



A supervised learning algorithm for the inverse source problem of fractional wave equations

Abumoslem Mohammadi¹ and Abolfazl Tari Marzabad^{2,*}

¹Department of Mathematics, Imam Ali University, Tehran, Iran.

²Department of Mathematics, Shahed University, Tehran, Iran.

Abstract

Inverse problems for partial differential equations play an important role in a wide range of scientific disciplines and enable us to recover crucial information about underlying physical processes. In this paper, we present a machine-learning algorithm for solving inverse source problems of time fractional wave equations using support vector regression with polynomial kernels. This innovative approach leverages the power of machine-learning to estimate elusive source parameters, providing a highly accurate and efficient solution. By combining the principles of support vector regression and polynomial kernels, our method offers a promising alternative to traditional numerical techniques, achieving remarkable results while maintaining computational efficiency. Through comprehensive experiments and comparisons, we demonstrate the superior performance and potential of our approach in addressing inverse source problems of time fractional wave equations in linear and nonlinear cases.

Keywords. Inverse problem, Support vector regression, Time-fractional PDEs, Model optimization.

2010 Mathematics Subject Classification. 35R11, 65M70, 65M22, 76M22.

1. INTRODUCTION

Partial differential equations (PDEs) serve as a fundamental framework for modeling a wide range of physical phenomena in various scientific and engineering disciplines. Among the many challenges posed by PDEs, inverse source problems stand out as critical puzzles, offering profound insights into our ability to uncover hidden parameters that underlie complex systems. These problems entail the task of estimating source terms within PDEs, often representing physical quantities that cannot be directly measured. The accurate determination of these source terms is essential in numerous real-world applications, making inverse source problems an area of paramount importance in scientific research and engineering.

Inverse problems within the context of PDEs have become a significant focal point of research due to their direct relevance to understanding and controlling physical systems. They arise in a multitude of fields, such as photo-acoustic imaging [8], geophysics and light propagation in optical tomography [10], environmental modeling and material science [6], image reconstruction [7] and seismic inverse problem [14] where the challenge lies in uncovering the sources of phenomena that manifest as observed data. By solving these inverse problems, researchers can gain invaluable insights into the underlying mechanisms governing these systems.

As an important case of inverse problems in PDEs, the time fractional wave equations with unknown source concerns the challenges to systems that exhibit complex evolution behavior [1, 15, 22, 24]. Time fractional wave equations capture phenomena with memory and long-range dependencies, which are prevalent in wave propagation, acoustics, and optics. Solving inverse source problems in such contexts is vital for reconstructing source distributions in both time and space, thus enabling a deeper understanding of wave phenomena.

Received: 16 November 2024 ; Accepted: 08 September 2025.

* Corresponding author. Email: tari@shahed.ac.ir.

Over the years, classic numerical methods have been developed to solve inverse source problems in PDEs. These approaches often involve discretization schemes and regularization techniques. While they have made significant contributions in this field, they also have limitations such as sensitivity to noise, overfitting, and computational costs.

In recent years, the advent of machine-learning has motivated scholars to address the solution of differential equations and PDEs by developing some learning techniques with appropriate kernels [20, 28, 29]. These methods can bypass some of the limitations of classical approaches, offering a fresh perspective on solving functional problems including PDEs and integral equations.

Despite the limitations of classical numerical methods for solving PDEs, such as difficulties in handling complex geometries, high-dimensional domains, and nonlinearities, machine-learning techniques offer flexible algorithms capable of addressing a broader class of problems with arbitrary domain shapes and increased complexity.

In this paper, we aim to introduce a novel machine-learning algorithm designed to address the numerical simulation of inverse source problems in time fractional wave equations by developing and support vector regression with a polynomial kernel.

Fractional wave equations are used to model wave propagation in complex media with non-local and memory-dependent properties. In some materials, the response to a wave at a given point depends not only on the current state but also on the history of the wave. Fractional derivatives allow for the incorporation of memory effects into wave equations, leading to more accurate models [16].

The inverse problems given as time-fractional wave equations have interesting applications such as reconstructing the source functions, [1, 11, 16, 22, 24]. Let $\Lambda \subset \mathbb{R}^d$ be a bounded Lipschitz domain, I be the time interval up to the final time T_f , $I = [0, T_f]$, $\Omega = \Lambda \times I$ and $\partial\Omega = \Gamma$ be the boundary. We consider the inverse time-fractional wave equation as

$$\partial_t^2 v + \tilde{g}(\partial_t^\alpha v) - \nabla \cdot (a \nabla v) + cv + K * v = f + \tilde{s}, \quad \forall (x, t) \in \Omega, \quad (1.1)$$

where the convolution $*$ is defined by

$$(K * v)(t) = \int_0^t K(t-s)v(s)ds,$$

with unknown source f . Here, $\partial_t^2 v$ represents the second derivative with respect to time and $\partial_t^\alpha v$ indicates the fractional derivative in terms of Caputo definition. The variables v , a , c , and \tilde{s} depend on space and time and $K = K(t)$, $f = f(x)$. The function \tilde{g} may represent a nonlinear behaviour [27]. Following the reference, we consider the linear and nonlinear cases $\tilde{g}(v) = v$ or v^2 .

The initial and boundary conditions are given as

$$v(x, 0) = h_0(x), \quad \text{for } x \in \Lambda, \quad (1.2)$$

$$v_t(x, 0) = h_1(x), \quad \text{for } x \in \Lambda, \quad (1.3)$$

$$v|_\Gamma = 0, \quad \text{for } t \in I. \quad (1.4)$$

Since f is an unknown source, an additional condition is assumed as

$$v(x, T_f) = \tilde{\psi}_{T_f}(x). \quad (1.5)$$

So we need to recover the source $f = f(x) \in L^2(\Lambda)$, and the wave solution $v \in H^1(\Omega)$.

Our aim is to develop a numerical simulation with a supervised least squares algorithm with the orthogonal polynomial kernel for the time-fractional inverse source problem (1.1)-(1.5) with fast convergence. In the training process, the parameters of the numerical model are trained by a complete residual function corresponding to the problem in a finite-dimensional polynomial space.

