

Robust fast method for variable-order time-fractional diffusion equations without regularity assumptions

Jia-li Zhang¹ · Zhi-wei Fang² · Hai-wei Sun³

Received: date / Accepted: date

Abstract In this paper, we develop a robust fast method for mobile-immobile variable-order (VO) time-fractional diffusion equations (tFDEs), superiorly handling the cases of small or vanishing lower bound of the VO function. The valid fast approximation of the VO Caputo fractional derivative is obtained using integration by parts and the exponential-sum-approximation method. Compared with the general direct method, the proposed algorithm (*RF-L1* formula) reduces the acting memory from $\mathcal{O}(n)$ to $\mathcal{O}(\log^2 n)$ and computational cost from $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log^2 n)$, respectively, where n is the number of time levels. Then *RF-L1* formula is applied to construct the fast finite difference scheme for the VO tFDEs, which sharp decreases the memory requirement and computational complexity. The error estimate for the proposed scheme is studied only under some assumptions of the VO function, coefficients, and the source term, but without any regularity assumption of the true solutions. Numerical experiments are presented to verify the effectiveness of the proposed method.

Keywords variable-order Caputo fractional derivative · exponential-sum-approximation method · fast algorithm · convergence

Mathematics Subject Classification (2010) 35R11, 65M06, 65M12

1 Introduction

Fractional operators have been extensively studied due to their broad applications in both mathematics and physical science. Numerous researchers revealed that the fractional calculus can

This work is supported in part by research grants of the Science and Technology Development Fund, Macau SAR (file no. 0122/2020/A3), and University of Macau (file no. MYRG2020-00224-FST).

1
Department of Mathematics, University of Macau, Macao
E-mail: zhangjl2628@163.com

2
Department of Mathematics, University of Macau, Macao
E-mail: fzw913@yeah.net

3
Corresponding Author. Department of Mathematics, University of Macau, Macao
E-mail: HSun@um.edu.mo

better characterize complex phenomena in fields such as the biology, the ecology, the diffusion, and the control system [3, 18, 23, 25, 30, 31]. In particular, the variable-order (VO) fractional operators are more efficient since many important dynamical problems exhibit the order of the fractional operator varying with time, space, or some other variables; see [20, 36, 37]. Recently, the VO fractional derivatives have been widely applied to model phenomena in fields of science and engineering; see, for details, [6–8, 14, 15, 27, 29, 33, 35, 46]. In this paper, we consider the VO mobile-immobile time-fractional diffusion equations (tFDEs) [11, 44, 45]

$$\frac{\partial}{\partial t}u(x, t) + \zeta {}_0^C\mathcal{D}_t^{\alpha(t)}u(x, t) = \frac{\partial}{\partial x}\left[p(x)\frac{\partial u(x, t)}{\partial x}\right] + f(x, t), \quad x \in \Omega, \quad t \in (0, T], \quad (1)$$

$$u(x, 0) = \varphi(x), \quad x \in \overline{\Omega}, \quad (2)$$

$$u(x, t) = 0, \quad x \in \partial\Omega, \quad t \in (0, T], \quad (3)$$

where $\Omega = (x_l, x_r)$, $\zeta > 0$ is the mobile/immobile capacity coefficient, and $f(x, t)$ is the source term. Moreover, $f(x, t)$, $\varphi(x)$ and $p_* \leq p(x) \leq p^*$ are given sufficiently smooth functions. The VO Caputo fractional derivative is defined by [7]

$${}_0^C\mathcal{D}_t^{\alpha(t)}u(x, t) := \frac{1}{\Gamma(1 - \alpha(t))} \int_0^t \frac{u'(x, \tau)}{(t - \tau)^{\alpha(t)}} d\tau, \quad (4)$$

where $0 \leq \alpha_* \leq \alpha(t) \leq \alpha^* < 1$ is the VO function depending on the time variable $t \in [0, T]$ and $\Gamma(\cdot)$ is the Gamma function. The VO tFDEs (1)–(3) describe the dynamic mass exchange between mobile and immobile phases and thus improve the modeling of anomalously diffusive transport [28, 35]. Several papers have considered numerical methods for the VO tFDEs; see the references in [9, 13, 38, 43]. However, most papers ignored the possible presence of an initial layer in the solution near $t = 0$ and presented convergence analyses that makes the unrealistic assumption that the true solution is smooth on the closed domain. Nevertheless, it is well known that the solutions to tFDEs exhibit initial singularities that may affect the accuracy of the numerical approximations. In [44, 45], the authors showed that such singularity may not be physical relevant in the diffusion processes and could be eliminated in VO functional models by imposing the integer limit of the VO function at $t = 0$. More precisely, the solution has full regularity like its integer-order analogue if the VO function has an integer limit at $t = 0$; or exhibits singular behaviors at $t = 0$ if the VO function has a non-integer value at $t = 0$. Taking into account the initial behavior of the VO tFDEs, a fully discretized finite element approximation to (1)–(3) is developed and analyzed [44].

As a result of the nonlocality of the fractional operators, using $L1$ formula [19, 22, 26, 34, 44] to discretize the VO Caputo fractional derivative is too expensive in storage and complexity, which requires $\mathcal{O}(n)$ storage and $\mathcal{O}(n^2)$ computational cost, where n is the total number of the time levels. The constant-order (CO) fractional operators also suffer from such difficulty. Hence many efforts have been made to speed up the evaluation of the CO Caputo fractional derivative [2, 4, 12, 16, 17, 21, 24, 41]. Nevertheless, the coefficient matrices of the numerical schemes for the VO fractional problems lose the Toeplitz-like structure and the VO fractional derivative is no longer a convolution operator. Those fast methods for the CO fractional derivative cannot be directly applied to VO cases. Limited fast methods have been presented so far for the VO fractional derivative. Recently, Fang, Sun and Wang [10] developed a fast algorithm for the VO Caputo fractional derivative based on a shifted binary block partition and uniform polynomial approximations. In [42], a fast $L1$ formula ($F-L1$ formula) is proposed using the exponential-sum-approximation (ESA) technique for the VO Caputo fractional derivative. $F-L1$ formula achieves the optimal convergence and significantly reduces the storage requirement to $\mathcal{O}(\log^2 n)$ and the computational cost to $\mathcal{O}(n \log^2 n)$, respectively. However,

a small value of α_* causes unreliable parameters in the fast algorithm, which leads to inaccurate approximations in the view of scientific computation. Even worse, F - $L1$ formula fails to approximate the VO fractional derivative with a vanishing α_* since $\alpha_* = 0$ indicates an infinite parameter in the algorithm.

In this paper, to overcome the weakness mentioned above, we develop a valid method to accelerate the approximation of the VO Caputo fractional derivative. At the time level t_k , using integration by parts, we transform the singular kernel in (4) from $(t_k - \tau)^{-\alpha(t_k)}$ to $(t_k - \tau)^{-1-\alpha(t_k)}$. It makes sense that $(t_k - \tau)^{-1-\alpha(t_k)}$ can be approached by a linear combination of exponentials based on the ESA technique, which supplies an opportunity to construct the fast algorithm. Compared with $L1$ formula, it reduces the acting memory from $\mathcal{O}(n)$ to $\mathcal{O}(\log^2 n)$ and computational cost from $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log^2 n)$. Significantly, although RF - $L1$ formula and F - $L1$ formula are both based on the ESA method, RF - $L1$ formula provides a powerful way to fast approximate the VO Caputo fractional derivative, which can attack the problem efficiently even if $\alpha_* = 0$ while F - $L1$ formula cannot work. Then RF - $L1$ formula is applied to construct a fast finite difference scheme (RF - $L1$ scheme) for the VO tFDEs, which sharp decreases the memory requirement and computational complexity. We present the optimal order convergence rate of the proposed scheme, assuming only the smoothness of the coefficients, source term, the VO function and the spatial domain, but not the regularity of the true solution. The numerical experiments show that RF - $L1$ scheme achieves temporal first-order accuracy with less CPU time and memory than the existing methods. Specifically, RF - $L1$ scheme performs well for $\alpha_* = 0, 0.05$ while F - $L1$ scheme fails to solve the problem.

The structure of the paper is as follows. In Section 2, we refer the regularity and well-posedness of the solution of (1)–(3) and recall $L1$ and F - $L1$ formulas. In Section 3, we propose RF - $L1$ formula for the VO Caputo fractional derivative based on the ESA technique. In Section 4, RF - $L1$ formula is applied to construct a fast difference scheme to solve the VO tFDEs (1)–(3). The convergence of the scheme is analysed without the smoothness assumption of the true solution. In Section 5, numerical results are reported to demonstrate the efficiency of the proposed scheme. Concluding remarks are given in Section 6.

2 Preliminaries

2.1 Well-posedness

It is well known that the solutions to tFDEs exhibit initial singularities that may affect the accuracy of the numerical approximations. The regularity of all typical solutions of (1)–(3) is investigated in [44, 45].

In this paper, we use the standard Hilbert space $L_2(\Omega)$ with norm $\|\cdot\|_{L_2}$ and inner product $(\cdot, \cdot)_{L_2}$. For convenience we may drop the subscript L_2 in $(\cdot, \cdot)_{L_2}$ and $\|\cdot\|_{L_2}$ as well as the notation Ω when no confusion occurs. Moreover, c denotes generic positive constants that are independent of T and any mesh used to solve (1)–(3) numerically. Note that c may be assumed to be different values at different occurrences. For the remainder of this paper, we make the following assumptions to ensure the essential feature of the solution.

Condition A: $\alpha(t) \in C[0, T]$, $0 \leq \alpha(t) \leq \alpha^* < 1$ on $[0, T]$, $\lim_{t \rightarrow 0^+} (\alpha(t) - \alpha(0)) \ln t$ exists.

