



A static memory sparse spectral method for time-fractional PDEs

Timon S. Gutleb*, José A. Carrillo

Mathematical Institute, University of Oxford, Oxford, OX2 6GG, England, United Kingdom

ARTICLE INFO

Keywords:

Time-fractional
Caputo derivative
Fractional derivative
Fractional PDE
History-free
Sparse
Banded
Spectral method
Orthogonal polynomials

ABSTRACT

We discuss a method which provides accurate numerical solutions to fractional-in-time partial differential equations posed on $[0, T] \times \Omega$ with $\Omega \subset \mathbb{R}^d$ without the excessive memory requirements associated with the nonlocal fractional derivative operator. Our approach combines recent advances in the development and utilization of multivariate sparse spectral methods as well as fast methods for the computation of Gauss quadrature nodes with recursive non-classical methods for the Caputo fractional derivative of general fractional order $\alpha > 0$. An attractive feature of the method is that it has minimal theoretical overhead when using it on any domain $\Omega \subset \mathbb{R}^d$ on which an orthogonal polynomial basis is available. We discuss the memory requirements of the method, present several numerical experiments demonstrating the method's performance in solving time-fractional PDEs on intervals, triangles and disks and derive error bounds which suggest sensible convergence strategies. As an important model problem for this approach we consider a type of wave equation with time-fractional dampening related to acoustic waves in viscoelastic media with applications in the physics of medical ultrasound and outline future research steps required to use such methods for the reverse problem of image reconstruction from sensor data.

1. Introduction

The Caputo derivative [1] of order $\alpha > 0$, $\alpha \notin \mathbb{N}$

$$\frac{\partial^\alpha}{\partial t^\alpha} f(t) = D_C^\alpha f(t) = \frac{1}{\Gamma(\lceil \alpha \rceil - \alpha)} \int_0^t (t-s)^{(\lceil \alpha \rceil - \alpha - 1)} f^{(\lceil \alpha \rceil)}(s) ds, \quad (1)$$

is a fractional generalization of the conventional integer order derivative and appears in many mathematical models in the natural science and medical applications [2–9]. There are multiple non-equivalent generalizations of the fractional derivative which have received the attention of mathematical research. In this paper, we focus on Caputo type (as opposed to e.g. Riemann-Liouville type) fractional derivatives exclusively.

In the general case of non-integer α the Caputo derivative is nonlocal and computing numerical solutions to fractional differential equations frequently more closely resembles integral rather than differential equation solvers. Treated naively, the nonlocality forces us to use an approach in which the history of the function at each point in time t must be stored – in principle indefinitely. Since fractional differential equations appear frequently in models used in the natural sciences (some of which we will mention below)

* Corresponding author.

E-mail address: timon.gutleb@maths.ox.ac.uk (T.S. Gutleb).

where the desired spacetime dimension is usually $3 + 1$ there has been a great deal of interest in the literature in mitigating the memory accumulation problem inherent to time-fractional differential equations.

In this paper we introduce a nonclassical recursive and static memory ('history-free') method for time-fractional partial differential equations on any domain on which sparse spectral methods for classical PDEs are available, with minimal theoretical overhead when adapting it to new problems and domains.

A full discussion of numerical methods available for the treatment of fractional PDEs is impossible to give in the context of a single paper and entire monographs have been written on the topic, see e.g. [10–12]. We nevertheless aim to provide a brief overview of popular approaches for time-fractional PDEs and in particular aim to mention instances where spectral methods were used in such a context: Classical finite difference time-stepping approaches for time-fractional differential equations end up with $O(N^2)$ computational complexity and $O(N)$ memory requirements for N time steps, since as mentioned they effectively require remembering the solution at each previous time-step [10]. Coupling higher dimensional spatial discretizations with such methods thus quickly becomes computationally untenable as the complexity is multiplied by the number of spatial grid points. The computational complexity may be improved by using a sum of exponentials (SOE) approach as described e.g. in [10, Section 4.2], which when used in conjunction with appropriately chosen memory cutoffs N_ℓ can reduce the memory footprint to $O(N_\ell)$ and the computational complexity to $O(NN_\ell)$. Somewhat simpler approaches which also rely on what has been termed the "short memory principle" were discussed in great detail in [13]. A classical reference for a critique as well as proposed improvements for the short memory principle in the form of a *logarithmic* memory principle can be found in [14]. We skip over a contemporary approach to the above references, the nonclassical method due to Yuan and Agrawal [15], as it is the topic of section 3.

Spectral methods for fractional calculus problems of various kinds have recently been successfully used for a number of different problems: A sparse spectral method using a non-orthogonal sum space basis for ODEs involving Caputo or Riemann–Liouville fractional derivatives was introduced by Hale and Olver in [16]. More recently, a spectral method for such ODEs was proposed by Pu and Fascondini in [17] which improved upon the stability of Olver and Hale's method by relying on specially constructed orthogonal functions at the cost of no longer resulting in sparse systems. Despite their efficiency and accuracy, neither of these spectral methods allow memory free computations for time-fractional PDEs, which is a potentially critical limitation for real world applications in higher dimensions. A relevant model of such an equation which we consider in some detail in this paper is a lossy wave equation model for acoustic waves in viscoelastic media as e.g. discussed in [3,2] for the purposes of developing medical ultrasound treatment devices. Colbrook and Ayton [18] recently introduced a competitive contour method for time-fractional PDEs based on taking inverse Laplace transforms. Among the advantages of Colbrook and Ayton's method are tight error control and excellent convergence properties. As a drawback one must be able to make a suitable choice of contour which requires knowledge about the singularities of the involved Laplace transform.

The primary motivation of this paper is the development of efficient and accurate solvers for the above-mentioned viscoelastic fractional wave equation [3] (see also [2]) with applications in medical ultrasound treatment devices:

$$\frac{1}{c_0^2} \frac{\partial^2}{\partial t^2} f - \Delta f - \tau \frac{\partial^\alpha}{\partial t^\alpha} \Delta f = 0.$$

The main domain of interest for this problem is a union of three-dimensional ball and stacked spherical shells of varying thickness as an initial model for the human skull and brain. State-of-the-art methods for this problem, cf. [3,19], replace the time-fractional derivative with a spatially nonlocal fractional Laplacian $(-\Delta)^\alpha$ [20,21]. While this results in a history-free method, it also introduces an inherent systematic approximation error prior to any considerations of error in the numerical solver. While the full multi-domain model presents additional difficulties which we intend to address in a follow-up paper, we discuss a closely related stand-in on the two-dimensional disk in the numerical experiments in Section 7.4 to demonstrate our method's applicability for this class of problems.

This paper is structured as follows: In Section 2 we give a brief introduction into the theory of and recent advances made in sparse spectral methods. Section 3 provides a brief history of nonclassical methods for Caputo derivatives. Section 4 shows how to construct recursive history-free sparse spectral methods for time-fractional Caputo PDEs. In Section 5 we consider the question of how much stored data the introduced static memory method requires. Finally, we give an error analysis of the method in Section 6 and present five detailed numerical experiments in Section 7 which discuss in practice properties such as error, stability as well as applicability in higher dimensional domains. We conclude with an outlook on future research directions.

2. Sparse spectral methods

In this section we introduce and discuss the building blocks of a class of numerical algorithms for solving differential and integral equations called sparse spectral methods. The specific type of spectral method we describe in this paper belongs to the class of *ultraspherical* spectral methods as described in a 2013 paper by Olver and Townsend [22], where the 'ultraspherical' is a reference to the ultraspherical polynomials (aka *Gegenbauer* polynomials) used therein. A more recent thorough treatment of these methods can be found in [23]. The basic idea behind these methods is to expand functions in bases of orthogonal polynomials in such a way that the linear operators appearing in the problem, e.g. differentiation and integration, have banded matrix forms.

We call a complete set of degree-ordered polynomials $\{p_j(x)\}_{j \in \mathbb{N}_0}$ *orthogonal* with respect to a positive Borel measure $\mu(x)$ if the polynomials satisfy

$$\langle p_n, p_m \rangle_\mu = \int_{\mathbb{R}} p_n(x) p_m(x) d\mu(x) = c_{nm} \delta_{nm},$$

where δ_{nm} is the Kronecker delta and c_{nm} are constants independent of x . If $c_{nm} = 1$ for all n, m then the polynomials are further referred to as *orthonormal*. Throughout this paper we adopt the notation

$$\mathbf{P}(x) = (p_0(x) \ p_1(x) \ p_2(x) \ \cdots).$$

The advantages of this notation is that polynomial expansion of a sufficiently well-behaved function in an orthogonal polynomial basis can be written concisely as

$$f(x) = \mathbf{P}(x)\mathbf{f} = \sum_{k=0}^{\infty} f_k p_k(x).$$

In a computing context these function approximations are truncated at finite degree K . In addition to the above characterization, it is natural to speak of orthogonal polynomials as living on a characteristic domain $\Omega = \text{supp}(\mu)$ determined by the support of the measure. For our purposes the measure $\mu(x)$ will always have an associated density function such that we have $d\mu(x) = w(x)dx$ where dx is the Lebesgue measure and we call $w(x)$ the weight function or simply the weight associated with the orthogonal polynomials. Classical examples of orthogonal polynomials are the well-known Legendre, Chebyshev, Gegenbauer, Jacobi, Laguerre and Hermite polynomials. Of particular importance for spectral methods are the Jacobi polynomials $\mathbf{P}^{(a,b)}(x)$, a two-parameter family of polynomials (with Legendre, Chebyshev, Gegenbauer as special cases) which are orthogonal with respect to the weight function $w(x) = (1-x)^a(1+x)^b$ with $a, b > -1$ on $\Omega = [-1, 1]$.