The structure of the paper is as follows. Section 2 gives the basic definitions and relations in fractional calculus, supervised learning algorithms, and orthogonal kernels. In section 3, we present a machine-learning numerical method for recovering source and wave functions in the time-fractional inverse source wave equation. Finally, some examples demonstrate the efficacy of the proposed method in 4.



2. PRELIMINARIES

In this section, we cover the essential concepts needed for upcoming sections. These include basic definitions of fractional calculus machine-learning and kernel methods for classification and regression.

2.1. Fractional derivatives. Fractional derivatives offer a versatile approach for modeling scientific and engineering phenomena, offering better compatibility with experimental results compared to classical derivatives. Fractional diffusion and wave equations have been found to yield more realistic results and have diverse applications, from population dynamics to mathematical finance (e.g., Black–Scholes equations). Limitations of the analytical methods for fractional PDEs motivate ongoing research focusing on developing numerical techniques for addressing the challenges concerning the solution profile, computational complexity, stability, and convergence.

The Caputo and Riemann–Liouville fractional derivative of order α , denoted by ${}_a^c D_t^{(\alpha)} f$ and ${}_a^{RL} D_t^{(\alpha)} f$, respectively are defined as [3, 9, 19, 21]

$${}_0^c \partial_t^\alpha u(x, t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{\partial u(x, \tau)}{\partial \tau} \frac{d\tau}{(t-\tau)^\alpha}, \quad (2.1)$$

$${}_0^{RL} \partial_t^\alpha u(x, t) = \frac{1}{\Gamma(1-\alpha)} \frac{\partial}{\partial t} \int_0^t \frac{u(x, \tau)}{(t-\tau)^\alpha} d\tau, \quad (2.2)$$

where $0 < \alpha < 1$. The two definitions are related by

$${}_0^{RL} \partial_t^\alpha u(x, t) = {}_0^c \partial_t^\alpha u(x, t) + \frac{u(x, 0)}{\Gamma(1-\alpha)t^\alpha}.$$

2.2. LS-SVR and polynomial kernels. Machine-learning techniques, including Support Vector Machines (SVM) and Support Vector Regression (SVR), have garnered widespread acclaim in the research community and industry for their capacity to classify data and predict patterns in extensive datasets. SVM is designed for classification tasks, while SVR is tailored for 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, transform these inequalities into equalities, simplifying quadratic programming into a system of linear equations. This adaptation streamlines training and computation processes, ensuring efficiency. Further details can be found in references [2, 4].

For a given data set (x_i, y_i) , $i = 1, \dots, N$, where the real numbers y_i 's represent the target value for the independent features $x_i \in \mathbb{R}^{n_d}$, LS-SVR determines the weights w_i 's and the bias $b \in \mathbb{R}$ in $y(x) = w^T \phi(x) + b$ using known functions ϕ_i , by solving the following quadratic programming:

$$\begin{aligned} \min \quad & \frac{1}{2} w^T w + \frac{\gamma}{2} e^T e, \\ \text{s.t.} \quad & y_i = w^T \phi(x_i) + b + e_i, \quad i = 1, \dots, N, \end{aligned} \quad (2.3)$$

in which $\gamma \in \mathbb{R}^+$ denotes the tuning parameter, $w = [w_1, \dots, w_M]^T$, $\phi = [\phi_1, \dots, \phi_M]^T$.

By using the Lagrangian to combine the objective function and a linear combination of the constraints with the dual variables α_i , the quadratic programming problem is reformulated into a dual problem, which is equivalent to a linear system. The following theorem establishes the equivalence between the primal problem, expressed as the quadratic programming formula (2.3), and a linear system [17, 18].

Theorem 2.1. *The problem (2.3) is equivalent to*

$$\begin{bmatrix} W + \frac{1}{\gamma} I_N & \mathbf{1}_N \\ \mathbf{1}_N^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ b \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}. \quad (2.4)$$

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

$$y(x) = \sum_{j=1}^N \alpha_j K(x, x_j) + b. \quad (2.5)$$



Kernels are integral in machine-learning, particularly in SVR, as they enable the transformation of data into higher-dimensional feature spaces, facilitating more accurate predictions and improved separability of data points. This transformation is essential for handling complex and nonlinear relationships within datasets, allowing for more robust and effective modeling in various regression tasks [13].

In the context of the machine-learning framework, when utilizing orthogonal polynomials as kernels in SVR, the kernel function can be defined as follows: $K(x, x') = \phi^T(x) \cdot \phi(x')$ where $\phi(x)$ represents the feature map transforming the input space into the orthogonal polynomial space. Jacobi polynomials $P_n^{(\alpha, \beta)}(x)$, with $\alpha, \beta > -1$ are orthogonal on the interval $[-1, 1]$ with respect to the weight function $\chi^{(\alpha, \beta)}(x) = (1-x)^\alpha(1+x)^\beta$ [5, 26]. The set of polynomials with degree less than or equal to M on the unit interval $[0, 1]$ denoted by \mathcal{P}_M has a basis given by shifted Jacobi polynomials $P_n^{(\alpha, \beta)}(2x-1)$. These functions lead to reduced computations due to the sparse nature of the matrices involved in the variational formulation of the method. The end-point values for handling boundary conditions by these functions are provided by [25]

$$P_n^{(\alpha, \beta)}(-1) = (-1)^n \binom{n+\beta}{n}, P_n^{(\alpha, \beta)}(1) = \binom{n+\alpha}{n}. \quad (2.6)$$

Let $\alpha = -\frac{1}{2}, \beta = -\frac{1}{2}$, as a special case, Chebyshev polynomials are computed by the recursive formula

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geq 1, \quad (2.7)$$

starting with $T_0(x) = 1, T_1(x) = x$. By (2.6), they have the boundary values $T_n(-1) = (-1)^n, T_n(1) = 1$. The shifted Chebyshev polynomials that are orthogonal on the desired domain $I = [0, T_f]$ are given by

$$\tilde{T}_n(t) = T_n\left(\frac{2t}{T_f} - 1\right), \quad n = 0, 1, 2, \dots, \quad (2.8)$$

with

$$\int_I \frac{\tilde{T}_m(t)\tilde{T}_n(t)}{\sqrt{T_f t - t^2}} dt = h_n \delta_{m,n}, \quad (2.9)$$

where $h_n = \frac{b_n \pi}{2}$, $b_0 = 2, b_j = 1, j \geq 1$.

