \\
&= (u, {}_T\partial_t^{\alpha} v)_I.
\end{aligned}
$$

□

We use (31) to get the following result.

Lemma 2. Let $0 < \alpha < 1$, $u \in {}_0H^1(I), v \in {}_0H^{\alpha/2}(I)$. Then,

$$(\partial_t^{\alpha} u, v)_I = (\partial_t^{\alpha/2} u_{,t}\, \partial_T^{\alpha/2} v)_I. \tag{32}$$

Proof. First, we consider a sequence $v_n \in C_0^{\infty}(I)$ such that

$$\|v_n - v\|_{H^{\alpha/2}(I)} \to 0, \quad n \to \infty.$$

By (31), we have

$$
\begin{aligned}
(\partial_t^{\alpha} u, v_n)_I &= (\partial_t^{\alpha/2} \partial_t^{\alpha/2} u, v_n)_I \\
&= (\partial_t^{\alpha/2} u_{,t}\, \partial_T^{\alpha/2} v_n)_I.
\end{aligned}
$$

By using the Schwarz inequality, we get

$$|(\partial_t^{\alpha} u, v_n)_I - (\partial_t^{\alpha} u, v)_I| \le \|\partial_t^{\alpha} u\| \|v - v_n\|.$$

This tends to zero as $n \to \infty$, which leads to (32).

Now, for a constant $a \in \mathbb{R}$, we have the standard weak formulation for the problem (17)–(20) as follows. Find $u \in B^{\alpha/2}(\Omega)$ with

$$-(\partial_t u, \partial_t v)_{\Omega} + \zeta(\partial_t^{\alpha/2} u_{,t}\, \partial_T^{\alpha/2} v)_{\Omega} + a(\partial_x u, \partial_x v)_{\Omega} + c(u, v)_{\Omega} + (K * u, v)_{\Omega}$$

$$= (f, v)_\Omega + (s, v)_\Omega, \tag{33}$$

for all v in the infinite-dimensional space $B^{\alpha/2}(\Omega)$. To be practical, we present an LS-SVR formulation by orthogonally projecting this form into trial and test spaces with finite dimensions. To discretize the problem (17)–(20) in the space-time domain, the approximate solution in space is sought in $P_M^0(\Lambda)$, and the space of polynomials on Λ with degree at most M having compact support on Λ. The approximate space in time includes the functions $v \in P_N$ with $v(t = 0) = 0$. The space-time trial space is written as

$$S_{M,N} = P_M^0(\Lambda) \otimes P_N^E(I). \tag{34}$$

For simplicity, let $S_L = S_{M,N}$ be the finite-dimensional space in both directions. □

The LS-SVR for the problem (17)–(20) is as follows. By letting a tolerance in the bilinear form, we seek $u_L \in S_L$ by considering

$$\begin{aligned} &\min \ \frac{1}{2} w^T w + \frac{\gamma}{2} e^T e \\ &\text{s.t.} \quad \mathcal{A}(u_L, v_L) - \mathcal{F}(v_L) = e_i, \quad i = 0, \cdots, M, \quad \text{for all } v_L \in S_L, \end{aligned} \tag{35}$$

in which

$$u_L(x, t) = \sum_{m=0}^{M} \sum_{n=0}^{N} u_{m,n} \phi_m(x) \psi_n(t), \tag{36}$$

$$f_L(x) = \sum_{m=0}^{M} f_m \phi_m(x), \tag{37}$$

with the trial functions. To facilitate the computations in inner products and make use of the orthogonality in (35), the test functions are chosen as $v_L = \phi_i(x)\psi_j(t) \in S_L$. So, we have a sparse system of linear equations as constraints in (35), and we call the optimization problem (35) with $u_L \in S_L$ as (36), $f_L \in P_M^0(\Lambda)$ as (37), the space-time Galerkin LS-SVR for the inverse problem (17)–(20). The constraints in (35) are written as follows:

$$\begin{aligned} &-(\partial_t u_L, \partial_t v_L)_\Omega + \zeta(\partial_t^{\alpha/2} u_{L,t}\, \partial_T^{\alpha/2} v_L)_\Omega + a(\partial_x u_L, \partial_x v_L)_\Omega + c(u_L, v_L)_\Omega \\ &+(K * u_L, v_L)_\Omega = (f_L, v_L)_\Omega + (s, v_L)_\Omega + e. \end{aligned} \tag{38}$$

Note that there are alternatives such as collocating at some points and/or using other test functions. However, from a computational point of view, Galerkin LS-SVR (35) is preferred. We present analytical relations for computing the inner products in (38). The relation (38) is written in the form

$$-P^{0,0}UQ^{1,1} + \zeta P^{0,0}UQ^{\alpha/2,\alpha/2} + P^{0,0}UQ^{\alpha/2,\alpha/2} + aP^{1,1}UQ^{0,0} + cP^{0,0}UQ^{0,0} + P^{0,0}UQ_c^{0,0} = F.$$

In this relation, the matrices P, Q, U, and F are given by

$$\begin{aligned} P^{r,s} &= (\partial_x^r \phi, \partial_x^s \phi)_\Lambda, \\ Q^{r,s} &= (\partial_x^r \psi, \partial_x^s \psi)_I, \\ Q_c^{0,0} &= (\partial_x^r \psi, K * \partial_x^s \psi)_I, \\ U &= [u_{m,n} : 0 \le m, n \le M], \\ F &= (f_L + \tilde{s}, \phi_i(x)\psi_j(t))_\Omega. \end{aligned}$$

We treat the inner products in the first term analytically and the second term numerically for the sake of efficiency. The analytical part provides the use of orthogonality, which leads to sparse matrices for an efficient computational cost.