Finally, we comment on how to solve integro-differential equations using sparse spectral methods. First, we note that differentiation

$$\frac{d}{dx}f(x) = \mathbf{P}(x)\mathbf{D}\mathbf{f},$$

as well as integration are known to satisfy recurrences and are thus banded operators when mapping between appropriate polynomial spaces, see e.g. the explicit forms given in [24, 18.9]. As an explicit example we consider the entries of \mathbf{D} for Jacobi polynomials and observe that the banded first derivative acts a shift operator between bases as follows [24, 18.9.15]:

$$\frac{d}{dx}P_n^{(a,b)}(x) = \frac{(n+a+b+1)}{2}P_{n-1}^{(a+1,b+1)}(x).$$

Note from the above that obtaining banded derivative operators requires careful choice of target basis (it will generally not be the same as the original basis) and may thus necessitate the use of banded conversion operators [24, 18.9] elsewhere in the equation where no derivatives operators are present. We will see examples of how this is done in practice in the numerical experiments of this paper. To solve differential equations one usually also requires a way to enforce boundary conditions. In the context of sparse spectral methods this is usually achieved by appending point-evaluation functionals, which take the form of row vectors acting on coefficients, to the linear operator and the corresponding boundary values to the right-hand side, cf. e.g. [22,25]. Alternatively one can also build the boundary conditions directly into the basis, cf. [26].

In Section 7 we describe numerical experiments of our method on triangle and disk domains. We thus note that the discussions in this section can be straightforwardly generalized to higher dimensional domains by using multivariate orthogonal polynomials $\mathbf{P}(\mathbf{x})$, cf. [26]. In general higher dimensional domains each variable x, y etc. has its own associated block-banded multiplication operator X, Y and so on and PDEs with non-zero boundary conditions may be solved using restriction operators analogously to the use of point evaluation operators mentioned above.

We close this section with an overview of recent progress on these banded spectral methods on various domains of interest: Since the seminal report on ultraspherical spectral methods in [22] a flurry of papers have been published to expand the range of problems that they can be applied to by deriving recurrence relationships which result in banded operators for various types of problems, see e.g. [27–33]. A directly related line of recent research has concerned the construction of efficiently computable orthogonal polynomial bases on various multivariate domains for the purpose of using them in such sparse spectral methods, recent successful examples of which include disks [34], triangles and simplices [26,35,36], disk slices and trapeziums [37], spherical caps [38], wedges [39], surfaces of revolution [40] as well as quadratic and cubic curves [41,42]. For a recent reference on multivariate orthogonal polynomials on classical higher dimensional domains such as balls, simplices and cubes we refer to the monograph on the subject by Dunkl and Xu [43]. We refer to [44] for a detailed discussion of computational aspects relating to constructing orthogonal polynomials with respect to non-classical weight functions, which can be used to construct orthogonal polynomials on non-classical domains such as annuli and spherical bands.

Finally, we mention as part of our motivation for using sparse spectral methods the possibility of adapting such methods in the form of spectral element methods which instead of being posed on a single domain work by deconstructing a domain into a mesh of simpler form – usually by means of triangles as done in [26]. For the motivating equation of interest, the natural domain decomposition is to split the sphere model of the skull into ball and spherical shell domains along the discontinuities. Importantly, the option of using spectral element methods with minimal theoretical modifications means that one can alleviate some of the usual criticisms levied against sparse spectral methods in requiring significant amounts of regularity from both the domain and the solution since a function then only needs to be well approximated by piecewise polynomials on a mesh, allowing the use of discontinuities as well as mesh boundary conditions.

3. Previous work on non-classical methods for Caputo derivatives

The method we describe in this paper belongs to a family of non-classical numerical algorithms for time-fractional differential equations whose origin can be traced back to a 2002 paper by Yuan and Agrawal [15]. In this section we provide an overview of the development of these methods without claiming historical completeness. Put concisely, Yuan and Agrawal made use of the following result:

Theorem 1 (Generalized Yuan–Agrawal–Caputo derivative). *Let $\alpha > 0$, $\alpha \notin \mathbb{N}$ and $f \in C^{[\alpha]}[0, T]$. Then the Caputo fractional derivative of f can be expressed as*

$$\frac{\partial^\alpha}{\partial t^\alpha} f(t) = \int_0^\infty \phi_f(w, t) dw,$$

where the function $\phi_f : (0, \infty) \times [0, T] \rightarrow \mathbb{R}$ is defined by

$$\phi_f(w, t) := \frac{(-1)^{[\alpha]} 2 \sin(\pi\alpha)}{\pi} w^{2\alpha-2[\alpha]+1} \int_0^t e^{-w^2(t-\tau)} \frac{\partial^{[\alpha]}}{\partial \tau^{[\alpha]}} f(\tau) d\tau.$$

Furthermore, for fixed $w > 0$ the function $\phi_f(w, t)$ satisfies the (non-fractional) differential equation

$$\frac{\partial}{\partial t} \phi_f(w, t) = -w^2 \phi_f(w, t) + \frac{(-1)^{[\alpha]} 2 \sin(\pi\alpha)}{\pi} w^{2\alpha-2[\alpha]+1} \frac{\partial^{[\alpha]}}{\partial t^{[\alpha]}} f(t), \quad (2)$$

with initial condition $\phi(w, 0) = 0$.

Strictly speaking Yuan and Agrawal themselves only proved this result for $0 < \alpha < 1$, with an extension for $1 < \alpha < 2$ being provided by Trinks and Ruge [45] in the same year. The fully general version we reproduced above was proved by Diethelm in 2008 [46]. The Yuan–Agrawal method for solving equations involving the Caputo fractional derivative computes $\phi_f(w, t)$ via the differential equation (2) and then uses Gauss–Laguerre quadrature for the half-infinite integral. As all of the resulting terms then only depend on the functions f and ϕ_f at the previous time step this eliminates the need to store the past of f in memory in exchange for the additional computational effort of computing ϕ_f and its half-infinite integral. A similar way of rewriting the Caputo fractional derivative using a slightly different integrand function $\phi_*(w, t)$ was introduced by Chatterjee [47] and also later analyzed by Diethelm [46] but we will not further discuss this alternative in this paper.

The Yuan–Agrawal method initially received mixed responses in the literature due to its poor convergence properties [48,49], specifically in relation to the number of quadrature points required for the Gauss–Laguerre step, but also resulted in a number of papers being written with the specific intention of improving it. Lu and Hanyga split the half-infinite integral into a region of $[0, 1]$ and $[1, \infty)$ and used Gauss–Jacobi and Gauss–Laguerre respectively to improve the convergence behavior of the method [50,51,49]. A full explanation of the poor convergence property and error analysis was later given by Diethelm [46]. We also note a more recent exploration of positive and negative examples of these non-classical approaches in [52]. In [46], Diethelm argued for a modification of the Yuan–Agrawal method which substantially improved its convergence properties by using a transformation of the function $\phi_f : (0, \infty) \times [0, T] \rightarrow \mathbb{R}$ to a function $\bar{\phi}_f : (-1, 1) \times [0, T] \rightarrow \mathbb{R}$ inspired by Gautschi’s classical work on quadrature methods for half-infinite integrals [53]:

$$\bar{\phi}_f(\kappa, t) := 2(1-\kappa)^{-\bar{\alpha}}(1+\kappa)^{\bar{\alpha}-2} \phi_f\left(\frac{1-\kappa}{1+\kappa}, t\right). \quad (3)$$

We note the connection of this idea to Cayley transforms which map the right half plane to the unit disk, cf. [54]. In one dimension this changes the integral on the half-infinite Laguerre domain to an integral on the Jacobi domain:

$$\int_0^\infty \phi_f(w, t) dw = \int_{-1}^1 (1-\kappa)^{\bar{\alpha}}(1+\kappa)^{-\bar{\alpha}} \bar{\phi}_f(\kappa, t) d\kappa,$$

with $\bar{\alpha} := 2\alpha - 2[\alpha] + 1 \in (-1, 1)$. Diethelm’s method thus uses Gauss–Jacobi quadrature instead of Gauss–Laguerre quadrature which as analyzed in [46] substantially improves the method’s convergence in terms of the number of quadrature points needed.