3. A POLYNOMIAL-KERNEL LEARNING METHOD FOR THE FRACTIONAL INVERSE WAVE PROBLEM (1.1)-(1.5)

Here, we first review the Sobolev spaces and specific spaces necessary for our formulation, along with their inner products and associated norms [12, 23, 26]. Then, we introduce the variational formulation, an ansatz in the form of a polynomial in two dimensions, and a machine-learning approach for the training of the parameters of the ansatz.

3.1. Universal approximation theorem and approximation in function spaces. The universal approximation theorem states that a feedforward neural network with a single hidden layer containing a finite number of neurons can approximate continuous functions on compact subsets of Euclidean space, under mild assumptions on the activation function. This theorem provides theoretical support for the expressive power of neural networks in function approximation tasks.

The function space $L^p(\Lambda)$, $p \geq 1$ on $\Lambda \subset \mathbb{R}^d$, includes the functions u with p th power Lebesgue integrable on Λ . For $p = 2$, it is a Hilbert space with the inner product $(u, v) = \int_\Lambda uv d\Lambda$ and the associated norm is given by $\|u\|_2 = (u, u)^{1/2}$. Also, the Sobolev space $W^{k,p}(\Lambda)$ contains the functions u with the weak derivatives $D^s u \in L^p(\Lambda)$ for $|s| \leq k$. The standard norm of the functions in this space denoted by $\|u\|_{k,p}$ is given by $\|u\|_{k,p} = (\sum_{i=0}^k \|u^{(i)}\|_p^p)^{1/p}$. When $p = 2$, it is denoted by $H^k(\Lambda)$ which is a Hilbert space. The function space $C_0^\infty(\Lambda)$ is the set of smooth functions on I with compact support with the closure in terms of $\|\cdot\|_2, H_0^2(\Lambda)$.

The initial conditions in (1.1)-(1.5) are not homogeneous. Let

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

then, the new problem takes the form

$$\partial_t^2 u + g(\partial_t^\alpha u) - \nabla \cdot (a \nabla u) + \beta u + K * u = f + s, \quad \forall (x, t) \in \Omega, \quad (3.1)$$



with the homogeneous initial conditions on Λ

$$u(x, 0) = 0, u_t(x, 0) = 0, \quad \forall x \in \Lambda, \tag{3.2}$$

and the BCs

$$u|_{\Gamma} = B(t), \quad \forall t \in I, \tag{3.3}$$

with the final condition at $t = T_f$ as

$$u(x, T_f) = \psi_{T_f}(x) := \tilde{\psi}_{T_f}(x) - h_0(x) - T_f h_1(x). \tag{3.4}$$

In our numerical experiments, we assume that $B(t) \equiv 0$, as given in [27].

Here, we present the variational form of (3.1)-(3.4), and an algorithm based on the SVR with an orthogonal kernel for simulating the problem numerically.

3.2. SVR for the inverse time-fractional wave Equation (3.1)-(3.4). First, let $\mathcal{L}u = f$, where \mathcal{L} is a linear differential operator on one-dimensional space with suitable homogeneous boundary conditions. The approximate solution using a kernel function written by dual variables is assumed as $u_M = \sum_{j=0}^M \alpha_j K(x, x_j)$, with a kernel $K(\cdot, \cdot)$ and the weights are determined by the following quadratic programming with equality constraints:

$$\begin{aligned} \min \quad & \frac{1}{2} w^T w + \frac{\gamma}{2} e^T e, \\ \text{s.t.} \quad & (w^T \mathcal{L}\phi(x), \psi_i) - (f(x), \psi_i) = e_i, \quad i = 0, \dots, M, \end{aligned} \tag{3.5}$$

Since the basis functions ϕ_i 's satisfy the BCs (3.3), these conditions are held for u_M as well. The following result gives a simple linear system to obtain the dual variables.

Theorem 3.1. *The parameters $\alpha = [\alpha_0, \dots, \alpha_M]^T$ as the weights of the kernels in the approximate solution are given by*

$$(W + \frac{1}{\gamma} I_M) \alpha = \mathbf{b}, \tag{3.6}$$

with $W_{k,i} = \sum_{j=0}^M (\mathcal{L}\phi_j, \psi_k)(\mathcal{L}\phi_j, \psi_i)$, $b_i = (f, \psi_i)$, $k, i = 0, \dots, M$.

Proof. First, let

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

then, by the KKT conditions, we get

$$w_k = \sum_{i=0}^M \alpha_i (\mathcal{L}\phi_k, \psi_i), \tag{3.7}$$

$$\gamma e_k + \alpha_k = 0, \tag{3.8}$$

$$(w^T \mathcal{L}\phi, \psi_k) - (f, \psi_k) - e_k = 0, \tag{3.9}$$

for $k = 0, \dots, M$. By (3.8) and (3.9), we have

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

Now, by (3.7), we get

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

This is (3.6). □



Here, we introduce an algorithm to numerically simulate the fractional inverse source problem in least squares SVR. The algorithm utilizes orthogonal Chebyshev polynomials as the kernel. We employ combinations of $\phi_m(x) = T_m(\frac{2}{\pi}x - 1)$ and $\psi_n(t) = T_n(2t/T - 1)$ for spatial and temporal approximation in space $[0, \pi]$, and time interval $[0, T]$. These combinations satisfy the specified boundary conditions.