4 Numerical experiments

In this section, we consider some linear and nonlinear test problems to illustrate the convergence of the proposed space-time Galerkin LS-SVR algorithm for the inverse source problems. The purpose is to check the convergence behavior of the numerical solutions for increasing degrees of polynomials M and N and some fractional derivatives α. We plot the errors with three measures L^∞, L^2, and H^1-errors in semi-log scale to illustrate the exponential convergence. The figures include the convergence behavior in both spatial and temporal dimensions with a fixed M and N, respectively. The exponential convergence is confirmed since the logarithm of errors behaves as linear functions versus the polynomial degree. The L^∞, L^2, and H^1-errors are computed by

$$L^{\infty}(u_L) \approx \max_{0\leq i,j\leq \mathcal{M}} |U(x_i,t_j) - u_L(x_i,t_j)|, \tag{39}$$

$$L^{2}(u_L) \approx (\frac{1}{\mathcal{M}^2}\sum_{i,j=1}^{\mathcal{M}} |U(x_i,t_j) - u_L(x_i,t_j)|^2)^{1/2}, \tag{40}$$

$$H^{1}(u_L) \approx (L^2(u_L) + L^2(\partial_x u_L))^{\frac{1}{2}}, \tag{41}$$

with $x_i = \frac{i}{\mathcal{M}}$, $i = 0,\ldots,\mathcal{M}$. The errors are computed on a grid of 101×101 equidistant points on the $x-t$ domain for all numerical examples. That is, we set the number of test points, $\mathcal{M} = 100$. For the errors of the source function f, we use similar measures but on space only. We use the experimental order of convergence (EOC) in time as

$$EOC_i = \frac{\ln(e_{i+1}/e_i)}{\ln(N_i/N_{i+1})},$$

in which e_i is the error with N_i as the number of training points, $\Delta t = 1/N_i$, $i = 1,2,\ldots$ and a fixed number of spatial basis functions M. The EOC in space is computed in a similar manner.

Example 1. As the first example, we consider the linear case, $g(u) = u$, with the following data:

$$h_0 = 0, h_1 = \sin(x), u|_{\Gamma} = 0,\ a = 1,\ c = 2,\ K = -1,$$
$$\psi_T = \sin(x), s = 2\sin(x)\cos(t) + 2\sin(t)\cos(x),$$

on the domain $(x,t) \in [0,\pi] \times [0,\pi/2]$ with the exact solution $u(x,t) = \sin(t)\sin(x)$ and $f = -\sin(x)$. The problem (1)–(5) with these data was also considered for $\alpha = 1$ in [29]. With same parameters, the Landweber-type algorithm discussed in [29], gives the following errors with seven iterations

$$\|u_7(x,T) - \psi_T(x)\| \leq 10^{-4},$$
$$\|f_7 - f\| \leq 0.018.$$

The proposed scheme, that is, the space-time Galerkin LS-SVR algorithm, gives the following results after seven iterations

$$\|u_7(x,T) - \psi_T(x)\| \leq 1.31e-05,$$
$$\|f_7 - f\| \leq 2.27e-03.$$

We report the numerical errors as well as the convergence rates in terms of L^∞, L^2, and H^1 measures versus M with $N = 16$ and $\alpha = 1$, to demonstrate the spatial convergence of our proposed scheme for the wave and source functions in Tables 2 and 3, respectively. Running time is also reported for these results. It is worth noting that the results of Tables 2 and 3 are obtained at once, giving wave and source functions, respectively. Therefore, the time is reported only for one of these tables.

Table 2: L^∞, L^2, and H^1 errors and spatial convergence rates of the wave function versus M with $N = 16$ and $\alpha = 1$ for Example 1.

M	L^∞	EOC	L^2	EOC	H^1	EOC	time (s)
2	5.22e-02		2.33e-02		1.01e-01		2.01
4	1.44e-03	5.1799	7.16e-04	5.0242	5.29e-03	4.2549	10.61
6	1.98e-05	10.5723	9.12e-06	10.7610	1.34e-04	9.0655	29.53
8	1.48e-07	17.0196	6.99e-08	16.9324	2.68e-06	13.5984	63.94
10	1.34e-09	21.0830	5.10e-10	22.0504	3.07e-08	20.0289	113.37
12	1.86e-11	23.4600	8.21e-12	22.6471	2.92e-10	25.5333	194.41

The spectral convergence of the method is seen in Figures 1 and 2 in which a linear semi-log plot verifies the exponential convergence for the wave and source functions, respectively, for $\alpha = 1$.