Condition B: $p(x) \in C^1(\overline{\Omega})$.

Next, we introduce some notations. Let $\{(\xi_i, \phi_i) : i = 1, 2, \dots\}$ be the eigenvalues and eigenfunctions for the Sturm-Liouville problem

$$\begin{aligned}\mathcal{L}\phi_i(x) &= \xi_i\phi_i(x), \quad x \in \Omega, \\ \phi_i(x) &= 0, \quad x \in \partial\Omega,\end{aligned}$$

where the eigenfunctions are normalised by requiring $\|\phi_i\| = 1$. The fractional power operator \mathcal{L}^γ for any $\gamma \geq 0$ and the fractional Sobolev spaces can be defined by the theory of sectorial operators [1, 39]

$$\begin{aligned}\mathcal{L}^\gamma v &:= \sum_{i=1}^{\infty} \xi_i^\gamma (v, \phi_i) \phi_i, \quad v = (v, \phi_i) \phi_i, \\ \check{H}^\gamma(\Omega) &= \{v \in L_2(\Omega) : |v|_{\check{H}^\gamma}^2 := (\mathcal{L}^\gamma v, v) = \sum_{i=1}^{\infty} \xi_i^\gamma (v, \phi_i)^2 < \infty\},\end{aligned}$$

with the norm being defined by $\|v\|_{\check{H}^\gamma} = (\|v\|^2 + |v|_{\check{H}^\gamma}^2)^{1/2}$. Moreover, $\check{H}^\gamma(\Omega)$ is a subspace of the fractional Sobolev space $H^\gamma(\Omega)$ characterized by [1, 39]

$$\check{H}^\gamma(\Omega) = \{v \in H^\gamma(\Omega) : \mathcal{L}^s v(x) = 0, x \in \partial\Omega, s < \gamma/2\},$$

and the seminorms $|v|_{\check{H}^\gamma}$ and $|v|_{H^\gamma}$ are equivalent in \check{H}^γ .

We cite the well-posedness and smoothing properties of problem (1)–(3).

Lemma 1 [44, 45] *If **Conditions A** and **B** hold and $\varphi \in \check{H}^{\gamma+2}$, $f \in H^1(0, T; \check{H}^\gamma)$ for $\gamma > 1/2$. Then problem (1)–(3) has unique solution $u \in C^1([0, T]; \check{H}^\gamma)$ and*

$$\|u\|_{C^1([0, T]; \check{H}^s)} \leq c(\|\varphi\|_{\check{H}^{s+2}} + \|f\|_{H^1(0, T; \check{H}^s)}), \quad 0 \leq s \leq \gamma.$$

Lemma 2 [44, 45] *If **Conditions A** and **B** hold, $\alpha(t) \in C^1[0, T]$ and $\varphi \in \check{H}^{s+4}$, $f \in H^1(0, T; \check{H}^{s+2}) \cap H^2(0, T; \check{H}^s)$ for $s \geq 0$. If $\alpha(0) > 0$, then $u \in C^2([0, T]; \check{H}^s)$ and for $0 < \theta \ll 1$,*

$$\|u\|_{C^2([\theta, T]; \check{H}^s)} \leq c\theta^{-\alpha(0)} (\|\varphi\|_{\check{H}^{s+4}} + \|f\|_{H^1(0, T; \check{H}^{s+2})} + \|f\|_{H^2(0, T; \check{H}^s)});$$

if $\alpha(0) = 0$, then $u \in C^2([0, T]; \check{H}^s)$ and

$$\|u\|_{C^2([0, T]; \check{H}^s)} \leq c(\|\varphi\|_{\check{H}^{s+4}} + \|f\|_{H^1(0, T; \check{H}^{s+2})} + \|f\|_{H^2(0, T; \check{H}^s)}).$$

To be specific, in addition to the assumption that the data in (1)–(3) are sufficiently smooth, the condition $\alpha(0) = 0$ admits a temporal C^2 solution with respect to the spatial norm $\|\cdot\|_{\check{H}^s(\Omega)}$ for large s defined by the eigenpairs of the diffusion operator in (1)–(3). Otherwise, if $\alpha(0) > 0$, then $u_{tt} \in C(0, T]$ satisfies the pointwise-in-time estimate $\|u_{tt}\|_{\check{H}^s(\Omega)} \leq ct^{-\alpha(0)}$, which indicates the singularity as that of the solutions to the CO tFDEs.

2.2 Basic approximations to VO Caputo fractional derivative

In this section, we first recall $L1$ formula and $F-L1$ formula for the VO Caputo fractional derivative (5) with $t \in [0, T]$. For a positive integer n , let $\Delta t = T/n$ be the time step and we further define $t_k = k\Delta t$ for $k = 0, 1, \dots, n$. At each time level t_k , we have

$${}_0^C \mathcal{D}_t^{\alpha(t_k)} u(x, t_k) = \frac{1}{\Gamma(1 - \alpha(t_k))} \int_0^{t_k} \frac{u'(x, \tau)}{(t_k - \tau)^{\alpha(t_k)}} d\tau. \quad (5)$$

For convenience, denote $\alpha_k = \alpha(t_k)$. To discretize the VO Caputo fractional derivative, denote the linear interpolation for $u(x, \tau)$ over the interval $[t_{k-1}, t_k]$ with $1 \leq k \leq n$ by

$$L_{1,k}(\tau) = \frac{t_k - \tau}{\Delta t} u(x, t_{k-1}) + \frac{\tau - t_{k-1}}{\Delta t} u(x, t_k), \quad \tau \in [t_{k-1}, t_k],$$

giving a first-order approximation to $u'(x, \tau)$ on $[t_{k-1}, t_k]$ by

$$L'_{1,k}(\tau) = \frac{u(x, t_k) - u(x, t_{k-1})}{\Delta t}.$$

Then, the piecewise approximating function is defined by

$$L_1(\tau) = \{L_{1,k}(\tau) | \tau \in [t_{k-1}, t_k], k = 1, 2, \dots, n\}.$$

Thus, $L1$ formula to (5) is obtained as [44]

$$\begin{aligned} {}_0 \mathcal{D}_t^{\alpha_k} u(x, t_k) &= \frac{1}{\Gamma(1 - \alpha_k)} \int_0^{t_k} \frac{L'_1(\tau)}{(t_k - \tau)^{\alpha_k}} d\tau \\ &= \frac{\Delta t^{-\alpha_k}}{\Gamma(2 - \alpha_k)} \left(a_0^{(k)} u(x, t_k) - \sum_{l=1}^{k-1} \left(a_{k-l-1}^{(k)} - a_{k-l}^{(k)} \right) u(x, t_l) - a_{k-1}^{(k)} u(x, t_0) \right), \end{aligned} \quad (6)$$

where $a_l^{(k)} = (l+1)^{1-\alpha_k} - l^{1-\alpha_k}$. The truncation error of $L1$ formula is estimated by the following lemma.

Lemma 3 (see [44]) *Suppose **Conditions A** and **B** hold, $\alpha(t) \in C^1[0, T]$ and $\varphi \in \check{H}^4$, $f \in H^1(0, T; \check{H}^2) \cap H^2(0, T; L_2)$. Let the VO Caputo fractional derivative at t_k be as (5) and $L1$ formula be defined by (6). If $\alpha(0) > 0$,*

$$\left\| {}_0^C \mathcal{D}_t^{\alpha_k} u(x, t_k) - {}_0 \mathcal{D}_t^{\alpha_k} u(x, t_k) \right\| \leq ck^{-\alpha^*} \Delta t^{1-\alpha^*};$$

if $\alpha(0) = 0$,

$$\left\| {}_0^C \mathcal{D}_t^{\alpha(t)} u(x, t) - {}_0 \mathcal{D}_t^{\alpha(t)} u(x, t) \right\|_{\widehat{L}_\infty(0, T; L_2)} := \max_{1 \leq k \leq n} \left\| {}_0^C \mathcal{D}_t^{\alpha_k} u(x, t_k) - {}_0 \mathcal{D}_t^{\alpha_k} u(x, t_k) \right\| \leq c\Delta t.$$

Due to the nonlocality of the VO fractional derivative, using $L1$ formula to calculate the value at the current time level, it needs to compute the sum of a series including the values of all previous time levels. Therefore, $L1$ formula requires large memory and computational cost. A fast algorithm (denoted by $F-L1$ formula) is developed in [42] to discretize the VO Caputo fractional derivative (4). For the expected accuracy ϵ , $F-L1$ formula for the VO Caputo fractional derivative is given by

$${}_0^F \mathcal{D}_t^{\alpha_k} u(x, t_k) = \frac{T^{-\alpha_k}}{\Gamma(1 - \alpha_k)} \sum_{i=\underline{N}+1}^{\overline{N}} \tilde{\theta}_i^{(k)} \tilde{F}_{k,i} + \frac{u(x, t_k) - u(x, t_{k-1})}{\Delta t^{\alpha_k} \Gamma(2 - \alpha_k)},$$

where $\tilde{F}_{1,i} = 0$ and

$$\tilde{F}_{k,i} = e^{-\tilde{\lambda}_i \Delta t/T} \tilde{F}_{k-1,i} + T \frac{e^{-\tilde{\lambda}_i \Delta t/T} - e^{-2\tilde{\lambda}_i \Delta t/T}}{\tilde{\lambda}_i \Delta t} \left(u(x, t_{k-1}) - u(x, t_{k-2}) \right), \quad k = 2, 3, \dots, n,$$

in which the quadrature exponents and weights are given by

$$\tilde{\lambda}_i = e^{i\tilde{h}}, \quad \tilde{\theta}_i^{(k)} = \frac{\tilde{h} e^{\alpha_k i \tilde{h}}}{\Gamma(\alpha_k)}$$

with

$$\begin{aligned} \tilde{h} &= \frac{2\pi}{\log 3 + \alpha_* \log(\cos 1)^{-1} + \log \epsilon^{-1}}, \\ \underline{N} &= \left\lceil \frac{1}{\tilde{h}} \frac{1}{\alpha_*} (\log \epsilon + \log \Gamma(1 + \alpha_*)) \right\rceil, \\ \overline{N} &= \left\lceil \frac{1}{\tilde{h}} \left(\log \frac{T}{\Delta t} + \log \log \epsilon^{-1} + \log \alpha_* + 2^{-1} \right) \right\rceil. \end{aligned} \quad (7)$$