In 2010 Birk and Song [55] published a paper which provides an alternative point of view on the Yuan–Agrawal and Diethelm methods. Rather than working in the time domain, Birk and Song derive a frequency domain non-classical method for Caputo derivatives. It is widely known, see e.g. [56,57,55], that the Caputo derivative has the Fourier multiplier form

$$\mathcal{F}\left\{\frac{\partial^\alpha}{\partial t^\alpha} f(t)\right\}(\omega) = (i\omega)^\alpha \mathcal{F}\{f(t)\}(\omega), \quad (4)$$

if $\forall t < 0 : f(t) = 0$, with the Fourier transform defined by

$$\mathcal{F}\{f(t)\}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\omega t} f(t) dt.$$

Table 1

Parameter values for the L -point Gauss–Laguerre and respectively Gauss–Jacobi methods corresponding to Yuan–Agrawal’s, Diethelm’s and Birk–Song’s methods in Equation (6). Note that λ_j and p_j in each row denote the appropriate weights and abscissae of the quadrature methods, cf. [55, Section 4.2].

method	quadrature type	A_j	s_j
Yuan–Agrawal [15]	Gauss–Laguerre	$\frac{(-1)^{ \alpha } 2 \sin(\pi\alpha)}{\pi} e^{p_j} \lambda_j p_j^{\tilde{\alpha}}$	p_j
Diethelm [46]	$(1-x)^{\tilde{\alpha}}(1+x)^{-\tilde{\alpha}}$	$\frac{(-1)^{ \alpha } \pi \sin(\pi\alpha)}{\pi} \frac{4\lambda_j}{(1+p_j)^2}$	$\frac{1-p_j}{1+p_j}$
Birk–Song [55]	$(1-x)^{2\tilde{\alpha}+1}(1+x)^{1-2\tilde{\alpha}}$	$\frac{(-1)^{ \alpha } \sin(\pi\alpha)}{\pi} \frac{8\lambda_j}{(1+p_j)^4}$	$\frac{(1-p_j)^2}{(1+p_j)^2}$

Birk and Song [55] show that we can express $(i\omega)^\alpha$ by the following integral:

$$(i\omega)^\alpha = \frac{(-1)^{|\alpha|} 2 \sin(\pi\alpha)}{\pi} \int_0^\infty \frac{(i\omega)^{[\alpha]}}{i\omega + p^2} p^{\tilde{\alpha}} dp, \quad (5)$$

where $\tilde{\alpha} := 2\alpha - 2[\alpha] + 1$. Note that $\tilde{\alpha} \in (-1, 1)$ for any non-integer order of fractional differentiation $\alpha > 0$, $\alpha \notin \mathbb{N}$. This result has a natural interpretation as the frequency domain equivalent of the time domain Yuan–Agrawal form, cf. [55, Appendix A]. Birk and Song [55] note that the approaches of Yuan–Agrawal and Diethelm can both be understood as resulting from a quadrature rule being applied to Equation (5):

$$(i\omega)^\alpha \approx (i\omega)^{[\alpha]} \sum_{j=1}^L \frac{A_j}{i\omega + s_j^2}, \quad (6)$$

with A_j, s_j taking different values depending on whether an L -point Gauss–Laguerre or Gauss–Jacobi quadrature is used, see Table 1. The third row in Table 1 corresponds to an alternative Gauss–Jacobi quadrature option discussed by Birk and Song in [55] which they show has advantages in some frequency domains and disadvantages in others.

4. Towards a recursive non-classical sparse spectral method

While the approach discussed in the previous section eliminates the requirement to memorize all previous states of the function $f(t)$ to which we apply the Caputo derivative, this is achieved by introducing additional internal variables. A recursive way to alleviate this drawback was suggested by Birk and Song [55]. They use Equation (4) and the inverse Fourier transform of the rational approximation in (6) to obtain

$$\frac{\partial^\alpha}{\partial t^\alpha} f(t) \approx \sum_{j=1}^L A_j \int_0^t e^{-s_j^2(t-\tau)} \frac{\partial^{[\alpha]}}{\partial \tau^{[\alpha]}} f(\tau) d\tau = \sum_{j=1}^L A_j \psi_j(t), \quad (7)$$

with

$$\psi_j(t) := \int_0^t e^{-s_j^2(t-\tau)} \frac{\partial^{[\alpha]}}{\partial \tau^{[\alpha]}} f(\tau) d\tau. \quad (8)$$

We note that one can arrive at an equivalent expression starting from Yuan–Agrawal’s approach without needing to pass through frequency space. The integrals under the sum, i.e. $\psi_j(t)$, are readily found to satisfy the following recurrence relationship when advancing by a time step Δt :

$$\psi_j(t) = e^{-s_j^2 \Delta t} \psi_j(t - \Delta t) + \int_{t-\Delta t}^t e^{-s_j^2(t-\tau)} \frac{\partial^{[\alpha]}}{\partial \tau^{[\alpha]}} f(\tau) d\tau. \quad (9)$$

Note that we have $\psi_j(0) = 0$ and the right-hand side of (9) only depends on the integral which was already computed in the previous time step as well as an incrementing term without long term memory of either $f(t)$ or $\psi_j(t)$.

Birk and Song show in [55] how the results discussed until this point can be used to solve ODEs involving the Caputo derivative: For simplicity we use $\alpha \in (0, 1)$ for the remainder of the paper such that

$$\frac{\partial^{[\alpha]}}{\partial \tau^{[\alpha]}} = \frac{\partial}{\partial \tau}, \quad \alpha \in (0, 1),$$

as the general α method becomes somewhat obfuscated by notation. Generalizations for higher orders of α are straightforward to obtain.

Approximating $f(t)$ via linear interpolation during each time step, i.e.:

$$f(\tau) \approx f((n-1)\Delta t) \left(1 - \frac{\tau}{\Delta t}\right) + f(n\Delta t) \frac{\tau}{\Delta t}, \quad 0 \leq \tau \leq \Delta t, \quad (10)$$

one readily obtains the following approximation from (9):

$$\psi_j(n\Delta t) \approx \psi_j^n = e^{-s_j^2 \Delta t} \psi_j((n-1)\Delta t) + \frac{(1-e^{-s_j^2 \Delta t})}{s_j^2} \frac{f(n\Delta t) - f((n-1)\Delta t)}{\Delta t}. \quad (11)$$

By the above discussion we thus have

$$\left(\frac{\partial^\alpha f}{\partial t^\alpha} \right) (n\Delta t) \approx \sum_{j=1}^L A_j \left(e^{-s_j^2 \Delta t} \psi_j((n-1)\Delta t) + \frac{(1-e^{-s_j^2 \Delta t})}{s_j^2} \frac{f(n\Delta t) - f((n-1)\Delta t)}{\Delta t} \right). \quad (12)$$

Remark 1. Birk and Song [55] refer to the linear interpolation in (10) as ‘trapezoidal rule’ but while linear interpolation is behind the trapezoidal rule applied to

$$\int_{t-\Delta t}^t f(\tau) d\tau,$$

this is not equivalent to applying the trapezoidal rule to integrate

$$\int_{t-\Delta t}^t e^{-s_j^2(t-\tau)} \frac{\partial}{\partial \tau} f(\tau) d\tau.$$

Naturally the linear interpolation can be replaced with a higher order interpolation to obtain a higher order method at the cost of storing additional function values. In fact any numerical method capable of numerically integrating the right-most term of (9) can be substituted in this approach. In this paper we restrict ourselves to using linear interpolation for notational simplicity but recommend using higher order methods in step size sensitive applications.

Since our aim is to derive a memory free solver for PDEs featuring time-fractional Caputo derivatives the minimal problem of interest involves $\frac{\partial^\alpha}{\partial t^\alpha} f(t, \mathbf{x})$, where f now depends on both space $\mathbf{x} \in \Omega \subset \mathbb{R}^d$ and time $t \in [0, T]$. To obtain a method in which each time step involves the solution of a *sparse* linear system, while simultaneously providing excellent spatial convergence properties we choose to combine Birk–Song’s recursive method with sparse spectral methods as described in Section 2.

At any fixed point in time $n\Delta t$ we represent the function $f(t, \mathbf{x})$ and the L auxiliary node functions $\psi_j(t, \mathbf{x})$ via their coefficient vectors in an orthogonal polynomial basis $\mathbf{P}(\mathbf{x})$ on the desired domain, i.e.:

$$f(n\Delta t, \mathbf{x}) = \mathbf{P}(\mathbf{x}) f(n\Delta t),$$

$$\psi_j(n\Delta t, \mathbf{x}) = \mathbf{P}(\mathbf{x}) \psi_j(n\Delta t).$$

We store only the current states of these functions such that the coefficient vectors \mathbf{f}^n and $\boldsymbol{\psi}_j^n$ approximating $f(n\Delta t)$ and $\psi_j(n\Delta t)$ are overwritten in each step. By Equation (12) we can thus approximate the Caputo derivative (with $\alpha \in (0, 1)$) at a point in time $n\Delta t$ as

$$\left(\frac{\partial^\alpha f}{\partial t^\alpha} \right) (n\Delta t, \mathbf{x}) \approx \mathbf{P}(\mathbf{x}) \sum_{j=1}^L A_j \left(e^{-s_j^2 \Delta t} \boldsymbol{\psi}_j^{n-1} + \frac{(1-e^{-s_j^2 \Delta t})}{s_j^2} (\mathbf{f}^n - \mathbf{f}^{n-1}) \right), \quad (13)$$

with the following recurrence for the auxiliary function coefficients:

$$\boldsymbol{\psi}(n\Delta t, \mathbf{x}) = \mathbf{P}(\mathbf{x}) \boldsymbol{\psi}(n\Delta t) \approx \mathbf{P}(\mathbf{x}) \boldsymbol{\psi}_j^n = \mathbf{P}(\mathbf{x}) \left(e^{-s_j^2 \Delta t} \boldsymbol{\psi}_j^{n-1} + \frac{(1-e^{-s_j^2 \Delta t})}{s_j^2} \frac{\mathbf{f}^n - \mathbf{f}^{n-1}}{\Delta t} \right). \quad (14)$$