Table 3: L^∞, L^2, and H^1 errors and spatial convergence rates for the source function f versus M with $N = 16$ and $\alpha = 1$ for Example 1 time is the same as 2.

M	$\|f - f_N\|_\infty$	EOC	$\|f - f_N\|_2$	EOC	$\|f - f_N\|_{H_1}$	EOC
2	5.04e-01		2.38e-01		6.06e-01	
4	5.78e-02	3.1243	2.72e-02	3.1293	1.35e-01	2.1664
6	2.64e-03	7.6115	1.18e-03	7.7385	9.95e-03	6.4314
8	6.95e-05	12.6432	3.11e-05	12.6392	3.34e-04	11.7984
10	1.06e-06	18.7460	4.73e-07	18.7586	6.07e-06	17.9606
12	1.08e-08	25.1560	4.83e-09	25.1438	4.59e-08	26.7914

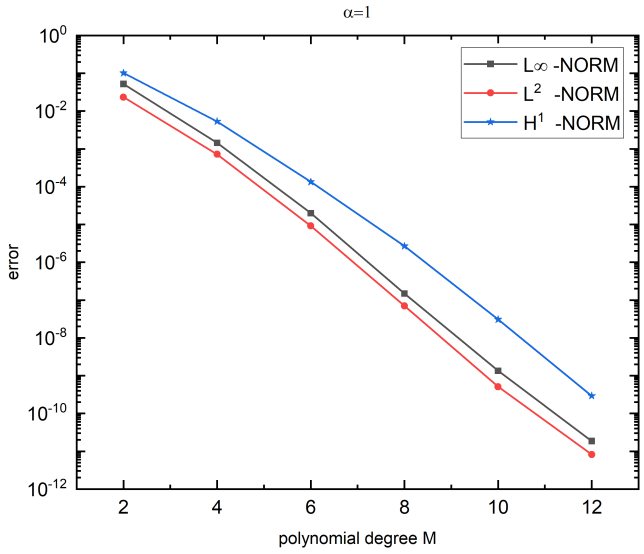

Figure 1: L^∞, L^2, and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 1$ for Example 1.

Tables 4, 5, and 6 demonstrate the spatial convergence of our proposed scheme as well as the running time for the wave function versus M with $N = 16$ for some fractional orders.

Table 4: L^∞, L^2, and H^1 errors and spatial convergence rates of the wave function versus M with $N = 16$ and $\alpha = 0.1$ for Example 1.

M	L_∞	EOC	L_2	EOC	H_1	EOC	time (s)
2	5.76e-02		3.03e-02		9.76e-02		2.06
4	1.61e-03	5.1609	7.61e-04	5.3153	5.15e-03	4.2442	10.98
6	7.34e-05	7.6161	3.48e-05	7.6086	1.41e-04	8.8737	26.01
8	4.61e-07	17.6246	2.10e-07	17.7636	2.71e-06	13.7367	55.78
10	3.03e-09	22.5184	1.24e-09	22.9986	3.12e-08	20.0063	162.14
12	2.87e-11	25.5561	1.46e-11	24.3627	2.75e-10	25.9509	257.39

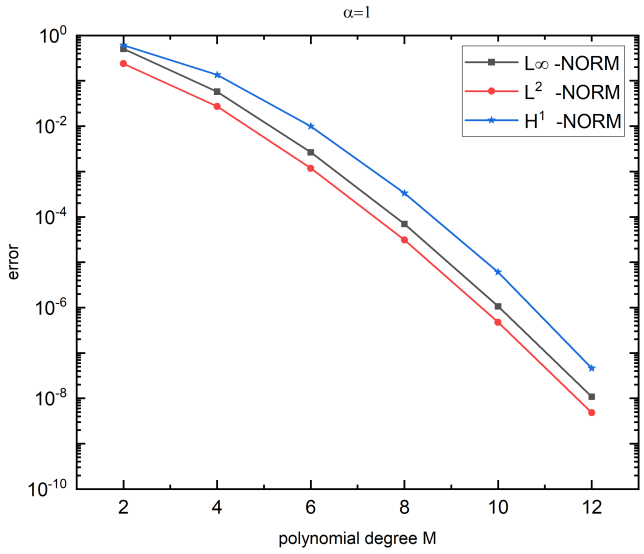

Figure 2: L^∞, L^2, and H^1 errors for the source function versus M with $N = 16$ and $\alpha = 1$ for Example 1.

The spectral convergence of the method is seen in Figures 3, 4, and 5, where a linear semi-log plot verifies the exponential convergence for the wave function for some fractional derivatives.

Table 5: L^∞, L^2, and H^1 errors and spatial convergence rates of the wave function versus M with $N = 16$ and $\alpha = 0.5$ for Example 1.