In addition,

$${}_{\mathcal{F}}\mathcal{D}_t^{\alpha_1} u(x, t_1) = \frac{u(x, t_1) - u(x, t_0)}{\Delta t^{\alpha_1} \Gamma(2 - \alpha_1)}.$$

Compared with $L1$ formula, F - $L1$ formula reduces the storage requirement from $\mathcal{O}(n)$ to $\mathcal{O}(\log^2 n)$ and the computational cost from $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log^2 n)$. It provides an efficient tool to approximate the VO Caputo fractional derivative. However, the fast method proposed in [42] cannot deal with the problem with a small α_* . In fact, a small value of α_* requires a large \underline{N} and even $\underline{N} \rightarrow \infty$ as $\alpha_* \rightarrow 0$, which affects the accuracy of the approximation. In order to overcome the shortage, we develop a robust fast approximation, which is called RF - $L1$ formula for the VO Caputo fractional derivative.

3 RF - $L1$ formula to VO Caputo fractional derivative

We first split the integral in (6) into two parts. Then that can be decomposed as

$$\begin{aligned} {}_0^C \mathcal{D}_t^{\alpha_k} u(x, t_k) &\approx \frac{1}{\Gamma(1 - \alpha_k)} \int_0^{t_{k-1}} \frac{L'_1(\tau)}{(t_k - \tau)^{\alpha_k}} d\tau + \frac{1}{\Gamma(1 - \alpha_k)} \int_{t_{k-1}}^{t_k} \frac{L'_{1,k}(\tau)}{(t_k - \tau)^{\alpha_k}} d\tau \\ &:= I_{his}(t_k) + I_{loc}(t_k), \end{aligned} \quad (8)$$

where we call $I_{his}(t_k)$ and $I_{loc}(t_k)$ the history and local parts, respectively. Since the local part contributes few memory and computational cost compared with the history part, we keep the local part be as in (8). For the history part, noting that $I_{his}(t_1) = 0$, we have

$${}_{\mathcal{R}}\mathcal{F}_0 \mathcal{D}_t^{\alpha_1} u(x, t_1) := I_{loc}(t_1) = {}_0 \mathcal{D}_t^{\alpha_1} u(x, t_1). \quad (9)$$

Using the integration by parts for $k = 2, 3, \dots, n$, we have

$$\begin{aligned} I_{his}(t_k) &= \frac{1}{\Gamma(1 - \alpha_k)} \left(\frac{u(x, t_{k-1})}{\Delta t^{\alpha_k}} - \frac{u(x, t_0)}{t_k^{\alpha_k}} - \alpha_k \int_0^{t_{k-1}} \frac{L_1(\tau)}{(t_k - \tau)^{1+\alpha_k}} d\tau \right) \\ &= \frac{1}{\Gamma(1 - \alpha_k)} \left(\frac{u(x, t_{k-1})}{\Delta t^{\alpha_k}} - \frac{u(x, t_0)}{t_k^{\alpha_k}} - \frac{\alpha_k}{T^{1+\alpha_k}} \int_0^{t_{k-1}} L_1(\tau) \left(\frac{t_k - \tau}{T} \right)^{-1-\alpha_k} d\tau \right). \end{aligned} \quad (10)$$

Noting that $1 + \alpha_k > 0$ and $0 < \Delta t/T \leq (t_k - \tau)/T$ for $\tau \in [0, t_{k-1}]$, with the help of ESA technique [5, 42], the kernel $((t_k - \tau)/T)^{-(1+\alpha_k)}$ in (10) can be approximated using a linear combination of exponentials. We have the following lemma to structure the robust fast formula.

Lemma 4 [5, 42] *For any constant $\beta_k \in [\beta_*, \beta^*] \subset [1, 2)$, $0 < \Delta t/T \leq (t_k - \tau)/T \leq 1$ for $\tau \in [0, t_{k-1}]$, $1 \leq k \leq n-1$ and the expected accuracy $0 < \epsilon \leq 1/e$, there exist a constant h , integers N^* and N_* , which satisfy*

$$\begin{aligned} h &= \frac{2\pi}{\log 3 + \beta^* \log(\cos 1)^{-1} + \log \epsilon^{-1}}, \\ N^* &= \left\lceil \frac{1}{h} \frac{1}{\beta_*} (\log \epsilon + \log \Gamma(1 + \beta^*)) \right\rceil, \\ N_* &= \left\lceil \frac{1}{h} \left(\log \frac{T}{\Delta t} + \log \log \epsilon^{-1} + \log \beta_* + 2^{-1} \right) \right\rceil, \end{aligned} \quad (11)$$

such that

$$\left| \left(\frac{t_k - \tau}{T} \right)^{-\beta_k} - \sum_{i=N_*+1}^{N^*} \theta_i^{(k)} e^{-\lambda_i(t_k - \tau)/T} \right| \leq \epsilon \left(\frac{t_k - \tau}{T} \right)^{-\beta_k},$$

where the quadrature exponents and weights are given by

$$\lambda_i = e^{ih}, \quad \theta_i^{(k)} = \frac{h e^{\beta_k ih}}{\Gamma(\beta_k)}.$$

Furthermore, the total number of terms in the summation can be estimated as

$$N_\epsilon = N^* - N_* \leq \frac{1}{10} \left(2 \log \frac{1}{\epsilon} + \log \beta^* + 2 \right) \left(\log \frac{T}{\Delta t} + \frac{1}{\beta_*} \log \frac{1}{\epsilon} + \log \log \frac{1}{\epsilon} + \frac{3}{2} \right).$$

So invoking Lemma 4 with $\beta_k = 1 + \alpha_k$ in (10), we obtain

$$\begin{aligned} &I_{his}(t_k) \\ &\approx \frac{1}{\Gamma(1 - \alpha_k)} \left(\frac{u(x, t_{k-1})}{\Delta t^{\alpha_k}} - \frac{u(x, t_0)}{t_k^{\alpha_k}} - \frac{\alpha_k}{T^{1+\alpha_k}} \sum_{i=N_*+1}^{N^*} \theta_i^{(k)} \int_0^{t_{k-1}} L_1(\tau) e^{-(t_k - \tau)/T \lambda_i} d\tau \right), \end{aligned} \quad (12)$$

where the quadrature weights and points are defined by

$$\lambda_i = e^{ih}, \quad \theta_i^{(k)} = \frac{h e^{(1+\alpha_k)ih}}{\Gamma(1 + \alpha_k)},$$

in which h , N_* and N^* are defined by (11). The related discretization formula is obtained (later in Lemma 5) as

$$I_{his}(t_k) = I_{his,\epsilon}(t_k) + \mathcal{O}(\epsilon \Delta t^{-\alpha_k}), \quad (13)$$

where $I_{his,\epsilon}(t_k)$ is defined by

$$I_{his,\epsilon}(t_k) = \frac{1}{\Gamma(1 - \alpha_k)} \left(\frac{u(x, t_{k-1})}{\Delta t^{\alpha_k}} - \frac{u(x, t_0)}{t_k^{\alpha_k}} - \frac{\alpha_k}{T^{1+\alpha_k}} \sum_{i=N_*+1}^{N^*} \theta_i^{(k)} F_{k,i} \right), \quad (14)$$

in which $F_{k,i}$ is given by

$$F_{k,i} = \int_0^{t_{k-1}} L_1(\tau) e^{-(t_k-\tau)\lambda_i/T} d\tau. \quad (15)$$

Note that $F_{1,i} = 0$ and $F_{k,i}$ can be calculated by the following recursive relation for $k = 2, 3, \dots, n$,

$$\begin{aligned} F_{k,i} &= e^{-\Delta t \lambda_i/T} F_{k-1,i} + \int_{t_{k-2}}^{t_{k-1}} L_{1,k-1}(\tau) e^{-(t_k-\tau)\lambda_i/T} d\tau \\ &= e^{-\Delta t \lambda_i/T} F_{k-1,i} \\ &\quad + T \frac{e^{-\Delta t \lambda_i/T}}{\Delta t \lambda_i^2} \left(-\Delta t \lambda_i e^{-\Delta t \lambda_i/T} + T - T e^{-\Delta t \lambda_i/T} \right) u(x, t_{k-2}) \\ &\quad + T \frac{e^{-\Delta t \lambda_i/T}}{\Delta t \lambda_i^2} \left(\Delta t \lambda_i - T + T e^{-\Delta t \lambda_i/T} \right) u(x, t_{k-1}). \end{aligned} \quad (16)$$

According to (13), replacing $I_{his}(t_k)$ in (8) by (14), we obtain *RF-L1* formula as

$$\begin{aligned} \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} u(x, t_k) &= I_{his, \epsilon}(t_k) + I_{loc}(t_k) \\ &= \frac{1}{\Gamma(1 - \alpha_k)} \left(\frac{u(x, t_{k-1})}{\Delta t^{\alpha_k}} - \frac{u(x, t_0)}{t_k^{\alpha_k}} - \frac{\alpha_k}{T^{1+\alpha_k}} \sum_{i=N_*+1}^{N^*} \theta_i^{(k)} F_{k,i} \right) \\ &\quad + \frac{u(x, t_k) - u(x, t_{k-1})}{\Delta t^{\alpha_k} \Gamma(2 - \alpha_k)}, \quad k = 2, 3, \dots, n. \end{aligned} \quad (17)$$

Recalling (9), we have

$$\mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_1} u(x, t_1) = \frac{u(x, t_1) - u(x, t_0)}{\Delta t^{\alpha_1} \Gamma(2 - \alpha_1)}. \quad (18)$$

Summarizing all this activity, we give the following algorithm to show the detailed instruction for the implementation of the robust fast algorithm for approximating the VO Caputo fractional derivative.