To illustrate the method in action we finish this section by working through a toy problem for time-fractional PDEs with constructed analytic solutions. The following time-fractional equation of motion is an ODE arising in the massless one-dimensional Kelvin-Voigt model:

$$kf + c \frac{\partial^\alpha}{\partial t^\alpha} f = g, \quad (15)$$

$$f(0) = 0, \quad (16)$$

where $c > 0$, $k > 0$ and the external force $g(t)$ is given by

$$g(t) = \begin{cases} 0, & \text{if } t = 0, \\ 1, & \text{if } t > 0. \end{cases}$$

This equation has the following known analytic solution, cf. [48,46,55]

$$f(t) = \frac{1}{k} \left(1 - E_{\alpha,1} \left(-\frac{kt^\alpha}{c} \right) \right).$$

Equation (15) was part of a critique of non-classical methods due to Schmidt and Gaul [48] and was also addressed again by Diethelm in [46] as well as by Birk and Song in [55]. With its known explicit solution and relevant historical role in early papers on nonclassical methods this equation has become a popular toy model for numerically evaluating the accuracy of fractional Caputo differential equation solvers. We will use it to derive fractional PDEs with known solutions to use as toy models to test the PDE approach with before moving on to more involved problems in the numerical experiments in Section 7. Multiplying Equation (15) by e^{-x} , we readily obtain

$$k \frac{\partial}{\partial x} f + c \frac{\partial^\alpha}{\partial t^\alpha} f = g, \quad (17)$$

satisfying $f(0, x) = 0$ and where we have defined

$$g(t, x) := g(t)e^x = \begin{cases} 0, & \text{if } t = 0, \\ e^x, & \text{if } t > 0. \end{cases}$$

An explicit solution on \mathbb{R} can be inferred via the equation's relationship with (15):

$$f(t, x) = f(t)e^{-x} = \frac{e^x}{k} \left(1 - E_{\alpha,1} \left(-\frac{kt^\alpha}{c} \right) \right).$$

Using Equation (13) one can then derive the following history-free nonclassical scheme for Equation (17):

$$\begin{aligned} & \mathbf{P}(x) \left(kD + \left(c \sum_{j=1}^L A_j \frac{(1 - e^{-s_j^2 \Delta t})}{s_j^2 \Delta t} \right) \mathbf{I} \right) \mathbf{f}^n \\ &= \mathbf{P}(x) \left(\mathbf{g}^n + \left(c \sum_{j=1}^L A_j \frac{(1 - e^{-s_j^2 \Delta t})}{s_j^2 \Delta t} \right) \mathbf{f}^{n-1} - \left(c \sum_{j=1}^L A_j e^{-s_j^2 \Delta t} \boldsymbol{\psi}_j^{n-1} \right) \right), \end{aligned}$$

where the $\boldsymbol{\psi}_j^{n-1}$ are computed recursively via (14). More precisely, to obtain a *banded* linear system we use the following bases and operators:

$$\begin{aligned} & \mathbf{P}^{(1,1)}(x) \left(kD + \left(c \sum_{j=1}^L A_j \frac{(1 - e^{-s_j^2 \Delta t})}{s_j^2 \Delta t} \right) \mathbf{C} \right) \mathbf{f}^n \\ &= \mathbf{P}^{(1,1)}(x) \mathbf{C} \left(\mathbf{g}^n + \left(c \sum_{j=1}^L A_j \frac{(1 - e^{-s_j^2 \Delta t})}{s_j^2 \Delta t} \right) \mathbf{f}^{n-1} - \left(c \sum_{j=1}^L A_j e^{-s_j^2 \Delta t} \boldsymbol{\psi}_j^{n-1} \right) \right), \end{aligned}$$

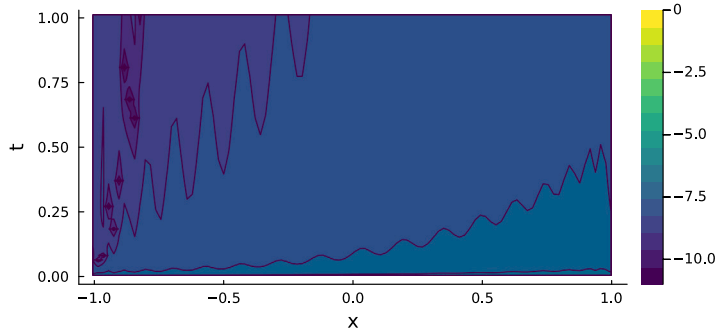
where \mathbf{C} is the banded conversion operator (a change of basis) from the Legendre polynomials $\mathbf{P}^{(0,0)}(x)$ to the Jacobi polynomials $\mathbf{P}^{(1,1)}(x)$ and \mathbf{f} and $\boldsymbol{\psi}$ are expanded in $\mathbf{P}^{(0,0)}(x)$. This change of basis operator is required since for D to be banded in the ultraspherical spectral method it has to map from the basis $\mathbf{P}^{(0,0)}(x)$ to $\mathbf{P}^{(1,1)}(x)$ as discussed in Section 2, cf. [23,22] and [24, 18.9]. As usual for this type of spectral method, the boundary conditions at each time step can be enforced either by using suitable weighted bases or by appending rows corresponding to the evaluation operator to the top of the LHS as well as the boundary condition values to the top of the RHS, see e.g. the methods discussed in [58,16,25] and the references therein. We opt to use the latter option along with the analytic solution values as both the $t = 0$ initial and spatial boundary conditions in this numerical experiment.

In Fig. 1 we show absolute and relative error contour plots for the problem in (17) with $k = 10$ and $c = 100$ for a degree $K = 40$ approximation computed with $L = 50$ quadrature points and $\Delta t = 2^{-20}$ for $t \in [0, 1]$, $x \in [-1, 1]$. Fig. 2 shows the error for $x \in [-1, 1]$ at time $t = 1$. Stepping through 1000 time steps with $L = 50$ and $K = 40$ takes roughly 29 milliseconds on a Lenovo ThinkPad X1 Carbon laptop with a 12th Gen Intel i5-1250P CPU.

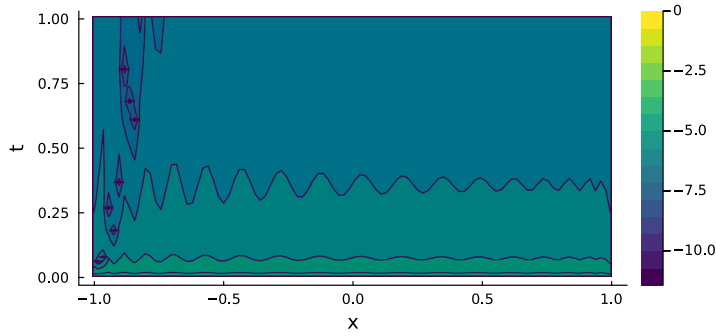
Remark 2. We note that while in the worked example this section we let c and k be constants independent of t and x due to the availability of analytic solutions, the nonclassical method we propose in this paper readily accommodates spacetime dependent coefficients by e.g. replacing the scalar coefficient $k(t, x)$ by a multiplication operator $K^n(X)$ at time $n\Delta t$. If $k(t, x)$ is polynomial in x , this operator does not break the sparsity of the method since polynomial multiplication operators are banded, cf. [23]. Well-behaved non-polynomial coefficient functions, while in theory resulting in dense multiplication operators, may in certain circumstances nevertheless decay rapidly off-band and can thus sometimes still be approximated by banded matrices. If $k(t, x) = k(x)$ is independent of t the banded multiplication operator may be stored in memory for significant performance improvements.

As a final remark for this section, we note that the domain $\Omega \subset \mathbb{R}^d$ determines what constitutes a suitable choice of basis functions and that the method can thus be used in any dimensions as long as a basis suitable for computation is available on the desired domain (although for dimensions ≥ 5 the curse of dimensionality may cause issues without use of further structure such as symmetries). Canonical higher dimensional domains such as higher dimensional cuboids can be treated using tensor products of Jacobi polynomials while arbitrary dimensional ball domains can make use of higher dimensional generalizations of Zernike polynomials. For further examples we refer to the literature overview in Section 2.

As we have seen, to ensure that each step is a sparse linear solve, the left and right-hand side orthogonal polynomial bases must be chosen carefully such that any other linear operators appearing in the fractional partial differential equation are accounted for



(a) absolute error



(b) relative error

Fig. 1. Contour plots of the absolute and relative errors for the fractional PDE in Equation (17) with $k = 10$, $c = 100$ on the spacetime box $[0, 1] \times [-1, 1]$. The legend is logarithmic, indicating the order of magnitude of the errors. The numerical solution was computed with $L = 50$ quadrature points, approximation degree $K = 40$ and with stepsize $\Delta t = 2^{-20}$. See Fig. 2 for a slice through $t = 1$.