M	L^∞	EOC	L^2	EOC	H^1	EOC
2	5.22e-02		2.26e-02		9.72e-02	
4	1.55e-03	5.0737	7.48e-04	4.9171	5.21e-03	4.2216
6	4.23e-05	8.8817	1.95e-05	8.9946	1.30e-04	9.1026
8	3.00e-07	17.2022	1.36e-07	17.2604	2.67e-06	13.5061
10	2.25e-09	21.9269	8.78e-10	22.5987	3.09e-08	19.9830
12	2.47e-11	24.7468	1.27e-11	23.2339	2.80e-10	25.7990

Now, we investigate the effect of increasing the number of training points in the time direction. To do so, we fix the number of used polynomials in the

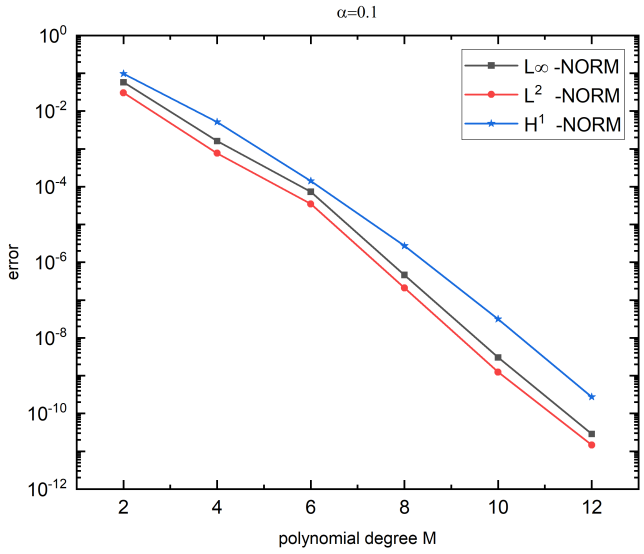

Figure 3: L^∞, L^2, and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 0.1$ for Example 1.

Table 6: L^∞, L^2, and H^1 errors and spatial convergence rates of the wave function versus M with $N = 16$ and $\alpha = 0.9$ for Example 1.

M	L_∞	EOC	L_2	EOC	H_1	EOC
2	5.22e-02		2.27e-02		1.01e-01	
4	1.46e-03	5.1600	7.24e-04	4.9706	5.28e-03	4.2577
6	2.24e-05	10.3021	1.06e-05	10.4175	1.33e-04	9.0793
8	1.72e-07	16.9261	7.97e-08	16.9991	2.67e-06	13.5854
10	1.51e-09	21.2212	5.69e-10	22.1478	3.07e-08	20.0121
12	2.47e-11	24.7468	1.27e-11	23.2339	2.80e-10	25.7990

space dimension $M = 15$. We report the L^∞, L^2, and H^1 errors to show the temporal convergence of our proposed scheme for the wave function versus N and $\alpha = 1$, $\alpha = 0.5$ in Tables 7 and 8, respectively.

The temporal convergence of our proposed scheme for the wave function versus N with $M = 15$ and $\alpha = 1$, $\alpha = 0.5$ in Figures 6 and 7, respectively.

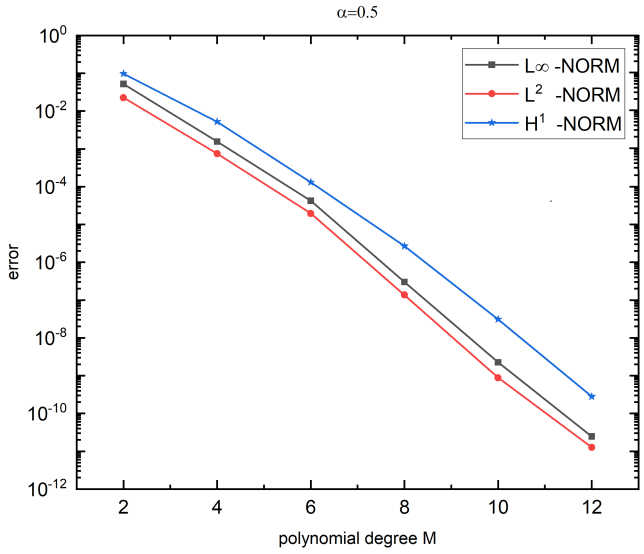

Figure 4: L^∞, L^2, and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 0.5$ for Example 1.

Table 7: L^∞, L^2, and H^1 errors and temporal convergence rates of the wave function versus N with $M = 15$ and $\alpha = 1$ for Example 1.

N	L_∞	EOC	L_2	EOC	H_1	EOC
2	6.85e-02		3.41e-02		4.83e-02	
4	1.94e-03	5.1420	8.92e-04	5.2566	1.60e-03	4.9159
6	7.69e-06	13.6400	3.24e-06	13.8554	5.08e-06	14.1873
8	2.26e-08	20.2645	1.21e-08	19.4316	1.63e-08	19.9592
10	6.32e-11	26.3481	3.28e-11	26.4876	4.55e-11	26.3562
12	2.26e-13	30.8988	1.06e-13	31.4540	1.41e-12	19.0549

The next example illustrates the convergence of the proposed method in space and time for the wave and source functions in a nonlinear case.