For any function space X , e.g., $X = L^2(\Lambda)$, we define the following space-time spaces for $s > 0$:

$$\begin{aligned} H^s(I; X) &= \{v : \|v(\cdot, t)\|_X \in H^s(I)\}, \\ H_0^s(I; X) &= \{v : \|v(\cdot, t)\|_X \in H_0^s(I)\}, \\ B^s(\Omega) &= H^s(I; L^2(\Lambda)) \cap L^2(I; H_0^1(\Lambda)) \end{aligned}$$

with $\Omega = \Lambda \times I$. So, the weak formulation of (3.1)-(3.4), becomes as

$$(\partial_t^\alpha u, v)_\Omega + (g(\partial_t^\alpha u), v)_\Omega - (\nabla \cdot (a\nabla u), v)_\Omega + c(u, v)_\Omega + (K * u, v)_\Omega = (f, v)_\Omega + (s, v)_\Omega, \quad (3.10)$$

for $v \in B^{\alpha/2}(\Omega)$. The corresponding norm is given by

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

For $f \in L^2(\Omega)$, we aim to find $u \in B^{\alpha/2}(\Omega)$ such that

$$\mathcal{A}(u, v) = \mathcal{F}(v), \quad \forall v \in B^{\alpha/2}(\Omega). \quad (3.11)$$

Here, \mathcal{A} is a computationally simple form of (3.10) and the the functional \mathcal{F} are given by

$$\begin{aligned} \mathcal{A}(u, v) &= -(\partial_t u, \partial_t v)_\Omega + (g(\partial_t^\alpha u), v)_\Omega + ((a\nabla u), \nabla v)_\Omega + c(u, v)_\Omega + (K * u, v)_\Omega, \\ \mathcal{F}(v) &= (f, v)_\Omega + (s, v)_\Omega. \end{aligned} \quad (3.12)$$

Subsequently, we introduce flexibility to the margin in LS-SVR by incorporating an error tolerance for the constraints (3.11). The bilinear map in the weak form captures the interactions between the unknown function and derivatives using appropriate test functions. The variational formulation, based on this bilinear form, results in an algebraic system of equations that can be solved using numerical methods. The Lax-Milgram lemma guarantees the existence and uniqueness of the solution, along with the continuous dependence of the solution on the provided data. This holds true when the bilinear form \mathcal{A} in (3.12) is a continuous, bounded, and weakly coercive function, i.e.,

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

The term $(\tilde{g}(\partial_t^\alpha u), v)_\Omega$ may be written in an efficient form by the left- and right-sided fractional Riemann-Liouville derivatives. The right derivative in terms of Riemann-Liouville definition, ${}^{RL}D_T^{(\alpha)} f$, is defined as

$${}^{RL}\partial_{T_f}^\alpha u(t) = \frac{-1}{\Gamma(1-\alpha)} \frac{\partial}{\partial t} \int_t^{T_f} \frac{u(\tau)}{(\tau-t)^\alpha} d\tau. \quad (3.13)$$

Following the work of [12], we obtain the following results.

Lemma 3.2. *Let $s \in (0, 1)$, $u(t) \in H^s(I)$, $v(t) \in C_0^\infty(I)$. Then,*

$$({}^{RL}\partial_t^\alpha u, v)_I = (u, {}_{T_f}\partial_t^\alpha v)_I. \quad (3.14)$$



Proof. By (2.2), and using integration by parts, we get

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

□

We use (3.14) to get the following result.

Lemma 3.3. *Let $0 < \alpha < 1$, $u \in {}_0H^1(I)$, $v \in {}_0H^{\alpha/2}(I)$. Then,*

$$(\partial_t^\alpha u, v)_I = (\partial_t^{\alpha/2} u, {}_{T_f}\partial_t^{\alpha/2} v)_I. \tag{3.15}$$

Proof. First, let $v_n \in C_0^\infty(I)$ converging to $v \in {}_0H^{\alpha/2}(I)$ with respect to $H^{\alpha/2}(I)$ -norm, i.e.,

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

By (3.14), we change the order of derivatives from left to right, i.e., ∂_t^α to ${}_{,t}\partial_{T_f}^{\alpha/2}$ as $(\partial_t^\alpha u, v_n)_I = (\partial_t^{\alpha/2} u, {}_{,t}\partial_{T_f}^{\alpha/2} v_n)_I$. Now by the Cauchy-Schwartz inequality, we get

$$|(\partial_t^\alpha u, v_n)_I - (\partial_t^\alpha u, v)_I| \leq \|\partial_t^\alpha u\| \|v - v_n\|.$$

As $n \rightarrow \infty$, we get

$$(\partial_t^\alpha u, v)_I = (\partial_t^{\alpha/2} u, {}_{,t}\partial_{T_f}^{\alpha/2} v)_I.$$

So the proof is completed. □

Now, let $a \in \mathbb{R}$, then, the problem (3.1)-(3.4) is written by the weak formulation as follows. Find $u \in B^{\alpha/2}(\Omega)$ such that

$$-(\partial_t u, \partial_t v)_\Omega + (\partial_t^{\alpha/2} u, {}_{,t}\partial_{T_f}^{\alpha/2} v)_\Omega + a(\partial_x u, \partial_x v)_\Omega + c(u, v)_\Omega + (K * u, v)_\Omega = (f, v)_\Omega + (s, v)_\Omega, \tag{3.16}$$

for all $v \in B^{\alpha/2}(\Omega)$. Now, we present a polynomial-kernel algorithm for learning the weights in the ansatz in finite dimensional spaces. To discretize the problem (3.1)-(3.4) in the space-time domain, let $P_M^0(\Lambda)$, be the space of polynomials on Λ with degree less than or equal to M on Λ and $P_N^E(I)$ be the approximate space in time including the functions $v \in P_N$ with $v(t=0) = 0$. Now, we consider the $x-t$ trial space in which the approximate solution is sought in

$$S_L = P_M^0(\Lambda) \otimes P_N^E(I). \tag{3.17}$$

By letting a tolerance in the bilinear form, the LS-SVR approximation $u_L \in S_L$ is as follows:

$$\begin{aligned}
 \min & \frac{1}{2} w^T w + \frac{1}{2} d^T d + \frac{\gamma}{2} e^T e \\
 \text{s.t.} & \mathcal{A}(u_L, v_L) - \mathcal{F}(v_L) = e_i, \quad i = 0, \dots, M, \quad \forall v_L \in S_L,
 \end{aligned} \tag{3.18}$$



in which

$$u_L(x, t) = \sum_{m=0}^M \sum_{n=0}^N w_{m,n} \phi_m(x) \psi_n(t), \quad (3.19)$$

$$f_L(x) = \sum_{m=0}^M d_m \phi_m(x), \quad (3.20)$$

Here, $w = [w_{m,n}; m = 0, \dots, M, n = 0, \dots, N]^T$ and $d = [d_0, \dots, d_M]$ as vectors of unknown coefficients. We set $v_L = v_{i,j} = \phi_i(x) \psi_j(t) \in S_L$, to make use of the orthogonality in (3.18) and getting a sparse system for the constraints (3.18). We call this approach with $u_L \in S_L$ as (3.19), $f_L \in P_M^0(\Lambda)$ as (3.20), the Galerkin LS-SVR algorithm with Chebyshev polynomials for the inverse problem (3.1)-(3.4) by Chebyshev polynomials. The constraints in (3.18) are written as follows

$$-(\partial_t u_L, \partial_t v_L)_\Omega + (\partial_t^{\alpha/2} u_{L,t} \partial_{T_f}^{\alpha/2} v_L)_\Omega + a(\partial_x u_L, \partial_x v_L)_\Omega + c(u_L, v_L)_\Omega + (K * u_L, v_L)_\Omega = (f_L, v_L)_\Omega + (s, v_L)_\Omega + e. \quad (3.21)$$

The orthogonal system in inner products (3.21), facilitate the computations and the relation (3.21) takes the matrix form

$$-P^{0,0} U Q^{1,1} + P^{0,0} U Q^{\alpha/2, \alpha/2} + P^{0,0} U Q^{\alpha/2, \alpha/2} + a P^{1,1} U Q^{0,0} + c P^{0,0} U Q^{0,0} + P^{0,0} U Q_c^{0,0} = F,$$

where P, Q, U and F are given by

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

The orthogonality of Chebyshev kernels enhances numerical stability in LS-SVR, contributing to robust and reliable computations.

4. NUMERICAL EXPERIMENTS

In this section, we examine various linear and nonlinear test problems to demonstrate the convergence of the proposed space-time Galerkin LS-SVR algorithm with Chebyshev polynomials for inverse source problems. The objective is to investigate the convergence patterns of numerical solutions as the degrees of polynomials M and N , along with certain fractional derivatives α , increase. We represent errors using three measures L^∞ , L^2 , and H^1 -errors—plotted in semi-log scale to showcase exponential convergence.

The plots depict convergence behavior separately for spatial and temporal dimensions while keeping M and N fixed, respectively. The confirmation of exponential convergence arises from the linear behavior of log-plots of errors against polynomial degrees. Calculations for L^∞ , L^2 , and H^1 -errors are carried out using the following measures

$$L^\infty(u_L) \approx \max_{0 \leq i, j \leq \mathcal{M}} |U(x_i, t_j) - u_L(x_i, t_j)|, \quad (4.1)$$

$$L^2(u_L) \approx \left(\frac{1}{\mathcal{M}^2} \sum_{i,j=1}^{\mathcal{M}} |U(x_i, t_j) - u_L(x_i, t_j)|^2 \right)^{1/2}, \quad (4.2)$$

$$H^1(u_L) \approx (L^2(u_L) + L^2(\partial_x u_L))^{\frac{1}{2}}, \quad (4.3)$$

with $x_i = \frac{i}{\mathcal{M}}$, $i = 0, \dots, \mathcal{M}$.

The errors are calculated across an equally spaced grid within the x-t domain for all numerical illustrations, where the number of test points on each space and time direction is set to $\mathcal{M} = 100$. Similar measures are employed for



TABLE 1. 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 4.1.

M	L^∞	EOC	L^2	EOC	H^1	EOC
2	5.22e-02		2.36e-02		7.81e-02	
4	1.72e-03	4.9236	7.73e-04	4.9322	3.80e-03	4.3613
6	2.09e-05	10.8772	9.37e-06	10.8832	9.07e-05	9.2121
8	2.07e-07	16.0413	8.02e-08	16.5486	1.78e-06	13.6642
10	1.65e-09	21.6540	5.93e-10	21.9907	1.99e-08	20.1379
12	1.85e-11	24.6310	7.65e-12	23.8616	2.00e-10	25.2310

TABLE 2. L^∞, L^2 and H^1 errors and spatial convergence rates for the source function versus M with $N = 16$ and $\alpha = 1$ for Example 4.1.

M	$\ f - f_N\ _\infty$	EOC	$\ f - f_N\ _2$	EOC	$\ f - f_N\ _{H^1}$	EOC
2	5.04e-01		2.02e-01		5.55e-01	
4	5.78e-02	3.1243	1.96e-02	3.3654	1.06e-01	2.3884
6	2.64e-03	7.6115	8.54e-04	7.7278	7.07e-03	6.6777
8	6.95e-05	12.6431	2.21e-05	12.7028	2.37e-04	11.8032
10	1.06e-06	18.7460	3.35e-07	18.7736	4.29e-06	17.9784
12	1.08e-08	25.1560	3.45e-09	25.0972	3.25e-08	26.7813

assessing the errors of the source function f , focusing solely on spatial characteristics. To verify the convergence in time, we set

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

as the experimental order of convergence of the proposed algorithm as the number of training points N increases. Here, e_i is the error associated with N_i , $i = 1, 2, \dots$. For the numerical tables related to temporal convergence, a constant number of spatial basis functions, denoted as M , is assumed. The EOC formula, used as a measure of convergence in space, employs M_i instead of N_i and a fixed N .