Algorithm 1 Robust fast algorithm to approximate VO Caputo fractional derivative gradually

- 1: Give the time step Δt , the expected accuracy ϵ and set h, N_*, N^* correspondingly
 - 2: Compute $\mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_1} u(x, t_1)$ by formula (18)
 - 3: Set $\{F_{1,i} = 0\}_{i=N_*+1}^{N^*}$ and $\{\lambda_i = e^{ih}\}_{i=N_*+1}^{N^*}$
 - 4: **for** $k = 2, 3, \dots, n$ **do**
 - 5: Set $\left\{ \theta_i^{(k)} = \frac{h e^{(1+\alpha_k)ih}}{\Gamma(1+\alpha_k)} \right\}_{i=N_*+1}^{N^*}$ and update $\{F_{k,i}\}_{i=N_*+1}^{N^*}$ by formula (16)
 - 6: Compute $\mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} u(x, t_k)$ by formula (17) using $\{F_{k,i}, \theta_i^{(k)}, \lambda_i\}_{i=N_*+1}^{N^*}$
 - 7: **end for**
-

Remark 1 At each time level, it only needs $\mathcal{O}(1)$ computational cost to compute $F_{k,i}$ since $F_{k-1,i}$ is known in advance. In total the robust fast algorithm requires only $\mathcal{O}(\log^2 n)$ memory and $\mathcal{O}(n \log^2 n)$ computational cost when numerically discretize the VO Caputo fractional derivative. The proposed method provides an efficient tool to approximate the VO Caputo fractional derivative.

3.1 Truncation error

In this subsection, we study the truncation error of $RF-L1$ formula (17)–(18) to the VO Caputo fractional derivative. To investigate the truncation error of $RF-L1$ formula to the VO Caputo fractional derivative, we first give the following lemma to state the error bound of $RF-L1$ formula to $L1$ formula.

Lemma 5 *Suppose **Conditions A** and **B** hold, $\alpha(t) \in C^1[0, T]$, $\varphi \in \check{H}^4$, $f \in H^1(0, T; \check{H}^2) \cap H^2(0, T; L_2)$. Let $L1$ formula be as (6), $RF-L1$ formula be defined by (17)–(18), ϵ be the expected accuracy, then*

$$\left\| {}_0\mathcal{D}_t^{\alpha_k} u(x, t_k) - \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} u(x, t_k) \right\| = \mathcal{O}(\epsilon \Delta t^{-\alpha_k}). \quad (19)$$

Proof For $k = 1$, we obtain the lemma directly by (9). For $k = 2, 3, \dots, n$, obviously, the only difference between $\mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} u(x, t_k)$ and ${}_0\mathcal{D}_t^{\alpha_k} u(x, t_k)$ is the approximation to the history part. According to (10) and (12), the error can be estimated by

$$\begin{aligned} & \left\| {}_0\mathcal{D}_t^{\alpha_k} u(x, t_k) - \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} u(x, t_k) \right\| \\ &= \frac{\alpha_k}{T^{1+\alpha_k} \Gamma(1-\alpha_k)} \left\| \int_0^{t_k-1} L_1(\tau) \left(\sum_{i=N_*+1}^{N^*} \theta_i^{(k)} e^{-(t_k-\tau)\lambda_i/T} - \left(\frac{t_k-\tau}{T}\right)^{-1-\alpha_k} \right) d\tau \right\|. \end{aligned}$$

By Lemma 4, we have

$$\begin{aligned} \left\| {}_0\mathcal{D}_t^{\alpha_k} u(x, t_k) - \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} u(x, t_k) \right\| &\leq \frac{\alpha_k}{\Gamma(1-\alpha_k)} \epsilon \left\| \int_0^{t_k-1} \frac{L_1(\tau)}{(t_k-\tau)^{1+\alpha_k}} d\tau \right\| \\ &\leq \frac{1}{\Gamma(1-\alpha_k)} \epsilon \|u\|_{C([0, T]; \check{H}^s)} \Delta t^{-\alpha_k} \\ &= \mathcal{O}(\epsilon \Delta t^{-\alpha_k}), \end{aligned}$$

where we apply Lemma 1, which guarantees the boundedness of $\|u\|_{C([0, T]; \check{H}^s)}$. The proof is completed.

We obtain the following theorem to estimate the truncation error of $RF-L1$ formula to the VO Caputo fractional derivative.

Theorem 1 *Suppose **Conditions A** and **B** hold, $\alpha(t) \in C^1[0, T]$, $\varphi \in \check{H}^4$, $f \in H^1(0, T; \check{H}^2) \cap H^2(0, T; L_2)$. Let the VO Caputo fractional derivative at t_k be as (5), its $RF-L1$ formula be defined by (17)–(18) and $\epsilon \leq \mathcal{O}(\Delta t^{1+\alpha^*})$ be the expected accuracy. If $\alpha(0) > 0$,*

$$\left\| {}_0^C \mathcal{D}_t^{\alpha_k} u(x, t_k) - \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} u(x, t_k) \right\| \leq ck^{-\alpha^*} \Delta t^{1-\alpha^*};$$

if $\alpha(0) = 0$,

$$\left\| {}_0^C \mathcal{D}_t^{\alpha(t)} u(x, t) - \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha(t)} u(x, t) \right\|_{\hat{L}_\infty(0, T; L_2)} \leq c \Delta t.$$

Proof The triangle inequality leads to

$$\begin{aligned} \left\| {}_0^C \mathcal{D}_t^{\alpha_k} u(x, t_k) - \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} u(x, t_k) \right\| &\leq \left\| {}_0^C \mathcal{D}_t^{\alpha_k} u(x, t_k) - {}_0\mathcal{D}_t^{\alpha_k} u(x, t_k) \right\| \\ &\quad + \left\| {}_0\mathcal{D}_t^{\alpha_k} u(x, t_k) - \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} u(x, t_k) \right\|. \end{aligned}$$

The desired result now follows on recalling Lemma 3 and Lemma 5.

3.2 Properties of discrete kernels

To simplify the further study of the convergence of the finite difference scheme, we first rewrite (17)–(18) using (15) into another form. For convenience, we denote

$$s^{(k)} = \frac{\Delta t^{-\alpha_k}}{\Gamma(2 - \alpha_k)}, \quad k = 1, 2, \dots, n.$$

By rearranging, it can be rewritten as

$$\mathcal{R}\mathcal{F}_0 \mathcal{D}_t^{\alpha_k} u(x, t_k) = s^{(k)} \left(u(x, t_k) - \sum_{l=0}^{k-1} d_l^{(k)} u(x, t_l) \right), \quad k = 1, 2, \dots, n, \quad (20)$$

where

$$d_0^{(k)} = (1 - \alpha_k) \left(k^{-\alpha_k} + \frac{\alpha_k}{T^{1+\alpha_k}} \Delta t^{\alpha_k-1} \sum_{i=N_*+1}^{N^*} \theta_i^{(k)} \int_{t_0}^{t_1} (t_1 - \tau) e^{-\lambda_i(t_k-\tau)/T} d\tau \right),$$

$$d_{k-1}^{(k)} = \alpha_k + \frac{\alpha_k(1 - \alpha_k)}{T^{1+\alpha_k}} \Delta t^{\alpha_k-1} \sum_{i=N_*+1}^{N^*} \theta_i^{(k)} \int_{t_{k-2}}^{t_{k-1}} (\tau - t_{k-2}) e^{-\lambda_i(t_k-\tau)/T} d\tau,$$

and

$$d_l^{(k)} = \frac{\alpha_k(1 - \alpha_k)}{T^{1+\alpha_k}} \Delta t^{\alpha_k-1} \sum_{i=N_*+1}^{N^*} \theta_i^{(k)} \left(\int_{t_{l-1}}^{t_l} (\tau - t_{l-1}) e^{-\lambda_i(t_k-\tau)/T} d\tau \right. \\ \left. + \int_{t_l}^{t_{l+1}} (t_{l+1} - \tau) e^{-\lambda_i(t_k-\tau)/T} d\tau \right), \quad l = 1, 2, \dots, k-2.$$

In particular $d_0^{(1)} = 1$.

Meanwhile, these coefficients satisfy the following lemma.

Lemma 6 *Let $\{d_l^{(k)}\}_{l=0}^{k-1}$ ($k = 1, 2, \dots, n$) be defined by (20) and ϵ be the expected accuracy. Then, we have*

$$(1) d_l^{(k)} > 0;$$

$$(2) \sum_{l=0}^{k-1} d_l^{(k)} \leq 1 + \epsilon.$$

Proof (1) This conclusion can be obtained by a straight forward calculation.