Example 2. Consider the problem (1) with the same parameters but the nonlinear function $g(u) = u^2$,

$$s = \sin^2(x)\cos^2(t) + 2\sin(t)\sin(x) + \cos(t)\sin(x),$$

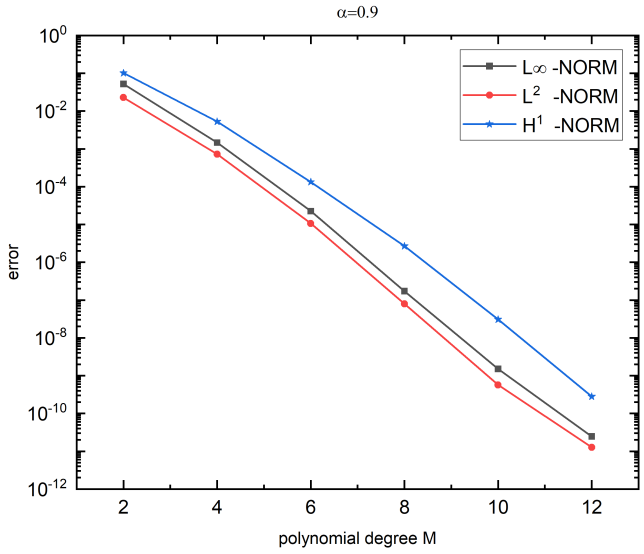

Figure 5: L^∞, L^2, and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 0.9$ for Example 1.

Table 8: L^∞, L^2, and H^1 errors and temporal convergence rates of the wave function versus N with $M = 15$ and $\alpha = 0.5$ for Example 1.

N	L^∞	EOC	L^2	EOC	H^1	EOC
2	6.85e-02		3.41e-02		4.83e-02	
4	2.38e-03	4.8471	1.11e-03	4.9411	1.72e-03	4.8115
6	9.07e-06	13.7370	3.86e-06	13.9628	5.89e-06	14.0008
8	2.94e-08	19.9238	1.53e-08	19.2246	2.16e-08	19.4948
10	8.05e-11	26.4426	4.25e-11	26.3781	5.83e-11	26.5069
12	4.18e-13	28.8530	1.88e-13	29.7322	1.49e-12	20.1119

on the same domain $[0, \pi] \times [0, \pi/2]$ as for Example 1 with the exact solution $u(x, t) = \sin(t)\sin(x)$ and $f(x) = -\sin(x)$. The problem (1)–(5) with these data was also considered for $\alpha = 1$ in [29]. The results (errors) obtained by the Landweber-type algorithm discussed in [29] with 13 iterations are

$$\|u_{13}(x, T) - \psi_T(x)\| \leq 0.001,$$

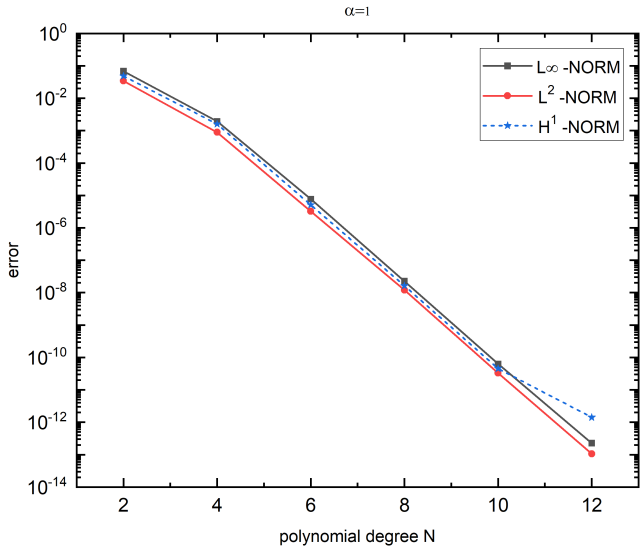

Figure 6: L^∞, L^2, and H^1 errors for the wave function versus N with $M = 15$ and $\alpha = 1$ for Example 1.

$$\|f_{13} - f\| \leq 0.02.$$

By the proposed scheme of this paper for this example with 13 iterations, the following results are obtained

$$\|u_{13}(x,T) - \psi_T(x)\| \leq 1.90e - 11,$$
$$\|f_{13} - f\| \leq 8.66e - 09.$$

We report the numerical errors as well as the convergence rates in terms of L^∞, L^2, and H^1 measures versus M with $N = 16$ and $\alpha = 1$ to demonstrate the spatial convergence of our proposed scheme for the wave and source functions in Tables 9 and 10, respectively.

The spectral convergence of the method is seen in Figures 8 and 9 in which linear semi-log plot verifies the exponential convergence for the wave and source functions, respectively, for $\alpha = 1$.

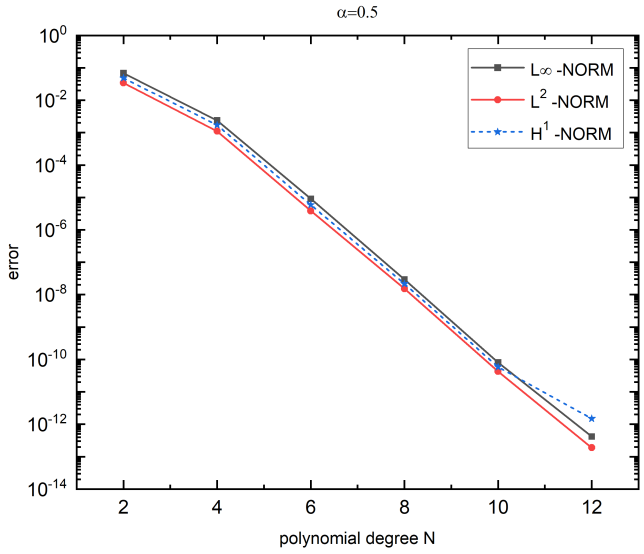

Figure 7: L^∞, L^2, and H^1 errors for the wave function versus N with $M = 15$ and $\alpha = 0.5$ for Example 1.