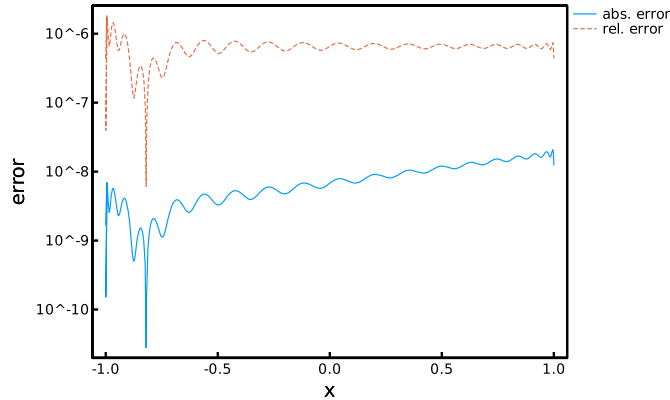


Fig. 2. Absolute and relative error slice at $t = 1$ for the experiment in Fig. 1.

consistently, typically also requiring sparse conversion operators in the process. Due to the many possible combinations of linear operators one may encounter when dealing with fractional PDEs, we show the procedure of ensuring sparsity in more cases in the numerical experiments in Section 7.

5. Memory cost of the history-free method

In this section we investigate how much memory the introduced static memory method requires. We focus only on the memory cost strictly associated with computing the Caputo fractional derivative as the exact cost of the method will otherwise naturally vary depending on the other present differential or integral operators. We count the minimal required values in storage for a solution

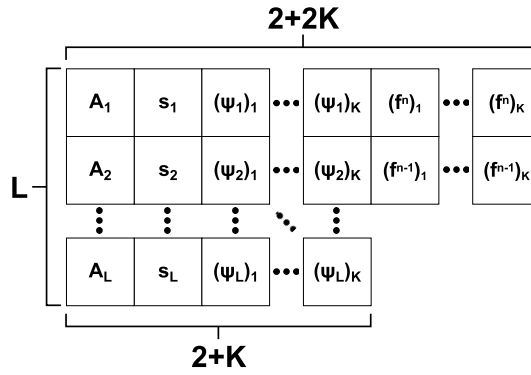


Fig. 3. Diagram of memory cost associated with solving the Caputo derivative part of a fractional PDE using the method described in this paper, where K is the length of the coefficient vectors in the orthogonal polynomial basis of choice and L is the number of quadrature points chosen in Eq. (13). The number of values required in storage is $L(2 + K) + 2K$.

loop in a diagram in Fig. 3, where K is the length of the coefficient vectors in the orthogonal polynomial basis of choice and L is the number of quadrature points chosen in Eq. (13). The number of values required in storage is found to be $L(2 + K) + 2K$ and we observe that this memory requirement is entirely independent of the number of time steps taken to reach the final desired time $T = N\Delta t$. We also note that for most problems of interest one expects $L \ll N$. In practice the method's efficiency and accuracy can benefit from keeping some additional repeatedly used static values in memory but this additional cost is also $O(L)$ rather than $O(N)$, which is the typical cost associated with classical methods cf. [10]. Whether the obtained memory footprint is smaller than that of a short memory principle approach will vary on a problem by problem basis, as the cutoff at which the incurred error remains sensible varies – the question boils down to how the smallest sensible memory cutoff N_ℓ compares to the required number of quadrature points L (note that nonclassical methods do not rely on short term memory cutoffs) and how much is known about the expected behavior of the solutions.

6. Error analysis and computational complexity

In this section we provide a discussion of the error as well as computational complexity associated with our method. A general discussion of the error of both the original Yuan-Agrawal method and Diethelm's modification was already given by Diethelm himself in Theorem 11 of [46]. Note that, disregarding the independent spatial error, Diethelm's proof assumes that the differential equation in Eq. (2) is solved for the auxiliary functions rather than the more efficient recurrence relationship approach. If we were to assume that the recursive approach performs at least as well as a differential equation solver in each step then Diethelm's error estimate holds similarly for our suggested solver as long as the independent spatial error is also taken into account. The numerical stability of recurrence relationships is a subtle matter which has received a large amount of attention in the literature, see e.g. [59]. In Section 7 we will present some numerical experiments regarding the stability of the recurrence relationship to convince ourselves of the applicability of the proposed method.

In what follows we provide an error analysis for the recursive approach to time-fractional Caputo derivatives including resolving spatial dependence in a basis of orthogonal polynomials. The results which follow apply for the direct computation of the Caputo derivative rather than fractional differential equations, keeping in mind that for a particular fractional PDE the error will also be affected by what method we use to deal with the remaining differential or integral operators. For the sake of brevity we also only discuss Diethelm's quadrature rule and fractional orders $0 \leq \alpha \leq 1$ but note that analogous results for the Birk–Song quadrature rule and higher orders can be derived in the same way.

Lemma 1. Let $\bar{P}_n^{(\bar{\alpha}, -\bar{\alpha})}(x)$ with $\bar{\alpha} := 2\alpha - 2[\alpha] + 1 \in (-1, 1)$ denote the monic Jacobi polynomial of order n and $\langle \cdot, \cdot \rangle_{(a,b)}$ denote the inner product with respect to the corresponding Jacobi weight, i.e.:

$$\langle f, g \rangle_{(a,b)} := \int_{-1}^1 f(t)g(t)(1-t)^a(1+t)^b dt.$$

Furthermore, let $\bar{\phi}_{T,x}(\kappa) := \bar{\phi}_f(\kappa, T, x)$ with $\bar{\phi}_f$ defined in Eq. (3). If $\bar{\phi}_{T,x} \in C^{2L}[0, T]$ then the quadrature error incurred from Diethelm's method as described in Section 3 is given by

$$\left(\frac{\partial^\alpha}{\partial t^\alpha} f \right)(T, x) - \sum_{j=1}^L A_j \psi_j(T, x) = \frac{\bar{\phi}_{T,x}^{(2L)}(\xi)}{(2L)!} \langle \bar{P}_L^{(\bar{\alpha}, -\bar{\alpha})}, \bar{P}_L^{(\bar{\alpha}, -\bar{\alpha})} \rangle_{(\bar{\alpha}, -\bar{\alpha})},$$

for some $\xi \in [0, T]$.

Proof. It is a classical result, cf. [60, Theorem 3.6.24], that if $g \in C^{2L}[a, b]$ quadrature methods with monic polynomials $\bar{P}_n(x)$ orthogonal with respect to $w(x)$ on $[a, b]$ satisfy

$$\int_a^b w(x)g(x)dx - \sum_{j=1}^L \lambda_j g(p_j) = \frac{g^{(2L)}(\xi)}{(2L)!} \langle \bar{P}_L, \bar{P}_L \rangle,$$

for some $\xi \in (a, b)$ where $\langle \cdot, \cdot \rangle$ is the inner product with respect to the appropriate orthogonality weight. To use this result we have to translate our compact notation into a slightly more cumbersome one: Using Table 1 and the definitions in Section 3 we find that for Diethelm's method

$$\sum_{j=1}^L A_j \psi_j(T, \mathbf{x}) = \sum_{j=1}^L \lambda_j \left(\frac{(-1)^{|\alpha|} \sin(\pi\alpha)}{\pi} \frac{4\psi_j(T, \mathbf{x})}{(1+p_j)^2} \right) = \sum_{j=1}^L \lambda_j \bar{\phi}_f(p_j, T, \mathbf{x}).$$

The stated error expression follows as an immediate corollary. \square

Definition 1. We denote by \mathcal{P}_K the canonical projection operator corresponding to truncation-to-degree- K , i.e.:

$$\mathbf{P}(\mathbf{x})\mathcal{P}_K g = \sum_{k=0}^K g_k P_k(\mathbf{x}).$$

Definition 2. We denote by $\text{err}_{\mathbf{P},K}(g(\mathbf{x}))$ the error incurred from truncating the polynomial expansion of a function g in the basis $\mathbf{P}(\mathbf{x})$ at degree K , i.e.:

$$\text{err}_{\mathbf{P},K}(g(\mathbf{x})) = |\mathbf{P}(\mathbf{x})g - \mathbf{P}(\mathbf{x})\mathcal{P}_K g|.$$

While for sufficiently nice functions one generally expects spectral convergence in the truncation degree when expanding functions in orthogonal polynomials, cf. [61, Chapter 21] and [62, Section 2.3], more generally the study of these truncation errors is an active field of research in its own right, see [63–69] and the references therein. We will make no attempt to fully cover the range of different error bounds for all the various different basis choices and domains our method can be applied with and instead settle for an abstract result in which the error caused by the polynomial approximation in space is separated from the time-fractional Caputo contribution.