(2) Summing up $d_l^{(k)}$ for l from 0 to $k-1$, and rearranging the integral terms, we have

$$\begin{aligned}
\sum_{l=0}^{k-1} d_l^{(k)} &= (1 - \alpha_k)k^{-\alpha_k} + \alpha_k \\
&\quad + \frac{\alpha_k(1 - \alpha_k)}{T^{1+\alpha_k}} \Delta t^{\alpha_k-1} \sum_{j=0}^{k-2} \int_{t_j}^{t_{j+1}} (t_{j+1} - t_j) \sum_{i=N_*+1}^{N^*} \theta_i^{(k)} e^{-\lambda_i(t_k-\tau)} d\tau \\
&= (1 - \alpha_k)k^{-\alpha_k} + \alpha_k + \frac{\alpha_k(1 - \alpha_k)}{T^{1+\alpha_k}} \Delta t^{\alpha_k} \int_0^{t_{k-1}} \sum_{i=N_*+1}^{N^*} \theta_i^{(k)} e^{-\lambda_i(t_k-\tau)} d\tau \\
&\leq (1 - \alpha_k)k^{-\alpha_k} + \alpha_k + (1 + \epsilon)(1 - \alpha_k)\alpha_k \Delta t^{\alpha_k} \int_0^{t_{k-1}} (t_k - \tau)^{-1-\alpha_k} d\tau \\
&= (1 - \alpha_k)k^{-\alpha_k} + \alpha_k + (1 + \epsilon)(1 - \alpha_k)(1 - k^{-\alpha_k}) \\
&= 1 + \epsilon - \alpha_k \epsilon + \alpha_k \epsilon k^{-\alpha_k} - \epsilon k^{-\alpha_k} \\
&\leq 1 + \epsilon,
\end{aligned}$$

where we use the estimate

$$(1 - \epsilon)t^{-1-\alpha_k} \leq \sum_{i=N_*+1}^{N^*} \theta_i^{(k)} e^{-\lambda_i t} \leq (1 + \epsilon)t^{-1-\alpha_k}.$$

The proof is completed.

4 Finite difference scheme for VO tFDE

In this section, we propose a fast finite difference scheme for solving (1)–(3) and prove its error estimates.

We first discretize the first-order partial derivative u_t by

$$\frac{\partial}{\partial t} u(x, t_k) = \frac{u(x, t_k) - u(x, t_{k-1})}{\Delta t} + E^k := \Delta_t u(x, t_k) + E^k, \quad (21)$$

where E^k satisfies the following lemma.

Lemma 7 [44] *Suppose **Conditions A** and **B** hold, $\alpha(t) \in C^1[0, T]$, $\varphi \in \check{H}^4$ and $f \in H^1(0, T; \check{H}^2) \cap H^2(0, T; L_2)$. If $\alpha(0) > 0$,*

$$\|E^k\| \leq ck^{-\alpha(0)} \Delta t^{1-\alpha(0)};$$

if $\alpha(0) = 0$,

$$\|E\|_{\hat{L}_\infty(0, T; L_2)} \leq c\Delta t.$$

Let m be a positive integer, $\Delta x = (x_r - x_l)/m$ be the size of spatial grid, and define a spatial partition $x_j = x_l + j\Delta x$ for $j = 0, 1, \dots, m$. Denote $x_{j+1/2} = (x_{j+1} + x_j)/2$ as the midpoint of the neighboring nodes x_{j+1} and x_j , $p_{j\pm 1/2} = p(x_{j\pm 1/2})$, $f_j^k = f(x_j, t_k)$ and $\varphi_j = \varphi(x_j)$.

The integer-order diffusion term is discretized by [40]

$$\begin{aligned} & \frac{\partial}{\partial x} \left[p(x_j) \frac{\partial u(x_j, t_k)}{\partial x} \right] \\ &= \frac{1}{\Delta x} \left(p_{j+1/2} \frac{u(x_{j+1}, t_k) - u(x_j, t_k)}{\Delta x} - p_{j-1/2} \frac{u(x_j, t_k) - u(x_{j-1}, t_k)}{\Delta x} \right) + G_j^k \\ &:= \Delta_x u(x_j, t_k) + G_j^k, \end{aligned} \quad (22)$$

where $G_j^k = c\Delta x^2$.

Then substituting (21), (22) and (20) into (1) gives

$$\begin{aligned} & \Delta_t u(x_j, t_k) + \zeta \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} u(x_j, t_k) \\ &= \Delta_x u(x_j, t_k) + f_j^k - (E^k(x_j) + \zeta R^k(x_j) - G_j^k), \quad 1 \leq j \leq m-1, \quad 1 \leq k \leq n. \end{aligned} \quad (23)$$

From the initial and boundary value conditions (2)–(3), we have

$$u(x_j, 0) = \varphi_j, \quad 0 \leq j \leq m, \quad (24)$$

$$u(x_l, t_k) = u(x_r, t_k) = 0, \quad 1 \leq k \leq n. \quad (25)$$

4.1 Finite difference schemes

Denote the approximate solution to $u(x_j, t_k)$ by U_j^k . Then,

$$\begin{aligned} \Delta_t U_j^k &= \frac{U_j^k - U_j^{k-1}}{\Delta t}, \\ \Delta_x U_j^k &= \frac{1}{\Delta x} \left(p_{j+1/2} \frac{U_{j+1}^k - U_j^k}{\Delta x} - p_{j-1/2} \frac{U_j^k - U_{j-1}^k}{\Delta x} \right). \end{aligned}$$

We obtain *RF-L1* scheme for the problem (1)–(3) as follows

$$\Delta_t U_j^k + \zeta \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} U_j^k = \Delta_x U_j^k + f_j^k, \quad 1 \leq j \leq m-1, \quad 1 \leq k \leq n, \quad (26)$$

$$U_j^0 = \varphi_j, \quad 0 \leq j \leq m, \quad (27)$$

$$U_0^k = U_m^k = 0, \quad 1 \leq k \leq n, \quad (28)$$

where

$$\begin{aligned} \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_k} U_j^k &= \frac{1}{\Gamma(1 - \alpha_k)} \left(\frac{U_j^{k-1}}{\Delta t^{\alpha_k}} - \frac{U_j^0}{t_k^{\alpha_k}} - \frac{\alpha_k}{T^{1+\alpha_k}} \sum_{i=N_*+1}^{N^*} \theta_i^{(k)} F_{k,i} \right) \\ &\quad + \frac{U_j^k - U_j^{k-1}}{\Delta t^{\alpha_k} \Gamma(2 - \alpha_k)}, \quad k = 2, 3, \dots, n, \\ \mathcal{R}_0^{\mathcal{F}} \mathcal{D}_t^{\alpha_1} U_j^1 &= \frac{U_j^1 - U_j^0}{\Delta t^{\alpha_1} \Gamma(2 - \alpha_1)}, \end{aligned}$$

in which $F_{1,i} = 0$ and for $k = 2, 3, \dots$,

$$F_{k,i} = e^{-\Delta t \lambda_i} F_{k-1,i} + T \frac{e^{-\Delta t \lambda_i}}{\Delta t \lambda_i^2} \left(-\Delta t \lambda_i e^{-\Delta t \lambda_i} + T - T e^{-\Delta t \lambda_i} \right) U_j^{k-2} \\ + T \frac{e^{-\Delta t \lambda_i}}{\Delta t \lambda_i^2} \left(\Delta t \lambda_i - T + T e^{-\Delta t \lambda_i} \right) U_j^{k-1}.$$

Similarly, we obtain L1 scheme as

$$\Delta_t U_j^k + \zeta_0 \mathcal{D}_t^{\alpha_k} U_j^k = \Delta_x U_j^k + f_j^k, \quad 1 \leq j \leq m-1, \quad 1 \leq k \leq n, \quad (29)$$

$$U_j^0 = \varphi_j, \quad 0 \leq j \leq m, \quad (30)$$

$$U_0^k = U_m^k = 0, \quad 1 \leq k \leq n, \quad (31)$$

where

$${}_0 \mathcal{D}_t^{\alpha_k} U_j^k = \frac{\Delta t^{-\alpha_k}}{\Gamma(2-\alpha_k)} \left(a_0^{(k)} U_j^k - \sum_{l=1}^{k-1} \left(a_{k-l-1}^{(k)} - a_{k-l}^{(k)} \right) U_j^l - a_{k-1}^{(k)} U_j^0 \right),$$

in which $a_l^{(k)} = (l+1)^{1-\alpha_k} - l^{1-\alpha_k}$.

And F-L1 scheme is obtained as

$$\Delta_t U_j^k + \zeta_0 \mathcal{F}_0 \mathcal{D}_t^{\alpha_k} U_j^k = \Delta_x U_j^k + f_j^k, \quad 1 \leq j \leq m-1, \quad 1 \leq k \leq n, \quad (32)$$

$$U_j^0 = \varphi_j, \quad 0 \leq j \leq m, \quad (33)$$

$$U_0^k = U_m^k = 0, \quad 1 \leq k \leq n, \quad (34)$$

where

$$\mathcal{F}_0 \mathcal{D}_t^{\alpha_k} U_j^k = \frac{T^{-\alpha_k}}{\Gamma(1-\alpha_k)} \sum_{i=\bar{N}+1}^{\bar{N}} \tilde{\theta}_i^{(k)} \tilde{F}_{k,i} + \frac{U_j^k - U_j^{k-1}}{\Delta t^{\alpha_k} \Gamma(2-\alpha_k)}, \quad k = 2, 3, \dots, n, \\ \mathcal{F}_0 \mathcal{D}_t^{\alpha_1} U_j^1 = \frac{U_j^1 - U_j^0}{\Delta t^{\alpha_1} \Gamma(2-\alpha_1)},$$

in which $\tilde{F}_{1,i} = 0$, and for $k = 2, 3, \dots$,

$$\tilde{F}_{k,i} = e^{-\tilde{\lambda}_i \Delta t / T} \tilde{F}_{k-1,i} + T \frac{e^{-\tilde{\lambda}_i \Delta t / T} - e^{-2\tilde{\lambda}_i \Delta t / T}}{\tilde{\lambda}_i \Delta t} (U_j^{k-1} - U_j^{k-2}).$$

4.2 Convergence

Next we estimate the error without any artificial regularity assumptions of the true solution.