Table 9: L^∞, L^2, and H^1 norm errors of the wave function versus M with $N = 16$ and $\alpha = 1$ for the time-fractional wave equation using space-time Galerkin LS-SVR.

M	L^∞	EOC	L^2	EOC	H^1	EOC
2	5.84e-02		3.08e-02		9.77e-02	
4	1.62e-03	5.1719	7.56e-04	5.3484	5.11e-03	4.2570
6	7.13e-05	7.7030	3.48e-05	7.5923	1.41e-04	8.8545
8	3.92e-07	18.0873	1.77e-07	18.3579	2.62e-06	13.8541
10	2.40e-09	22.8364	9.66e-10	23.3515	3.07e-08	19.9274
12	2.45e-11	25.1454	1.24e-11	23.8889	2.74e-10	25.8822

Table 11 demonstrates the convergence in time of the proposed scheme for the wave function, versus N with $M = 15$ for $\alpha = 0.5$.

The spectral convergence in time of the method is seen in Figure 10 in which a linear semi-log plot verifies the exponential convergence for the wave function for $\alpha = 0.5$.

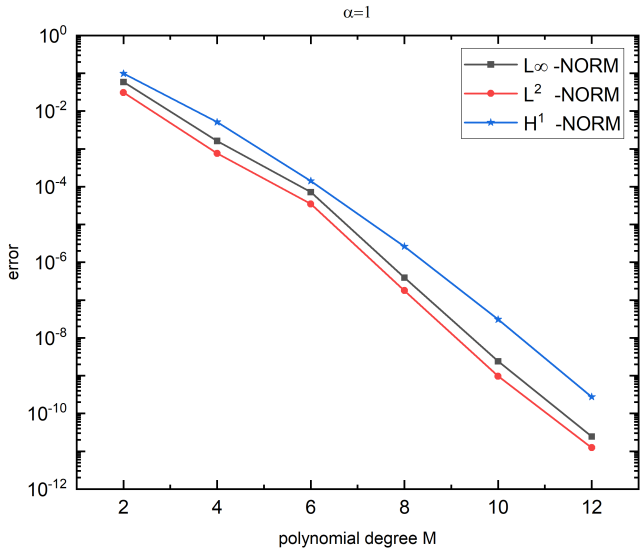

Figure 8: L^∞, L^2, and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 1$ for Example 2.

Table 10: L^∞, L^2, and H^1 errors and spatial convergence rates of the source function versus M with $N = 16$ and $\alpha = 1$ for Example 2.

M	$\|f - f_N\|_\infty$	EOC	$\|f - f_N\|_2$	EOC	$\|f - f_N\|_{H_1}$	EOC
2	4.76e-01		2.23e-01		5.12e-01	
4	5.48e-02	3.1187	2.55e-02	3.1285	1.24e-01	2.0458
6	2.63e-03	7.4894	1.20e-03	7.5379	1.07e-02	6.0425
8	6.95e-05	12.6300	3.12e-05	12.6864	3.51e-04	11.8784
10	1.06e-06	18.7460	4.73e-07	18.7730	6.08e-06	18.1757
12	1.08e-08	25.1560	4.83e-09	25.1438	5.05e-08	26.2766

As observed from the numerical results, the proposed LS-SVR, which employs orthogonal kernels with Legendre polynomials, yields approximate models with the desired accuracy and demonstrates favorable convergence behavior as the number of training points in both dimensions increases.

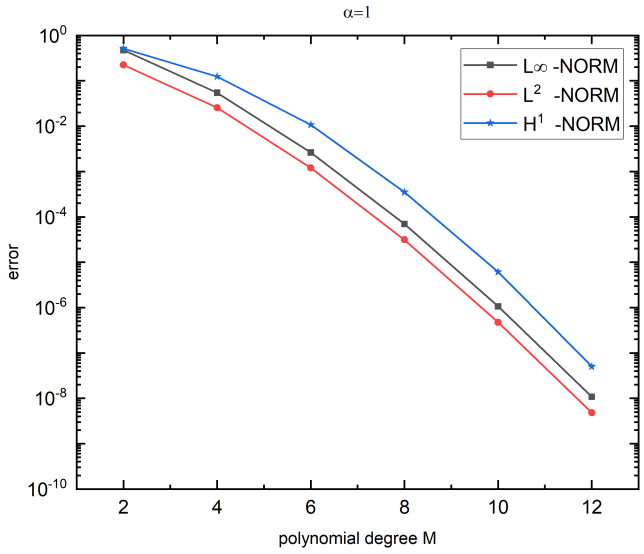

Figure 9: L^∞, L^2, and H^1 errors for the source function versus M with $N = 16$ and $\alpha = 1$ for Example 2.