Example 4.1. First, let $g(u) = u$, and

$$h_0 = 0, h_1 = \sin(x), u|_\Gamma = 0, a = 1, c = 2, K = -1,$$

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

on the space-time domain $(x, t) \in [0, \pi] \times [0, \pi/2]$. Let the exact solutions be $u = \sin(t) \sin(x)$, $f = -\sin(x)$ for the inverse source problem (3.1)-(3.4). The numerical errors and the experimental convergence rates of the LS-SVR are reported in terms of different measures, L^∞ , L^2 , and H^1 as M increases with fixed $N = 16$ or N increases with fixed $M = 15$ to demonstrate the convergence of the proposed method in space and time, respectively.

To show the spectral convergence, we illustrate the errors in the logarithmic scale in Figures 1-6. The linear semi-log plots establish the exponential convergence. This fast convergence in both space and time as shown in the figures occurs for the wave and source functions, simultaneously.

The authors in [27] have used a Landweber-type algorithm and report the errors with 7 iterations for $\alpha = 1$ as

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

$$\|f_7 - f\| \leq 1.8 \times 10^{-2}.$$



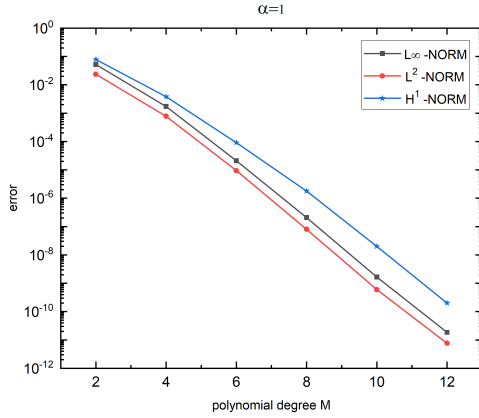


FIGURE 1. L^∞ , L^2 and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 1$ for Example 4.1.

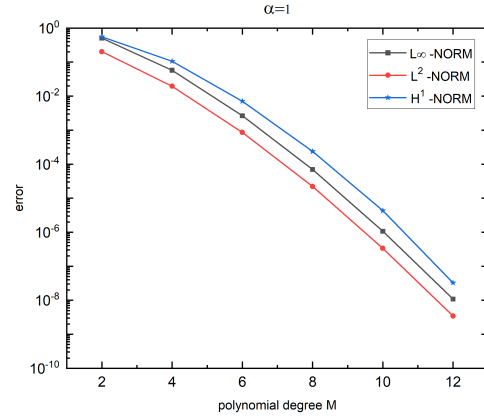


FIGURE 2. L^∞ , L^2 and H^1 errors for the source function versus M with $N = 16$ and $\alpha = 1$ for Example 4.1.

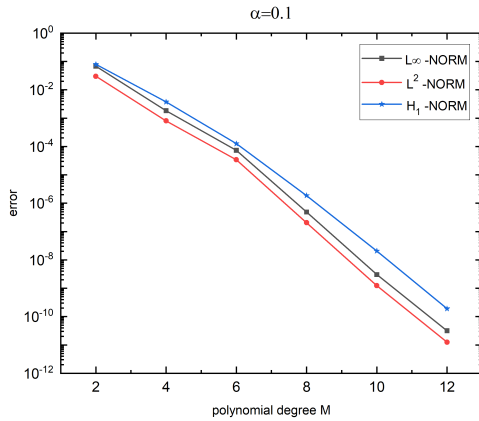


FIGURE 3. L^∞ , L^2 and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 0.1$ for Example 4.1.

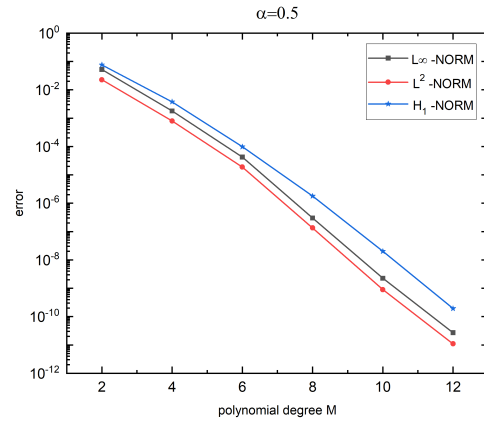


FIGURE 4. L^∞ , L^2 and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 0.5$ for Example 4.1.

On the other hand, using the proposed space-time LS-SVR with Chebyshev polynomials, the errors for the wave and source functions, after 7 iterations are given by

$$\begin{aligned} \|u_7(x, T) - \psi_T(x)\| &\leq 1.31 \times 10^{-5}, \\ \|f_7 - f\| &\leq 2.27 \times 10^{-3}. \end{aligned}$$

The authors in [27] have only discussed the problem in integer order derivatives and there are no data for fractional cases.

Figures 3, 4, and 5 show the spectral convergence of the method in which a linear semi-log plot verifies the exponential convergence for the wave function for fractional orders $\alpha = 0.1, 0.5, 0.9$.



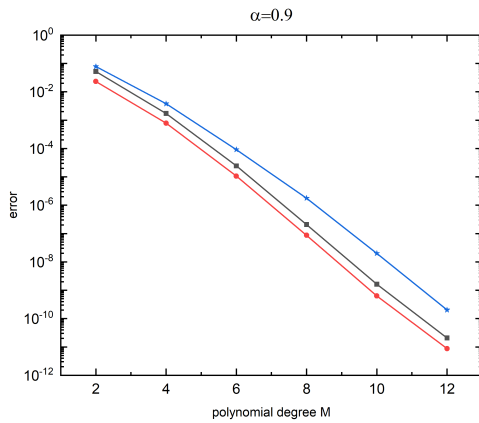


FIGURE 5. L^∞, L^2 and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 0.9$ for Example 4.1.