Theorem 2 Suppose **Conditions A** and **B** hold, $\alpha(t) \in C^1[0, T]$, $\varphi \in \check{H}^4$, $f \in H^1(0, T; \check{H}^2) \cap H^2(0, T; L_2)$. Suppose the expected accuracy $\epsilon \leq \mathcal{O}(\Delta t^{1+\alpha^*})$. Suppose $\{u(x_j, t_k) | 0 \leq j \leq m, 0 \leq k \leq n\}$ and $\{U_j^k | 0 \leq j \leq m, 0 \leq k \leq n\}$ are solutions of the problem (1)–(3) and RF-L1 scheme (26)–(28), respectively. Then

$$\|U - u\|_{\hat{L}_\infty(0, T; L_2)} \leq c(\Delta t + \Delta x^2).$$

Proof Let $r_j^k = u_j^k - u(x_j, t_k)$ for $0 \leq j \leq m$, $0 \leq k \leq n$. Subtracting (26)–(28) from (23)–(25), we obtain the error equation

$$\Delta_t r_j^k + \zeta \mathcal{R}\mathcal{F}_0 \mathcal{D}_t^{\alpha_k} r_j^k = \Delta_x r_j^k - (E^k(x_j) + \zeta R^k(x_j) - G_j^k), \quad 1 \leq j \leq m-1, 1 \leq k \leq n, \quad (35)$$

$$r_j^0 = 0, \quad 0 \leq j \leq m, \quad (36)$$

$$r_0^k = r_m^k = 0, \quad 1 \leq k \leq n. \quad (37)$$

Use $r_j^0 = 0$ and (20) to rearrange (35) as

$$\begin{aligned} & \frac{r_j^k - r_j^{k-1}}{\Delta t} + \zeta s^{(k)} \left(r_j^k - \sum_{l=1}^{k-1} d_l^{(k)} r_j^l \right) - \frac{1}{\Delta x} \left(p_{j+1/2} \frac{r_{j+1}^k - r_j^k}{\Delta x} - p_{j-1/2} \frac{r_j^k - r_{j-1}^k}{\Delta x} \right) \\ & = - (E^k(x_j) + \zeta R^k(x_j) - G_j^k). \end{aligned}$$

Making an inner product with r^k on both hand sides of the equality, and from Theorem 1, Lemma 7 and (22), we obtain

$$(1 + \zeta \Delta t s^{(k)}) \|r^k\| \leq \|r^{k-1}\| + \zeta \Delta t s^{(k)} \sum_{l=1}^{k-1} d_l^{(k)} \|r^l\| + \Delta t J^k, \quad (38)$$

where $J^k \leq c_1 k^{-\alpha(0)} n^{\alpha(0)-1} + c_2 k^{-\alpha^*} n^{\alpha^*-1} + c_3 \Delta x^2$. It is clear from (38) and $\|r^0\| = 0$ that

$$\|r^1\| \leq \Delta t J^1 \leq \Delta t (1 + \epsilon) J^1.$$

Assume that

$$\|r^{k_0}\| \leq \Delta t (1 + \epsilon)^{k_0} \sum_{q=1}^{k_0} J^q, \quad k_0 = 2, 3, \dots, k-1. \quad (39)$$

Using the mathematical induction, it is derived that

$$\begin{aligned} & (1 + \zeta \Delta t s^{(k)}) \|r^k\| \\ & \leq \Delta t (1 + \epsilon)^{k-1} \sum_{q=1}^{k-1} J^q + \zeta \Delta t s^{(k)} \sum_{l=1}^{k-1} d_l^{(k)} \left(\Delta t (1 + \epsilon)^l \sum_{q=1}^l J^q \right) + \Delta t J^k \\ & \leq \Delta t (1 + \epsilon)^k \sum_{q=1}^k J^q + \zeta \Delta t^2 s^{(k)} (1 + \epsilon)^{k-1} \sum_{l=1}^{k-1} d_l^{(k)} \sum_{q=1}^l J^q \\ & \leq \Delta t (1 + \epsilon)^k \sum_{q=1}^k J^q + \zeta \Delta t^2 s^{(k)} (1 + \epsilon)^k \sum_{q=1}^{k-1} J^q \\ & \leq \Delta t (1 + \epsilon)^k (1 + \zeta \Delta t s^{(k)}) \sum_{q=1}^k J^q. \end{aligned}$$

Thus (39) holds for $k = 1, 2, \dots, n$ by mathematical induction. It remains to bound the right-hand side of (39) for any $1 \leq k \leq n$. We use Theorem 1, Lemma 7 and (22) again to conclude that

$$\begin{aligned} \|r^n\| &\leq \Delta t(1 + \epsilon)^n \sum_{k=1}^n J^k \leq c(1 + \epsilon)^n \Delta t \sum_{k=1}^n \frac{1}{k^{\alpha_*} n^{1-\alpha_*}} + c(1 + \epsilon)^n \Delta t \sum_{k=1}^n \Delta x^2 \\ &\leq c_4(1 + \epsilon)^n \Delta t + c_5(1 + \epsilon)^n \Delta x^2 \\ &\leq c_6 e^T \Delta t + c_7 e^T \Delta x^2. \end{aligned}$$

We incorporate these estimates into (39) to complete the proof.

5 Numerical results

In this section, we test some problems and present the numerical results to verify the effectiveness of the proposed *RF-L1* scheme (26)–(28) compared with *L1* scheme (29)–(31) and *F-L1* scheme (32)–(34). All experiments are performed based on Matlab 2016b on a laptop with the configuration: Intel(R) Core(TM) i7-7500U CPU 2.70GHz and 8.00 GB RAM.

Example 1 To verify the efficiency of the robust fast algorithm for the VO Caputo fractional derivative, we first solve an ordinary differential equation

$$\begin{aligned} \frac{\partial}{\partial t} u(t) + \zeta {}_0^C \mathcal{D}_t^{\alpha(t)} u(t) &= 1, \quad t \in (0, T], \\ u(0) &= 1, \end{aligned}$$

where $\zeta = 1$, $T=1$ and the VO function is given by

$$\alpha(t) = \alpha(T) + (\alpha(0) - \alpha(T)) \left(1 - t/T - \frac{\sin(2\pi(1 - t/T))}{2\pi} \right). \quad (40)$$

In the calculations, we use the numerical solutions \widehat{U} to the corresponding problem discretized with $\Delta t = 1/2^{22}$ as the reference solutions. Define the error and the convergence rate in time by

$$E(\Delta t) = |u^n - \widehat{U}^n|, \quad R_t = \log_2 \frac{E(\Delta t)}{E(\Delta t/2)},$$

respectively.

We set the expected accuracy $\epsilon = (\Delta t/T)^2$ to keep the accuracy of the solution of *RF-L1* scheme as same as that of *L1* scheme. The numerical results of *L1* scheme, *F-L1* scheme and *RF-L1* scheme with different $\alpha(0)$ and $\alpha(T)$ are listed in Table 1. For $\alpha_* = \alpha(0) = 0.2$, Table 1 shows that compared with *L1* scheme, the two fast algorithms *F-L1* scheme and *RF-L1* scheme greatly reduce the computational cost. The CPU time reveal $\mathcal{O}(n \log^2 n)$ computational complexity of *F-L1* scheme and *RF-L1* scheme, and $\mathcal{O}(n^2)$ computational complexity of *L1* scheme, respectively. Although the computational complexity required in *F-L1* scheme and *RF-L1* scheme are both $\mathcal{O}(n \log^2 n)$, the numbers of exponentials needed in *RF-L1* scheme are very modest and not strongly influenced by α_* , which indeed contributes to reduce the memory and computational cost. Moreover, as we have mentioned above, *F-L1* scheme is not applicable for α_* of small value. Table 1 shows that *F-L1* scheme cannot achieve the ideal convergence rate for $\alpha_* = 0.05$ and even cannot work for $\alpha_* = 0$. *RF-L1* scheme always performs well with few CPU time and memory.

Example 2 We investigate the temporal and spatial convergence behaviors of *RF-L1* scheme. Consider (1)–(3) with the spatial domain $[x_l, x_r] = [0, 1]$, the time interval $[0, T] = [0, 1]$, $\zeta = 1$, $p(x) = 1$, $\varphi(x) = \sin(\pi x)$, $f = 0$ and the VO function is given by (40).

Set the expected accuracy $\epsilon = (\Delta t/T)^2$. We use the numerical solutions \widehat{U} to the corresponding tFDE models discretized with $\Delta x = (x_r - x_l)/2^{10}$ and $\Delta t = T/2^{18}$ as the reference solutions. Define the error, the convergence rate in time and in space by

$$E(\Delta x, \Delta t) = \max_{0 \leq j \leq m} |u_j^n - \widehat{U}_j^n|, \quad R_t = \log_2 \frac{E(\Delta x, \Delta t)}{E(\Delta x, \Delta t/2)}, \quad R_s = \log_2 \frac{E(\Delta x, \Delta t)}{E(\Delta x/2, \Delta t)},$$

respectively.