Table 11: L^∞, L^2, and H^1 errors of the source function and temporal convergence rates versus N with $M = 15$ and $\alpha = 1$ for Example 2.

N	L^∞	EOC	L^2	EOC	H^1	EOC
2	6.85e-02		3.41e-02		4.83e-02	
4	2.40e-03	4.8350	1.13e-03	4.9154	1.67e-03	4.8541
6	9.55e-06	13.6305	4.23e-06	13.7811	6.45e-06	13.7040
8	3.25e-08	19.7546	1.65e-08	19.2803	2.34e-08	19.5323
10	8.37e-11	26.7171	4.59e-11	26.3716	6.30e-11	26.5182
12	5.87e-13	27.2045	2.83e-13	27.9110	1.66e-12	19.9445

5 Conclusion

In this study, we addressed an inverse source problem concerning the time-fractional wave equation. We presented a space-time Galerkin LS-SVR for the numerical simulation of the problem. We employed an orthogonal polynomial

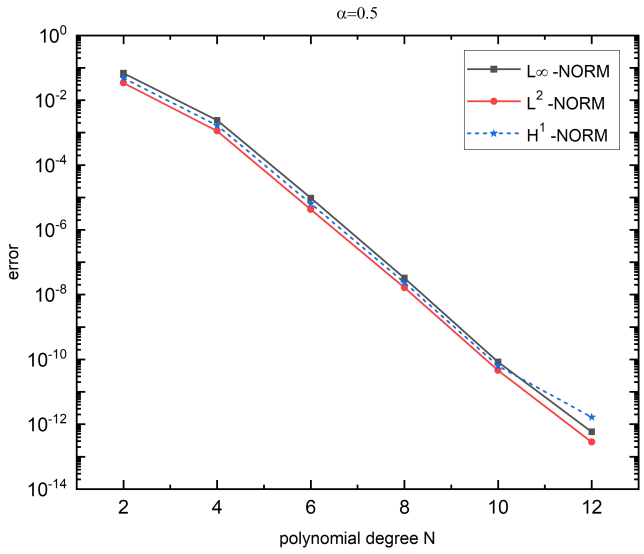

Figure 10: L^∞, L^2, and H^1 errors for the wave function versus N with $M = 15$ and $\alpha = 0.5$ for Example 2.

kernel in both space and time variables simultaneously. We presented the method as a quadratic programming problem and subsequently transformed it into a system of linear algebraic equations by introducing dual variables. The structure of the resulting system was also discussed, and some numerical test problems were given to demonstrate the efficacy of the proposed method. The method can be further developed for the reconstruction of sources in the presence of noisy data in future works as well as for other cases of inverse problems.

Acknowledgments

The authors are very grateful to the anonymous referees and the editor of this journal for carefully reading the paper and for their constructive comments and suggestions to improve the paper.

Funding: The authors did not receive support from any organization for the

submitted work.

Financial interests: The authors declare that they have no conflict of interest.

Availability of data and material: Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Code availability: Not applicable.

Authors' contributions: The order of the author list reflects contributions to the paper. All authors read and approved the final manuscript.

References

[1] Adler, A., Araya-Polo, M., and Poggio T., *Deep learning for seismic inverse problems: Toward the acceleration of geophysical analysis workflows,* IEEE Signal Process. Mag. 38(2) (2021), 89–119.

[2] Arridge, S.R. and Schotland, J.C., *Optical tomography: forward and inverse problems,* Inverse Probl. 25(12) (2009), 123010.

[3] Ashurov, R.R. and Faiziev, Y.É., *Inverse problem for finding the order of the fractional derivative in the wave equation,* Math.l Notes, 110(5-6) (2021), 842–852.

[4] Bourquin, F. and Nassiopoulos, A. *Inverse reconstruction of initial and boundary conditions of a heat transfer problem with accurate final state,* Int. J. Heat Mass Transf. 54(15-16) (2011), 3749–3760.

[5] Campbell, C. and Ying, Y. *Learning with support vector machines,* Springer Nature, 2022.

[6] Cui, M., Mei, J., Zhang, B-W, Xu, B.-B., Zhou, L., and Zhang, Y., *Inverse identification of boundary conditions in a scramjet combustor with a regenerative cooling system,* Appl. Therm. Eng. 134 (2018), 555–563.

[7] Del Aguila Pla, P. and Jaldén, J. *Cell detection by functional inverse diffusion and non-negative group sparsity—part ii: Proximal optimiza-*

tion and performance evaluation, IEEE Trans. Signal Process. 66(20) (2018), 5422–5437.

[8] Diethelm, K., *The analysis of fractional dierential equations: An application-oriented exposition using dierential operators of Caputo type,* Springer, 2010.

[9] Ernst, F. and Schweikard, A. *Fundamentals of Machine Learning: Support Vector Machines Made Easy,* Utb GmbH, 2020.

[10] Guo, B. *Spectral methods and their applications,* World Scientific, 1998.