Proposition 1. Let $f_{\mathbf{x}}(t) := f(t, \mathbf{x})$ with $f_{\mathbf{x}} \in C^{1+[\alpha]}[0, T]$ and $\bar{\phi}_{T,\mathbf{x}}(\kappa) := \bar{\phi}_f(\kappa, T, \mathbf{x})$ with $\bar{\phi}_{T,\mathbf{x}} \in C^{2L}[0, T]$. Furthermore, let f_{α}^N denote the coefficient vector such that

$$\left(\frac{\partial^{\alpha}}{\partial t^{\alpha}} f \right) (T, \mathbf{x}) = \mathbf{P}(\mathbf{x}) f_{\alpha}^N.$$

We assume the quadrature-related constants A_j and s_j are obtained with negligible error and define the following non-negative functions:

$$M(\mathbf{x}) = \max_{\xi_1 \in [0, T]} \left| \bar{\phi}_T^{(2L)}(\xi_1, \mathbf{x}) \langle P_L^{(\bar{\alpha}, -\bar{\alpha})}, P_L^{(\bar{\alpha}, -\bar{\alpha})} \rangle_w \right|,$$

$$C(\mathbf{x}) = \frac{1}{2} \max_j |A_j| \max_{\xi_2 \in [0, T]} \left| \left(\frac{\partial^2}{\partial t^2} f \right) (\xi_2, \mathbf{x}) \right|.$$

Then the error incurred from using the recursive degree- K -truncated spectral method as described in Section 4 to compute the time-fractional Caputo derivative of order $0 \leq \alpha \leq 1$ of a function $f(t, \mathbf{x})$ at time T using Diethelm's quadrature rule and a uniform time grid of step size $\Delta t = \frac{T}{N}$ is bounded by

$$\begin{aligned} & \left| \left(\frac{\partial^{\alpha}}{\partial t^{\alpha}} f \right) (T, \mathbf{x}) - \mathbf{P}(\mathbf{x})\mathcal{P}_K \sum_{j=1}^L A_j \psi_j^N \right| \\ & \leq \frac{M(\mathbf{x})}{(2L)!} + C(\mathbf{x})LT\Delta t + \text{err}_{\mathbf{P},K} \left(\sum_{j=1}^L A_j \psi_j(T, \mathbf{x}) \right). \end{aligned}$$

Proof. Using the triangle inequality we obtain

$$\begin{aligned} & \left| \left(\frac{\partial^{\alpha}}{\partial t^{\alpha}} f \right) (T, \mathbf{x}) - \mathbf{P}(\mathbf{x})\mathcal{P}_K \sum_{j=1}^L A_j \psi_j^N \right| = \left| \mathbf{P}(\mathbf{x})f_{\alpha}^N - \mathbf{P}(\mathbf{x})\mathcal{P}_K \sum_{j=1}^L A_j \psi_j^N \right| \\ & \leq \left| \mathbf{P}(\mathbf{x})f_{\alpha}^N - \mathbf{P}(\mathbf{x}) \sum_{j=1}^L A_j \psi_j^N \right| + \left| \mathbf{P}(\mathbf{x}) \sum_{j=1}^L A_j \psi_j^N - \mathbf{P}(\mathbf{x})\mathcal{P}_K \sum_{j=1}^L A_j \psi_j^N \right|. \end{aligned}$$

Note that the first term, $|\mathbf{P}(\mathbf{x})f_a^N - \mathbf{P}(\mathbf{x})\sum_{j=1}^L A_j \psi_j^N|$, now describes an error contribution at each spatial point \mathbf{x} which is solely incurred due to the recursive time-fractional Caputo derivative method. The second term is simply the approximation error of truncating the polynomial expansion of the function $\sum_{j=1}^L A_j \psi_j(T, \mathbf{x})$ in $\mathbf{P}(\mathbf{x})$ which we leave abstract.

To obtain a more explicit bound for the first term we use Lemma 1 to observe that the error added to the series approximation in (7) caused by perturbations $\delta_j(\mathbf{x})$ in the auxiliary functions $\psi_j(t, \mathbf{x})$ is

$$\begin{aligned} \left(\frac{\partial^\alpha}{\partial t^\alpha} f \right) (T, \mathbf{x}) - \sum_{j=1}^L A_j (\psi_j(T, \mathbf{x}) + \delta_j(\mathbf{x})) \\ = \left(\frac{\partial^\alpha}{\partial t^\alpha} f \right) (T, \mathbf{x}) - \sum_{j=1}^L A_j \psi_j(T, \mathbf{x}) - \sum_{j=1}^L A_j \delta_j(\mathbf{x}) \\ = \frac{\tilde{\phi}_{T, \mathbf{x}}^{(2L)}(\xi)}{(2L)!} \langle P_L^{(\tilde{\alpha}, -\tilde{\alpha})}, P_L^{(\tilde{\alpha}, -\tilde{\alpha})} \rangle_w - \sum_{j=1}^L A_j \delta_j(\mathbf{x}), \end{aligned}$$

for some $\xi \in [0, T]$. From this we obtain the error bound

$$\begin{aligned} \left| \left(\frac{\partial^\alpha}{\partial t^\alpha} f \right) (T, \mathbf{x}) - \sum_{j=1}^L A_j (\psi_j(T, \mathbf{x}) + \delta_j(\mathbf{x})) \right| \\ \leq \max_{\xi \in [0, T]} \left| \frac{\tilde{\phi}_{T, \mathbf{x}}^{(2L)}(\xi_1)}{(2L)!} \langle P_L^{(\tilde{\alpha}, -\tilde{\alpha})}, P_L^{(\tilde{\alpha}, -\tilde{\alpha})} \rangle_w \right| + L \max_j (|A_j| |\delta_j(\mathbf{x})|). \end{aligned}$$

We thus seek a bound on $|\delta_j(\mathbf{x})|$ caused by the approximation in (11). Comparing (9) with the approximation in (11) we find that the δ_j are the errors caused by integrating the derivative of the linear interpolation. If $f(x)$ is in $C^{q+1}[a, b]$ and $p_{f,q}(x)$ is its interpolating polynomial of degree $\leq q$ at $q+1$ points $x_0, \dots, x_q \in [a, b]$, then the remainder term of the interpolation is given by [24, 3.3.11]

$$f(x) - p_{f,q}(x) = \frac{\prod_{i=0}^q (x - x_i)}{(q+1)!} f^{(q+1)}(\xi),$$

for some $\xi \in [a, b]$. For our case of taking derivatives of the linear interpolation in each time step we thus obtain that there exists a $\xi \in [(n-1)\Delta t, n\Delta t]$ such that

$$\left(\frac{\partial}{\partial \tau} f \right) (\tau, \mathbf{x}) - \frac{f(n\Delta t) - f((n-1)\Delta t)}{\Delta t} = \frac{\Delta t + 2(\tau - n\Delta t)}{2} \left(\frac{\partial^2}{\partial \tau^2} f \right) (\xi, \mathbf{x}).$$

Since for all $\tau \in [(n-1)\Delta t, n\Delta t]$ we have $|\frac{\Delta t + 2(\tau - n\Delta t)}{2}| \leq \frac{\Delta t}{2}$ this yields the error bound

$$\left| \left(\frac{\partial}{\partial \tau} f \right) (\tau, \mathbf{x}) - \frac{f(n\Delta t) - f((n-1)\Delta t)}{\Delta t} \right| \leq \frac{\Delta t}{2} \max_{\xi \in [(n-1)\Delta t, n\Delta t]} \left| \left(\frac{\partial^2}{\partial \tau^2} f \right) (\xi, \mathbf{x}) \right|.$$

In a single time step the error in the integral is thus bounded by

$$\begin{aligned} \left| \int_{(n-1)\Delta t}^{n\Delta t} e^{-s_j^2(n\Delta t - \tau)} \left(\left(\frac{\partial}{\partial \tau} f \right) (\tau, \mathbf{x}) - \frac{f(n\Delta t) - f((n-1)\Delta t)}{\Delta t} \right) d\tau \right| \\ \leq \frac{\Delta t}{2} \max_{\xi \in [(n-1)\Delta t, n\Delta t]} \left| \left(\frac{\partial^2}{\partial \tau^2} f \right) (\xi, \mathbf{x}) \right| \int_{(n-1)\Delta t}^{n\Delta t} e^{-s_j^2(n\Delta t - \tau)} d\tau \\ = \frac{\Delta t}{2} \frac{(1 - e^{-s_j^2 \Delta t})}{s_j^2} \max_{\xi \in [(n-1)\Delta t, n\Delta t]} \left| \left(\frac{\partial^2}{\partial \tau^2} f \right) (\xi, \mathbf{x}) \right| \\ \leq \frac{(\Delta t)^2}{2} \max_{\xi \in [(n-1)\Delta t, n\Delta t]} \left| \left(\frac{\partial^2}{\partial \tau^2} f \right) (\xi, \mathbf{x}) \right|, \end{aligned}$$

resulting in the following error bound for $\delta_j(\mathbf{x})$:

$$|\delta_j(\mathbf{x})| \leq \frac{N(\Delta t)^2}{2} \max_{\xi \in [0, T]} \left| \left(\frac{\partial^2}{\partial \tau^2} f \right) (\xi, \mathbf{x}) \right| = \frac{T\Delta t}{2} \max_{\xi \in [0, T]} \left| \left(\frac{\partial^2}{\partial \tau^2} f \right) (\xi, \mathbf{x}) \right|.$$

Combining this with the above-derived error bounds yields the stated result. \square

Remark 3. We note that similarly to what was discussed in [60, Section 3.7], such error bounds for quadrature-type methods tend to not be useful for error estimations in actual applications as they include difficult to compute or estimate high order derivative terms. A more practical approach to estimating errors in an application is to check convergence with higher precision results. Nevertheless, the derived error bounds suggest a sensible convergence strategy: Prioritize convergence in L , while keeping $L\Delta t$ constant, then

adjust Δt further if required once convergence in L has been achieved. We verify that this is a sensible convergence strategy in practice in the numerical experiments in the next section.