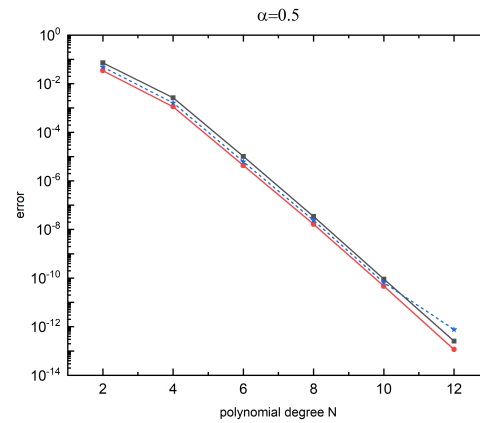


FIGURE 6. L^∞, L^2 and H^1 errors for the wave function versus N with $M = 15$ and $\alpha = 0.5$ for Example 4.1.

TABLE 3. 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 4.2.

M	L^∞	EOC	L^2	EOC	H^1	EOC
2	6.86e-02		3.04e-02		7.88e-02	
4	1.81e-03	5.2441	7.93e-04	5.2606	3.71e-03	4.4087
6	7.13e-05	7.9765	3.42e-05	7.7531	1.25e-04	8.3619
8	4.37e-07	17.7095	1.72e-07	18.3970	1.77e-06	14.7987
10	2.40e-09	23.3234	9.62e-10	23.2417	2.00e-08	20.0902
12	2.81e-11	24.3935	1.09e-11	24.5734	1.88e-10	25.5979

The temporal convergence of the LS-SVR scheme for the wave function, versus N with $M = 15$ is illustrated in Figure 6.

The following example demonstrates the convergence of the proposed method with the Chebyshev polynomial kernel in both space and time for the wave and source functions in the presence of nonlinearity.

Example 4.2. We consider the nonlinear case $g(u) = u^2$ for the inverse source problem with

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

on the same domain $[0, \pi] \times [0, \pi/2]$ as for Example 4.1 with the exact solution $u = \sin(t) \sin(x)$, $f = -\sin(x)$. L^∞ , L^2 , and H^1 norms are used to measure the accuracy of the numerical results of the proposed algorithm and to see the convergence as M increases while N is set to $N = 16$ and $\alpha = 1$. This is to verify the spatial convergence for the unknown wave and source in Tables 3 and 4, respectively.

Figure 8 and 7 represent a linear model of the logarithmic scale error for the wave and source, respectively. The semi-log plot approves the exponential convergence in both cases.

The authors in [27] implemented the Landweber-type algorithm with 13 iterations to the same problem to get

$$\begin{aligned} \|u_{13}(x, T) - \psi_T(x)\| &\leq 0.001, \\ \|f_{13} - f\| &\leq 0.02. \end{aligned}$$



TABLE 4. 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 4.2.

M	$\ f - f_N\ _\infty$	EOC	$\ f - f_N\ _2$	EOC	$\ f - f_N\ _{H^1}$	EOC
2	4.76e-01		1.93e-01		4.64e-01	
4	5.48e-02	3.1187	1.83e-02	3.3987	9.71e-02	2.2566
6	2.63e-03	7.4894	8.99e-04	7.4319	7.62e-03	6.2767
8	6.95e-05	12.6300	2.23e-05	12.8499	2.49e-04	11.8919
10	1.06e-06	18.7460	3.35e-07	18.8139	4.30e-06	18.1894
12	1.08e-08	25.1560	3.45e-09	25.0972	3.58e-08	26.2636

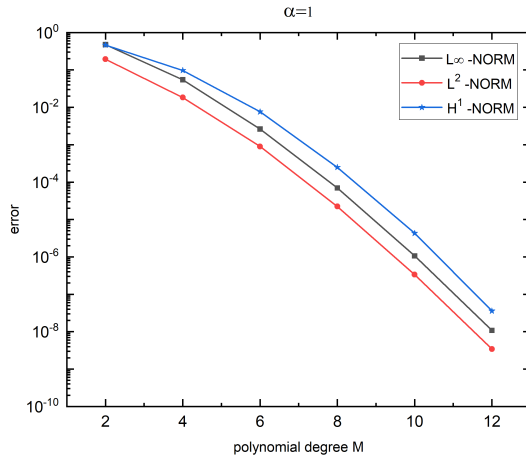


FIGURE 7. L^∞, L^2 and H^1 errors for the source function versus M with $N = 16$ and $\alpha = 1$ for Example 4.2.

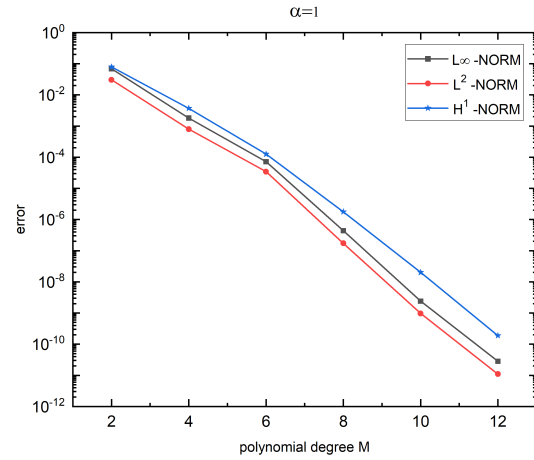


FIGURE 8. L^∞, L^2 and H^1 errors for the wave function versus M with $N = 16$ and $\alpha = 1$ for Example 4.2.

However, the Galerkin LS-SVR with Chebyshev polynomials provide high accurate results, e.g., we see

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

As seen in Figure 9, the linear logarithmic scale plot shows the exponential convergence in time for the wave function.

The outcomes reveal that the proposed LS-SVR based on Chebyshev kernels achieves highly accurate approximations in linear and nonlinear cases of fractional inverse wave problems and it demonstrates robust convergence behavior which is seen in the figures.