[11] Hu, W., Gu, Y., and Fan, C.M., *A meshless collocation scheme for inverse heat conduction problem in three-dimensional functionally graded materials,* Eng. Anal. Bound. Elem. 114 (2020), 1–7.

[12] Jiang, D., Liu, Y., and Yamamoto, M., *Inverse source problem for a wave equation with final observation data,* Mathematical Analysis of Continuum Mechanics and Industrial Applications, 26 (2017), 153–164.

[13] Kinash, N. and Janno, J., *An inverse problem for a generalized fractional derivative with an application in reconstruction of time-and space-dependent sources in fractional diffusion and wave equations,* Mathematics, 7(12) (2019), 1138.

[14] Li, X. and Xu, C., *A space-time spectral method for the time-fractional diffusion equation,* SIAM J. Numer. Anal. 47(3) (2009), 2108–2131.

[15] Liu, Y. and Xie, M., *Rebooting data-driven soft-sensors in process industries: A review of kernel methods,* J. Process Control, 89 (2020), 58–73.

[16] Lloyd, S., Schaal, C. and Jeong, C., *Inverse modeling and experimental validation for reconstructing wave sources on a 2D solid from surficial measurement,* Ultrasonics, 128 (2023), 106880.

[17] Mainardi, F., *Fractional calculus in wave propagation problems,* In Forum der Berliner Mathematischer Gesellschaft, vol. 19, pp. 20-52. 2011.

[18] Mehrkanoon, S. and Suykens, J.A., *Learning solutions to partial differential equations using LS-SVM,* Neurocomputing, 159 (2015), 105–116.

[19] Ohe, T., *Real-time reconstruction of moving point/dipole wave sources from boundary measurements,* Inverse Probl. Sci. Eng. 28(8) (2020), 1057–1102.

[20] Ongie, G., Jalal, A., Metzler, C.A., Baraniuk, R.G., Dimakis, A.G., and Willett, R., *Deep learning techniques for inverse problems in imaging,* IEEE J. Sel. Areas Inf. Theory, 1(1) (2020), 39–56.

[21] Pakniyat, A., Parand, K., and Jani, M., *Least squares support vector regression for differential equations on unbounded domains,* Chaos Solit. Fractals. 151 (2021), 111232.

[22] Parand, K., Aghaei, A.A., Jani, M. and Sahleh, R., *Solving integral equations by ls-svr. In Learning with Fractional Orthogonal Kernel Classifiers in Support Vector Machines: Theory, Algorithms and Applications,* Singapore: Springer Nature Singapore, 2023.

[23] Parand, K., Aghaei, A.A., Kiani, S., Zadeh, T.I. and Khosravi, Z., *A neural network approach for solving nonlinear differential equations of Lane–Emden type,* Eng. Comput. (2023), 1–7.

[24] Podlubny, I., *Fractional Differential Equations,* Academic Press, New York, 1999.

[25] Qiu, L., Hu, C. and Qin, Q.H., *A novel homogenization function method for inverse source problem of nonlinear time-fractional wave equation,* Appl. Math. Lett.109 (2020), 106554.

[26] Reddy, B.D., *Introductory functional analysis: with applications to boundary value problems and finite elements,* Springer Science and Business Media, 2013.

[27] Ren, K. and Triki, F., *A Global stability estimate for the photo-acoustic inverse problem in layered media,* Eur. J. Appl. Math. 30(3) (2019), 505–528.

[28] Sakamoto, K. and Yamamoto, M., *Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems,* J. Math. Anal. Appl. 382(1) (2011), 426–447.

[29] Seliga, L. and Slodicka, M., *An inverse source problem for a damped wave equation with memory,* J. Inverse Ill-Posed Probl. 24(2) (2016), 111–122.

[30] Shen, J. and Sheng, C.-T., *An efficient space–time method for time-fractional diffusion equation,* J. Sci. Comput. 81 (2019), 1088–1110.

[31] Shen J., Tang T., and Wang L.L., *Spectral methods: algorithms, analysis and applications,* Springer, 2011.

[32] Smith, R.C. and Demetriou, M.A., *Research directions in distributed parameter systems,* SIAM, 2003

[33] Suykens, J.A.K, Gestel, T.V., De Brabanter, J., De Moor, B., and Vandewalle, J., *Least squares support vector machines,* Singapore: World Scientific Publishing Company, 2002.

[34] Suykens, J.A.K. and Vandewalle, J., *Least squares support vector machine classifiers,* Neural Process. Lett. 9(3) (1999), 293–300.

[35] Taheri, T., Aghaei, A.A., and Parand, K., *Bridging machine learning and weighted residual methods for delay differential equations of fractional order,* Appl. Soft Comput. (2023) 110936.

[36] Yang, J.P. and Hsin, W.-C., *Weighted reproducing kernel collocation method based on error analysis for solving inverse elasticity problems,* Acta Mech. 230 (2019), 3477–3497.

[37] Zhu, X., and Law, S.-S., *Recent developments in inverse problems of vehicle–bridge interaction dynamics,* J. Civ. Struct. Health Monit. 6 (2016), 107–128.