As mentioned before, it is straightforward to see from the proof of Proposition 1 that choosing higher order accuracy integration methods for the integrals in the recurrence instead of using (10) would also lead to a higher order in Δt method for the Caputo derivative.

Finally, we address computational complexity and the associated memory footprint. We will suppose here that the desired accuracy is reached with K polynomial coefficients and L quadrature points. First we note that as seen in the numerical experiments in Section 7 and the worked example in Section 4, each time step involves two computations: First, the update of the auxiliary function via the recurrence and second the linear system solve. Additionally, at the beginning of each computation the quadrature-related constants must be computed, although this only has to be done a single time and will thus be considered pre-computation. The recursive update of the auxiliary function is $O(1)$ for each quadrature point and spatial index, cf. Fig. 3, meaning that the computational cost of this update has computational cost $O(KL)$. Meanwhile, the linear solve involves a sparse (banded or block-banded) operator of size $K \times K$ if set up correctly using the ultraspherical spectral method, cf. [22,23]. The exact computational cost of the second part depends on this sparsity pattern but if it is banded then one expects $O(K)$. Setting the pre-computation of the quadrature constants aside the computational cost in each time step is thus dominated by the computational cost $O(KL)$. Using the ultraspherical spectral method to resolve the spatial aspect of a time-fractional PDE the computational cost of our proposed solver is thus effectively determined by the cost of updating the auxiliary functions using the recursion, which is an efficient process. Coupling a memory cutoff method to the ultraspherical spectral method for comparison can be expected to be at best $O(KN_\ell)$ in each timestep where N_ℓ is the amount of time steps which are remembered, cf. [10, Section 4.2]. Whether the proposed method performs better is thus entirely dependent on the significance of the memory effect of a given equation – if the number of time steps N_ℓ required in memory in the short memory principle method exceed the number of quadrature points L needed to achieve the desired accuracy at time T , then the non-classical approach described in this paper will perform better and vice versa. The situation is naturally analogous for the memory footprint. We note however that there are reasons to expect the error behavior of memory cutoff methods to be rather poor in general cases, cf. [14]. Similar observations can be made for the logarithmic memory principle suggested in [14].

7. Numerical experiments

We present four numerical experiments in this section which respectively explore the following subjects:

1. Recursive nonclassical computation of time-fractional Caputo derivatives.
2. Stability of the recurrence relationship for the auxiliary functions ψ_j .
3. A time-fractional heat equation on a triangle.
4. A wave equation dampened by a time-fractional term on a disk.

We note that a worked example which may be considered a numerical experiment was already discussed in Section 4. The numerical experiments 3 and 4 were chosen such that the polynomial bases would not simply consist of tensor product domains. Throughout this paper, when discussing error compared to reference solutions we will be considering the pointwise absolute and relative errors

$$Err_{abs,f}(n, \mathbf{x}) = |\mathbf{P}(\mathbf{x})f^n - f(n\Delta t, \mathbf{x})|,$$

$$Err_{rel,f}(n, \mathbf{x}) = \frac{|\mathbf{P}(\mathbf{x})f^n - f(n\Delta t, \mathbf{x})|}{|f(n\Delta t, \mathbf{x})|},$$

and indicate which is being considered. We furthermore use Birk–Song’s quadrature rule as described in Table 1 throughout this section. As accuracy in the Gaussian quadrature related values A_j and s_j is critical for the method and these values may be precomputed for a given quadrature type without having to ever be re-computed (or one can use lookup tables), we initially use 128-bit floating point numbers and then project them down to 64-bit precision before starting the recurrence loop – all other values in the numerical experiments are generated and stored as 64-bit floating point numbers.

The implementation used for the numerical experiments in this section made use of the open source Julia [70] packages Multi-variateOrthogonalPolynomials.jl [71], FastGaussQuadrature.jl [72] as well as ApproxFun.jl [73].

7.1. Experiment 1: direct computation of Caputo derivatives

While constructed for the purpose of solving differential equations, the recursive non-classical approach can nevertheless straightforwardly be used to numerically approximate the Caputo derivative of a known function at a given point. In this section we compare some cases with known explicit Caputo derivatives with numerical approximations obtained via (12). This idea was also explored by Diethelm in [46] but his numerical experiments did not include error plots and were limited by small quadrature point counts, a problem which we are unencumbered thanks to [74–76] and implementations such as FastGaussQuadrature.jl [72], cf. [77]. We make use of the following explicitly given Caputo derivatives for these numerical experiments, cf. [78, Sec. 3.1]:

$$\frac{\partial^\alpha}{\partial t^\alpha}(t^2) = \frac{2t^{2-\alpha}}{\Gamma(3-\alpha)}, \quad \alpha \in [0, 1], \quad (18)$$

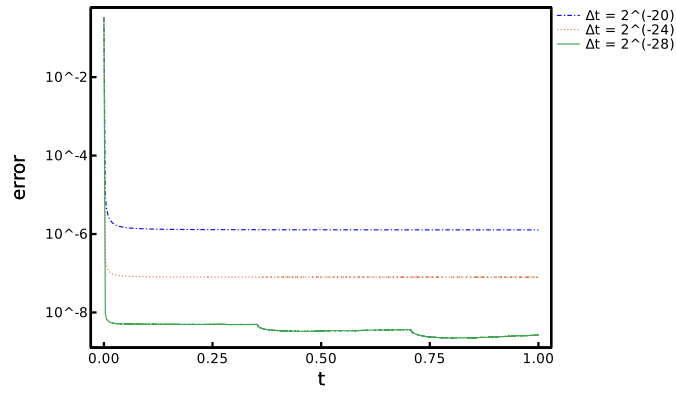
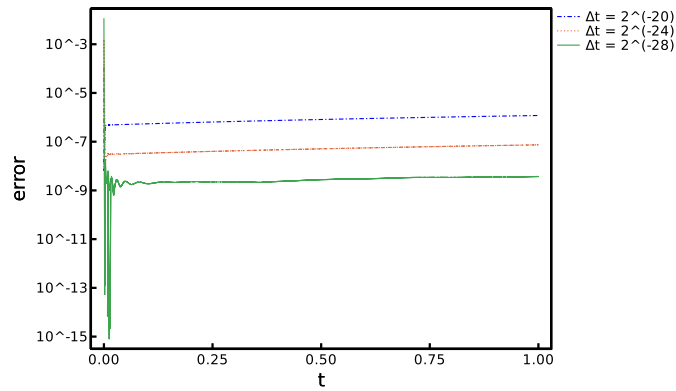
(a) $f(t) = t^2$, $L = 65$, $\alpha = \frac{2}{3}$ (b) $f(t) = e^t$, $L = 60$, $\alpha = \frac{1}{2}$

Fig. 4. Relative errors between analytic and recursively computed Caputo derivatives of indicated $f(t)$ with number of quadrature points L for various time step sizes Δt .

$$\frac{\partial^{1/2}}{\partial t^{1/2}} E_{a,1}(t) = -\frac{1}{\sqrt{t\pi}} + \sum_{k=0}^{\infty} \frac{k! t^{k-\frac{1}{2}}}{\Gamma(k+\frac{1}{2}) \Gamma(1+ak)}, \quad a > 0, \quad (19)$$

where $E_{a,b}(z)$ denotes the two-parameter Mittag-Leffler function [79,80], which is defined by (cf. [24, 10.46.3]):

$$E_{a,\beta}(z) := \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)}.$$

We furthermore have (cf. [81, Chapter 4])

$$E_{1,1}(z) = e^z,$$

which via (19) yields

$$\frac{\partial^{1/2}}{\partial t^{1/2}} e^t = e^t \operatorname{erf}(\sqrt{t}). \quad (20)$$

In Figs. 4 and 5 we show error plots comparing recursively computed numerical approximations to the explicitly known solutions in (18) and (20) with varying time step sizes Δt and quadrature nodes L respectively. Note that as suggested by the error bounds in Section 6 we observe linear improvement in accuracy in Δt once sufficient convergence in L has been achieved.

7.2. Experiment 2: stability of the auxiliary function recurrence

A remaining concern with the proposed method may be that floating point arithmetic can cause recurrence relationships to be unstable. In this section, we thus provide some numerical experiments in which we compare full domain numerical integration in

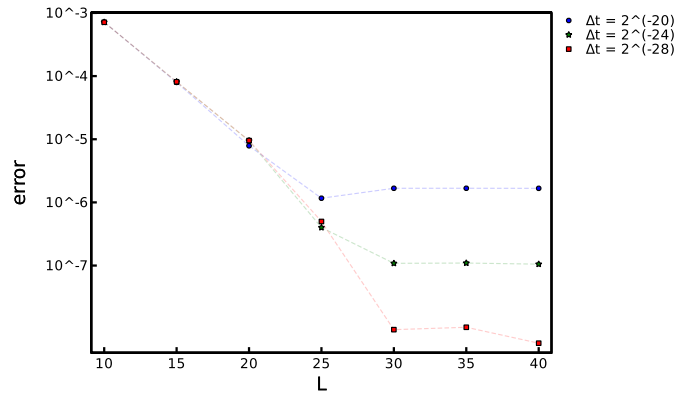
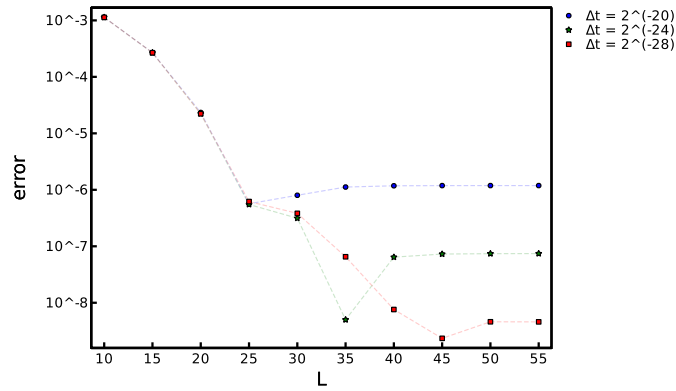
(a) $f(t) = t^2$, $\alpha = \frac{1}{4}$ (b) $f(t) = e^t$, $\alpha = \frac{1}{2}$

Fig. 5. Relative errors between analytic and recursively computed Caputo derivatives of indicated $f(t)$ at the end point of the time domain $t \in [0, 1]$ as a function of the number of quadrature points L used with various time step sizes Δt .