The error and temporal convergence order of *L1* scheme, *F-L1* scheme and *RF-L1* scheme with different $\alpha(0)$ and $\alpha(T)$ are listed in Table 2. Fine spatial size is fixed at $\Delta x = (x_r - x_l)/2^{10}$. *F-L1* scheme and *RF-L1* scheme achieve the same accuracy as *L1* scheme when $\alpha_* = \alpha(0) = 0.2$. Compared with *L1* scheme, *F-L1* scheme greatly save the computational cost, and *RF-L1* scheme further reduces CPU time and memory since much less N_ϵ is needed. Besides, *F-L1* scheme fails to solve the problem with $\alpha_* = 0$ and is not very effective for $\alpha_* = 0.05$. Moreover, *RF-L1* scheme is valid for $\alpha(0) = 0, 0.05$ and save much computational cost.

Table 3 lists the spatial convergence order of *L1* scheme, *F-L1* scheme and *RF-L1* scheme with $\alpha(0) = 0.05$ and $\alpha(T) = 0.5$. Fine temporal step is fixed at $\Delta t = T/2^{18}$ and spatial sizes refined from $\Delta x = (x_r - x_l)/2^3$ to $\Delta x = (x_r - x_l)/2^7$. It shows that three schemes achieve the second-order convergence in space. Nevertheless, CPU time and memory of *RF-L1* scheme are cheaper than those of *L1* scheme and *F-L1* scheme.

Table 1 Temporal convergence rates and the CPU time, memory of *L1* scheme, *F-L1* scheme and *RF-L1* scheme for Example 1

$(\alpha(0), \alpha(T))$	n	L1 scheme				F-L1 scheme					RF-L1 scheme				
		$E(\Delta t)$	R_t	CPU(s)	Memory	$E(\Delta t)$	R_t	CPU(s)	Memory	\tilde{N}_ϵ	$E(\Delta t)$	R_t	CPU(s)	Memory	N_ϵ
(0, 0.2)	2^{13}	2.1281e-5	-	3.53	1.97e+5	-	-	-	-	-	2.1281e-5	-	0.26	4.90e+3	98
	2^{14}	1.0620e-5	1.00	12.69	3.93e+5	-	-	-	-	-	1.0619e-5	1.00	0.50	5.58e+3	112
	2^{15}	5.2890e-6	1.00	46.27	7.87e+5	-	-	-	-	-	5.2889e-6	1.01	1.06	6.30e+3	127
	2^{16}	2.6237e-6	1.01	179.23	1.57e+6	-	-	-	-	-	2.6236e-6	1.01	2.20	7.06e+3	143
	2^{17}	1.2910e-6	1.02	733.52	3.15e+6	-	-	-	-	-	1.2910e-6	1.02	4.02	7.83e+3	159
(0.05, 0.5)	2^{13}	1.9849e-5	-	3.29	1.97e+5	1.4301e-5	-	0.97	5.55e+4	1153	1.9849e-5	-	0.30	4.76e+5	95
	2^{14}	9.9041e-6	1.00	11.84	3.93e+5	3.5916e-6	1.99	1.82	6.40e+4	1329	9.9040e-6	1.00	0.52	5.48e+5	110
	2^{15}	4.9329e-6	1.01	47.42	7.87e+5	1.4914e-6	1.27	3.09	7.31e+4	1519	4.9327e-6	1.01	1.20	6.10e+5	123
	2^{16}	2.4476e-6	1.01	187.85	1.57e+6	9.8072e-8	3.93	6.65	8.27e+4	1720	2.4473e-6	1.01	2.24	6.87e+6	139
	2^{17}	1.2051e-6	1.02	767.61	3.15e+6	2.0321e-7	-1.05	14.18	9.31e+4	1935	1.2049e-6	1.02	4.58	7.69e+6	156
(0.2, 0.6)	2^{13}	1.8761e-5	-	3.19	1.97e+7	1.8754e-5	-	0.65	1.54e+4	317	1.8761e-5	-	0.50	4.52e+5	90
	2^{14}	9.3607e-6	1.00	12.50	3.93e+7	9.3583e-6	1.00	0.95	1.77e+4	365	9.3605e-6	1.00	0.66	5.10e+5	102
	2^{15}	4.6624e-6	1.01	48.05	7.87e+7	4.6601e-6	1.01	2.01	2.02e+4	416	4.6622e-6	1.01	1.31	5.77e+5	116
	2^{16}	2.3139e-6	1.01	188.77	1.57e+8	2.3137e-6	1.01	4.06	2.28e+4	471	2.3135e-6	1.01	2.65	6.44e+5	130
	2^{17}	1.1399e-6	1.02	759.75	3.15e+8	1.1398e-6	1.02	8.20	2.56e+4	529	1.1397e-6	1.02	4.59	7.11e+6	144

Table 2 Temporal convergence rates and the CPU time, memory of $L1$ scheme, $F-L1$ scheme and $RF-L1$ scheme for Example 2 with $m = 2^{10}$

$(\alpha(0), \alpha(T))$	n	$L1$ scheme				$F-L1$ scheme					$RF-L1$ scheme				
		$E(\Delta x, \Delta t)$	R_t	CPU(s)	Memory	$E(\Delta x, \Delta t)$	R_t	CPU(s)	Memory	\tilde{N}_ϵ	$E(\Delta x, \Delta t)$	R_t	CPU(s)	Memory	N_ϵ
(0, 0.2)	2^{11}	6.5685e-6	-	17.34	1.69e+7	-	-	-	-	-	6.5685e-6	-	1.41	7.48e+5	73
	2^{12}	3.2568e-6	1.01	66.37	3.37e+7	-	-	-	-	-	3.2568e-6	1.01	3.06	8.46e+5	85
	2^{13}	1.6022e-6	1.02	284.29	6.73e+7	-	-	-	-	-	1.6022e-6	1.02	6.26	9.53e+5	98
	2^{14}	7.7515e-7	1.05	1170.43	1.34e+8	-	-	-	-	-	7.7515e-7	1.05	14.14	1.07e+6	112
	2^{15}	3.6171e-7	1.10	4599.13	2.69e+8	-	-	-	-	-	3.6171e-7	1.10	35.79	1.19e+6	127
(0.05, 0.5)	2^{11}	1.4465e-5	-	17.53	1.69e+7	1.2593e-5	-	43.91	7.03e+6	837	1.4465e-5	-	1.11	7.31e+5	71
	2^{12}	7.1687e-6	1.01	75.82	3.37e+7	6.1145e-6	1.04	99.49	8.28e+6	989	7.1687e-6	1.01	2.22	8.30e+5	83
	2^{13}	3.5253e-6	1.02	278.57	6.73e+7	3.2392e-6	0.92	232.70	9.63e+6	1153	3.5253e-6	1.02	5.86	9.29e+5	95
	2^{14}	1.7051e-6	1.05	1098.39	1.34e+8	1.3753e-6	1.24	530.39	1.11e+7	1329	1.7051e-6	1.05	15.34	1.05e+6	110
	2^{15}	7.9551e-7	1.10	4429.14	2.69e+8	6.1283e-7	1.17	1211.36	1.26e+7	1519	7.9551e-7	1.10	34.56	1.16e+6	123
(0.2, 0.6)	2^{11}	1.6780e-5	-	17.20	1.69e+7	1.6777e-5	-	12.49	2.05e+6	231	1.6780e-5	-	0.99	6.98e+5	67
	2^{12}	8.3079e-6	1.01	67.22	3.37e+7	8.3049e-6	1.01	30.18	2.38e+6	272	8.3078e-6	1.01	1.99	7.89e+5	78
	2^{13}	4.0826e-6	1.02	268.00	6.73e+7	4.0822e-6	1.02	66.48	2.75e+6	317	4.0826e-6	1.02	4.30	8.88e+5	90
	2^{14}	1.9736e-6	1.05	1073.83	1.34e+8	1.9735e-6	1.05	152.53	3.15e+6	365	1.9736e-6	1.05	9.58	9.86e+5	102
	2^{15}	9.2040e-7	1.10	4565.72	2.69e+8	9.2026e-7	1.10	343.53	3.57e+6	416	9.2040e-7	1.10	23.17	1.10e+6	116

Table 3 Spatial convergence rates and the CPU time, memory of $L1$ scheme, $F-L1$ scheme and $RF-L1$ scheme for Example 2 with $n = 2^{18}$

$(\alpha(0), \alpha(T))$	m	$L1$ scheme				$F-L1$ scheme					$RF-L1$ scheme				
		$E(\Delta x, \Delta t)$	R_s	CPU(s)	Memory	$E(\Delta x, \Delta t)$	R_s	CPU(s)	Memory	\tilde{N}_ϵ	$E(\Delta x, \Delta t)$	R_s	CPU(s)	Memory	N_ϵ
(0.05, 0.5)	2^3	9.2958e-4	-	5533.98	1.89e+7	9.2955e-4	-	82.48	2.33e+4	2161	9.2958e-4	-	15.90	4.20e+4	172
	2^4	2.3079e-4	2.01	7586.97	3.57e+7	2.3076e-4	2.01	102.65	3.72e+5	2161	2.3079e-4	2.01	18.05	5.40e+4	172
	2^5	5.7557e-5	2.00	11751.81	6.92e+7	5.7527e-5	2.00	150.06	6.51e+5	2161	5.7557e-5	2.00	26.15	7.79e+4	172
	2^6	1.4341e-5	2.00	20286.87	1.36e+8	1.4311e-5	2.01	968.32	1.21e+6	2161	1.4341e-5	2.00	49.96	1.26e+5	172
	2^7	3.5427e-6	2.02	37741.75	2.71e+8	3.5132e-6	2.03	2204.64	2.32e+6	2161	3.5427e-6	2.02	63.77	2.22e+5	172

6 Concluding Remarks

In this paper, a robust fast algorithm is developed to approximate the VO Caputo fractional derivative, which can handle the cases of small or vanishing lower bound of the VO function. The method is applied to construct a fast finite difference scheme for the VO tFDEs. Moreover, the convergence is studied without any regularity assumptions of the true solution. Numerical tests are reported to show the effectiveness of the proposed scheme and confirm the theoretical findings.