5. CONCLUSION

This study focused on an inverse source problem related to the time-fractional wave equation. We introduced a space-time Galerkin LS-SVR with Chebyshev kernels for numerically simulating the problem, utilizing an orthogonal polynomial kernel in both space and time variables. We formulated the method as a quadratic programming problem and then converted it into a system of linear algebraic equations through the introduction of dual variables. The resulting system's structure was also examined, and several numerical test problems were provided to illustrate the effectiveness of the proposed method. The results show a good exponential convergence of the proposed method for



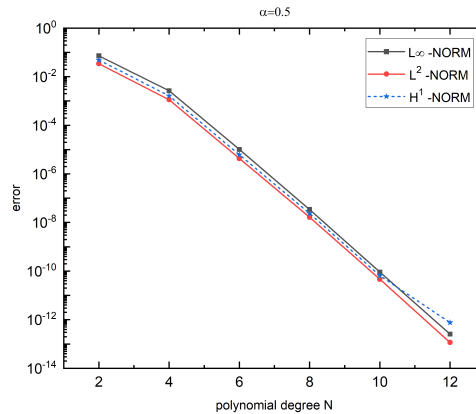


FIGURE 9. L^∞ , L^2 and H^1 errors for the wave function versus N with $M = 15$ and $\alpha = 0.5$ for Example 4.2.

different fractional orders and in both the linear and nonlinear cases, temporally and spatially. A further study requires the investigation of the method in the presence of noise and non-smooth data for the inverse source problems of time and space fractional wave equations.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous referees for valuable comments that helped the authors 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 is not applicable to this paper 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] R. Ashurov and Y. Faiziev, *Inverse problem for finding the order of the fractional derivative in the wave equation*, Mathematical Notes, 110(5-6) (2021), 842-852.
- [2] C. Campbell and Y. Ying, *Learning with support vector machines*, Springer Nature, 2022.
- [3] K. Diethelm, *The analysis of fractional differential equations: An application-oriented exposition using differential operators of Caputo type*, Springer, 2010.
- [4] F. Ernst and A. Schweikard, *Fundamentals of Machine Learning: Support Vector Machines Made Easy*, utb GmbH, 2020.
- [5] B. Guo, *Spectral methods and their applications*, World Scientific, 1998.
- [6] A. Hasanov, *Some new classes of inverse coefficient problems in non-linear mechanics and computational material science*, International Journal of Non-Linear Mechanics, 46(5) (2011), 667-684.
- [7] I. Häggström, C. Schmidlein, G. Campanella, and T. Fuchs, *DeepPET: A deep encoder-decoder network for directly solving the PET image reconstruction inverse problem*, Medical image analysis, 54 (2019), 253-262.
- [8] M. Idemen and A. Alkumru *On an inverse source problem connected with photo-acoustic and thermo-acoustic tomographies*, Wave Motion, 49(6) (2012), 595-604.



- [9] S. Kazemi and A. Tari, *Collocation Method for Solving Two-Dimensional Fractional Volterra Integro-Differential Equations*, Iran J. Sci. Technol. Trans. Sci., 46 (2022), 1629–1639.
- [10] T. Khan and A. Thomas, *Inverse problem in refractive index based optical tomography*, Inverse problems, 22(4) (2006), 1121-1137.
- [11] N. Kinash and J. Janno, *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.
- [12] X. Li and C. Xu, *A space-time spectral method for the time-fractional diffusion equation*, SIAM journal on numerical analysis, 47(3) (2009), 2108-2131.
- [13] Y. Liu and m. Xie, *Rebooting data-driven soft-sensors in process industries: A review of kernel methods*, Journal of Process Control, 89 (2020), 58-73.
- [14] S. Liu and J. Zhang, *Machine-learning-based prediction of regularization parameters for seismic inverse problems*, Acta Geophysica, 69 (2021), 809-820.
- [15] S. Lloyd, C. Schaal, and C. Jeong, *Inverse modeling and experimental validation for reconstructing wave sources on a 2D solid from surficial measurement*, Ultrasonics, 128(6) (2023), 106880.
- [16] F. Mainardi, *Fractional calculus in wave propagation problems*, arXiv preprint arXiv:1202.0261. 2012.
- [17] A. Mohammadi and A. Tari, *A new approach to numerical solution of the time-fractional KdV-Burgers equations using least squares support vector regression*, Journal of Mathematical Modeling, 12(4) (2024), 583-602.
- [18] A. Mohammadi and A. Tari, *A space-time least-squares support vector regression scheme for inverse source problem of the time-fractional wave equation*, Iranian Journal of Numerical Analysis and Optimization, 14(4) (2024), 1037-1068.
- [19] M. Mojahedfar and A. Tari, *Solving two-dimensional fractional integro-differential equations by Legendre wavelets*, Bull. Iranian Math. Soc. 43(7) (2017), 2419-2435.
- [20] A. Pakniyat, K. Parand, and M. Jani, *Least squares support vector regression for differential equations on unbounded domains*, Chaos, Solitons & Fractals, 151 (2021), 111232.
- [21] I. Podlubny, *Fractional differential equations*, Academic Press, New York, 1999.
- [22] L. Qiu, C. Hu, and Q. Qin, *A novel homogenization function method for inverse source problem of nonlinear time-fractional wave equation*, Applied Mathematics Letters, 109 (2020), 106554.
- [23] B. Reddy, *Introductory functional analysis: with applications to boundary value problems and finite elements*, Springer Science & Business Media, 2013.
- [24] K. Sakamoto and M. Yamamoto, *Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems*, Journal of Mathematical Analysis and Applications, 382(1) (2011), 426-447.
- [25] J. Shen and C. T. Sheng, *An efficient space-time method for time-fractional diffusion equation*, Journal of Scientific Computing, 81 (2019), 1088-1110.
- [26] J. Shen, T. Tang, and L. Wang, *Spectral methods: algorithms, analysis and applications*, Springer, 2011.
- [27] M. Slodicka and L. Seliga, *An inverse source problem for a damped wave equation with memory*, Journal of Inverse and Ill-posed problems, 24(2) (2016), 111-122.
- [28] J. Suykens, T. Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least squares support vector machines*, Singapore: World Scientific Publishing Company, 2002.
- [29] J. Suykens and J. Vandewalle, *Least squares support vector machine classifiers*, Neural Process Lett, 9(3) (1999), 293-300.