In our future work, fast methods for high-order scheme to approximate the VO tFDEs with the initial weak singularity are meaningful to study.

References

1. R.A. Adams and J.J.F. Fournier, Sobolev spaces, Elsevier, San Diego, 2003.
2. D. Baffet and J.S. Hesthaven, A kernel compression scheme for fractional differential equations, SIAM J. Numer. Anal., 55 (2017), pp. 496–520.

3. D.A. Benson, S.W. Wheatcraft, and M.M. Meerschaert, Application of a fractional advection dispersion equation, *Water Resour. Res.*, 36 (2000), pp. 1403–1412.
4. D. Bertaccini and F. Durastante, Block structured preconditioners in tensor form for the all-at-once solution of a finite volume fractional diffusion equation, *Appl. Math. Lett.*, 95 (2019), pp. 92–97.
5. G. Beylkin and L. Monzon, Approximation by exponential sums revisited, *Appl. Comput. Harmon. Anal.*, 28 (2010), pp. 131–149.
6. A.V. Chechkin, R. Gorenflo, and I.M. Sokolov, Fractional diffusion in inhomogeneous media, *J. Phys. A: Math. Gen.*, 38 (2005), pp. 679–684.
7. C.F.M. Coimbra, Mechanics with variable-order differential operators, *Ann. Phys.*, 12 (2003), pp. 643–736.
8. G. Diazand and C.F.M. Coimbra, Nonlinear dynamics and control of a variable order oscillator with application to the van der pol equation, *Nonlinear Dynam.*, 56 (2009), pp. 145–157.
9. R.L. Du, A.A. Alikhanov, and Z.Z. Sun, Temporal second order difference schemes for the multi-dimensional variable-order time fractional sub-diffusion equations, *Comput. Math. with Appl.*, 79 (2020), pp. 2952–2972.
10. Z.W. Fang, H.W. Sun, and H. Wang, A fast method for variable-order caputo fractional derivative with applications to time fractional diffusion equations, *Comput. Math. Appl.*, 80 (2020), pp. 1443–1458.
11. Z.J. Fu, S. Reutskiy, H.G. Sun, J. Ma, and M.A. Khan, A robust kernel-based solver for variable-order time fractional PDEs under 2D/3D irregular domains, *Appl. Math. Lett.*, 94 (2019), pp. 105–111.
12. H.F. Fu and H. Wang, A preconditioned fast finite difference method for space-time fractional partial differential equations, *Fract. Calc. Appl. Anal.*, 20 (2017), pp. 88–116.
13. X.M. Gu, H.W. Sun, Y. L. Zhao, and X. C. Zheng, An implicit difference scheme for time-fractional diffusion equations with a time-invariant type variable order, *Appl. Math. Lett.*, DOI: <https://doi.org/10.1016/j.aml.2021.107270>, 2021.
14. D. Ingman and J. Suzdalnitsky, Control of damping oscillations by fractional differential operator with time-dependent order, *Comput. Methods Appl. Mech. Eng.*, 193 (2004), pp. 5585–5595.
15. Y. Jia, M. Xu, and Y. Lin, A numerical solution for variable order fractional functional differential equation, *Appl. Math. Lett.*, 64 (2017), pp. 125–130.
16. S.D. Jiang, J.W. Zhang, Q. Zhang, and Z.M. Zhang, Fast evaluation of the caputo fractional derivative and its applications to fractional diffusion equations, *Commun. Comput. Phys.*, 21 (2017), pp. 650–678.
17. R.H. Ke, M.K. Ng, and H.W. Sun, A fast direct method for block triangular Toeplitz-like with tridiagonal block systems from time-fractional partial differential equations, *J. Comput. Phys.*, 303 (2015), pp. 203–211.
18. A.A. Kilbas, H.M. Srivastava, and J.J. Trujillo, *Theory and applications of fractional differential equations*, Elsevier, Amsterdam, 2006.
19. T.A. M. Langlands and B.I. Henry, The accuracy and stability of an implicit solution method for the fractional diffusion equation, *J. Comput. Phys.*, 205 (2005), pp. 719–736.
20. C.F. Lorenzo and T.T. Hartley, Variable order and distributed order fractional operators, *Nonlin. Dyn.*, 29 (2002), pp. 57–98.
21. X. Lu, H.W. Sun, and H.K. Pang, Fast approximate inversion of a block triangular Toeplitz matrix with applications to fractional subdiffusion equations, *J. Comput. Phys.*, 22 (2015), pp. 866–882.
22. H.L. Liao, D.F. Li, and J.W. Zhang, Sharp error estimate of the nonuniform L_1 formula for linear reaction-subdiffusion equations, *SIAM J. Numer. Anal.*, 56 (2018), pp. 1112–1133.
23. F.W. Liu, V. Anh, and I. Turner, Numerical solution of the space fractional Fokker-Planck equation, *J. Comput. Appl. Math.*, 166 (2004), pp. 209–219.
24. C. Lubich and A. Schädle, Fast convolution for nonreflecting boundary conditions, *SIAM J. Sci. Comput.*, 24 (2002), pp. 161–182.
25. F. Mainardi, M. Raberto, R. Gorenflo, and E. Scalas, Fractional calculus and continuous-time finance II: the waiting-time distribution, *Phys. A*, 287 (2000), pp. 468–481.
26. K.B. Oldham and J. Spanier, *The fractional calculus*, Academic Press, New York, 1974.
27. A.D. Obembe, M.E. Hossain, and S.A. Abu-Khamsin, Variable-order derivative time fractional diffusion model for heterogeneous porous media, *J. Petrol. Sci. Eng.*, 152 (2017), pp. 391–405.
28. S. Patnaik, J.P. Hollkamp, and F. Semperlotti, Applications of variable-order fractional operators: a review, *Proc. R. Soc. Lond. Ser. A. Math. Phys. Eng. Sci.*, DOI: <http://dx.doi.org/10.1098/rspa.2019.0498>, 2020.
29. H.T. C. Pedro, M.H. Kobayashi, J.M. C. Pereira, and C.F. M. Coimbra, Variable order modeling of diffusive-convective effects on the oscillatory flow past a sphere, *J. Vib. Control*, 14 (2008), pp. 1659–1672.
30. I. Podlubny, *Fractional differential equations*, Academic Press, New York, 1999.
31. M. Raberto, E. Scalas, and F. Mainardi, Waiting-times and returns in high-frequency financial data: an empirical study, *Phys. A*, 314 (2002), pp. 749–755.
32. K. Sakamoto and M. Yamamoto, Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems, *J. Math. Anal. Appl.*, 382 (2011), pp. 426–447.
33. I.M. Sokolov and J. Klafter, From diffusion to anomalous diffusion: a century after einsteins brownian motion, *Chaos*, 15 (2005), pp. 1–7.
34. Z.Z. Sun, *Numerical methods of partial differential equations*, Science Press, Beijing, 2005.

35. H.G. Sun, W. Chen, and Y.Q. Chen, Variable-order fractional differential operators in anomalous diffusion modeling, *Phys. A*, 388 (2009), pp. 4586–4592.
36. H.G. Sun, A. Chang, Y. Zhang, and W. Chen, A review on variable-order fractional differential equations: mathematical foundations, physical models, numerical methods and applications, *Fract. Calc. Appl. Anal.*, 22 (2019), pp. 27–59.
37. H.G. Sun, W. Chen, H. Wei, and Y.Q. Chen, A comparative study of constant-order and variable-order fractional models in characterizing memory property of systems, *Eur. Phys. J. Spec. Top.* 193 (2011), pp. 185–192.
38. H.G. Sun, W. Chen, C.P. Li, and Y.Q. Chen, Finite difference schemes for variable-order time fractional diffusion equation, *Int. J. Bifurcation Chaos*, 22 (2012), pp. 1250085, 16 pages.
39. V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer, New York, 1984.
40. J.A. Trangenstein, *Numerical solution of elliptic and parabolic partial differential equations*, Cambridge University Press, New York, 2013.
41. F.H. Zeng, I. Turner, and K. Burrage, A stable fast time-stepping method for fractional integral and derivative operators, *J. Sci. Comput.*, 77 (2018), pp. 283–307.
42. J.L. Zhang, Z.W. Fang, and H.W. Sun, Exponential-sum-approximation technique for variable-order time-fractional diffusion equations, *J. Appl. Math. Comput.*, DOI: <https://doi.org/10.1007/s12190-021-01528-7>, 2021.
43. X. Zhao, Z.Z. Sun, and G.E. Karniadakis, Second-order approximations for variable order fractional derivatives: Algorithms and applications, *J. Comput. Phys.*, 293 (2015), pp. 184–200.
44. X.C. Zheng and H. Wang, Optimal-order error estimates of finite element approximations to variable-order time-fractional diffusion equations without regularity assumptions of the true solutions, *IMA J. Numer. Anal.*, 41 (2021), pp. 1522–1545.
45. H. Wang and X.C. Zheng, Wellposedness and regularity of the variable-order time-fractional diffusion equations, *J. Math. Anal. Appl.*, 475 (2019), pp. 1778–1802.
46. P. Zhuang, F.W. Liu, V. Anh, and I. Turner, Numerical methods for the variable-order fractional advection-diffusion equation with a nonlinear source term, *SIAM J. Numer. Anal.*, 47 (2009), pp. 1760–1781.