(8) to the recurrence approach in (14) to see how this error behaves in practice, using the explicitly known solutions of the previous numerical experiment as a template.

Figs. 6 and 7 show the error between the recurrence and full domain approaches to computing the auxiliary functions ψ_j . Consistent with the error analysis in Section 6, we observe that if L is low, reducing Δt does not improve the error. As suggested by the error bounds discussion in Section 6, one should thus increase L to convergence before excessively refining the step size Δt .

7.3. Experiment 3: a time-fractional heat equation on the triangle

In this section we use the Prorol polynomials $\mathbf{P}^{(a,b,c)}(x, y)$ which are orthogonal on the triangle domain

$$\Omega_T = \{(x, y) : 0 \leq x, 0 \leq y \leq 1 - x\}$$

with respect to the weight function $w(x, y) = x^a y^b (1 - x - y)^c$, cf. [26,43], to solve the time-fractional heat equation

$$\frac{\partial^\alpha}{\partial t^\alpha} f - k \Delta f = 0, \quad (21)$$

$$f(0, x, y) = f_0(x, y), \quad (22)$$

with zero Dirichlet conditions on the boundary of Ω_T . While non-zero Dirichlet boundary conditions would require the use of restriction operators [26], zero Dirichlet conditions can be enforced by resolving $f(x, y)$ in the weighted Prorol polynomial basis

$$\mathbf{W}^{(1,1,1)}(x, y) = xy(1 - x - y) \mathbf{P}^{(1,1,1)}(x, y).$$

The Laplacian Δ then maps $\mathbf{W}^{(1,1,1)}(x)$ to $\mathbf{P}^{(1,1,1)}(x)$ and a sparse conversion operator C for this basis map is also available, see Fig. 8 for sparsity pattern plots. We thus obtain the recursive non-classical method

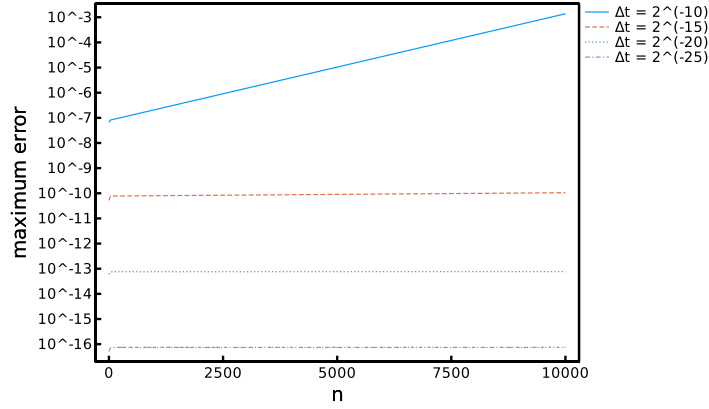


Fig. 6. Maximal-in- j absolute errors at each time step between auxiliary functions ψ_j after n time steps of indicated size Δt computed via the recurrence and full domain quadrature for the problem in (20) with $L = 30$ quadrature points using the Birk-Song quadrature rule. For a high n endpoint experiment see Fig. 7.

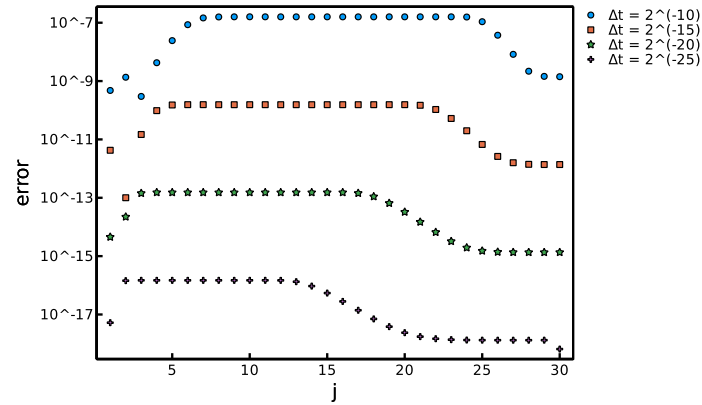


Fig. 7. Absolute errors between auxiliary functions ψ_j , with j on the x -axis, after 10 million time steps of indicated size Δt computed via the recurrence and full domain approach for the problem in (18) with $L = 30$ quadrature points and $\alpha = \frac{1}{2}$ using the Birk-Song quadrature rule. For an error over time experiment see Fig. 6.

$$\begin{aligned} \mathbf{P}^{(1,1,1)}(x, y) \left(-k\Delta + \left(\sum_{j=1}^L A_j \frac{(1 - e^{-s_j^2 \Delta t})}{s_j^2 \Delta t} \right) C \right) f^n \\ = \mathbf{P}^{(1,1,1)}(x, y) C \left(\left(\sum_{j=1}^L A_j \frac{(1 - e^{-s_j^2 \Delta t})}{s_j^2 \Delta t} \right) f^{n-1} - \left(\sum_{j=1}^L A_j e^{-s_j^2 \Delta t} \psi_j^{n-1} \right) \right), \end{aligned} \quad (23)$$

valid for the heat equation with zero Dirichlet boundary conditions on the triangle domain Ω_T and $k > 0$. Note that the solution obtained via the above scheme is

$$f(N\Delta t, x, y) \approx \mathbf{W}^{(1,1,1)}(x, y) f^N.$$

In Fig. 9 we plot numerical solutions obtained for the time-fractional heat equation on the triangle with $k = \frac{1}{500}$ with fractional order $\alpha = \frac{3}{5}$ and initial condition

$$f_0(x, y) = 100xy(1 - x - y)^2 e^{-10((x-0.6)^2 + (y-0.2)^2)}. \quad (24)$$

Fig. 10 shows the corresponding coefficient decay in the polynomial approximation which can be used as a proxy for whether the desired spatial convergence of a given application has been achieved, cf. [26]. Stepping through 1000 time steps with $L = 45$ and $K = 400$ takes roughly 143 milliseconds on a Lenovo ThinkPad X1 Carbon laptop with a 12th Gen Intel i5-1250P CPU.

7.4. Experiment 4: a wave equation dampened by a fractional term on the disk

We consider a problem closely associated with the motivating model used in [2,3] to describe medical ultrasound devices used on the human brain. As mentioned in the introduction, the equation of interest (cf. [3]) is

$$\frac{1}{c_0} \frac{\partial^2}{\partial t^2} f - \Delta f - \tau \frac{\partial^\alpha}{\partial t^\alpha} (\Delta f) = 0,$$

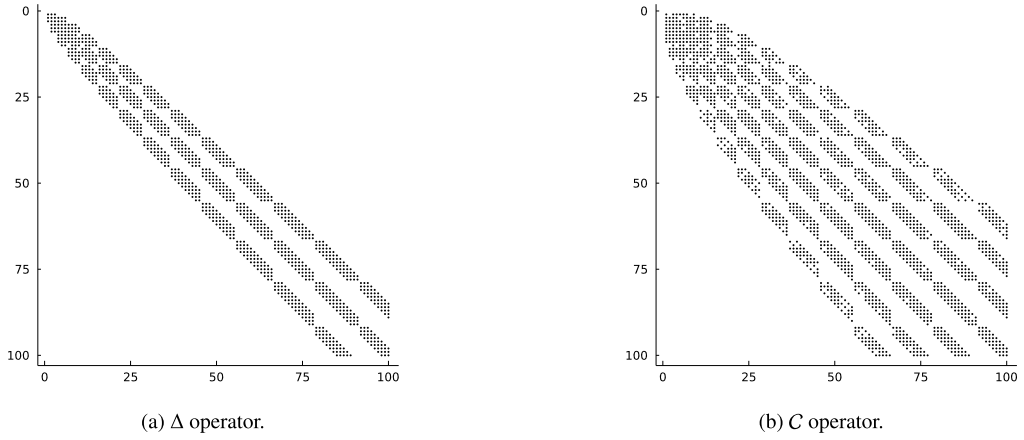


Fig. 8. Sparsity pattern (spy) plots for 100×100 blocks of the Laplacian Δ and conversion operator C appearing in Equation (23). See [26] for how to compute their entries.

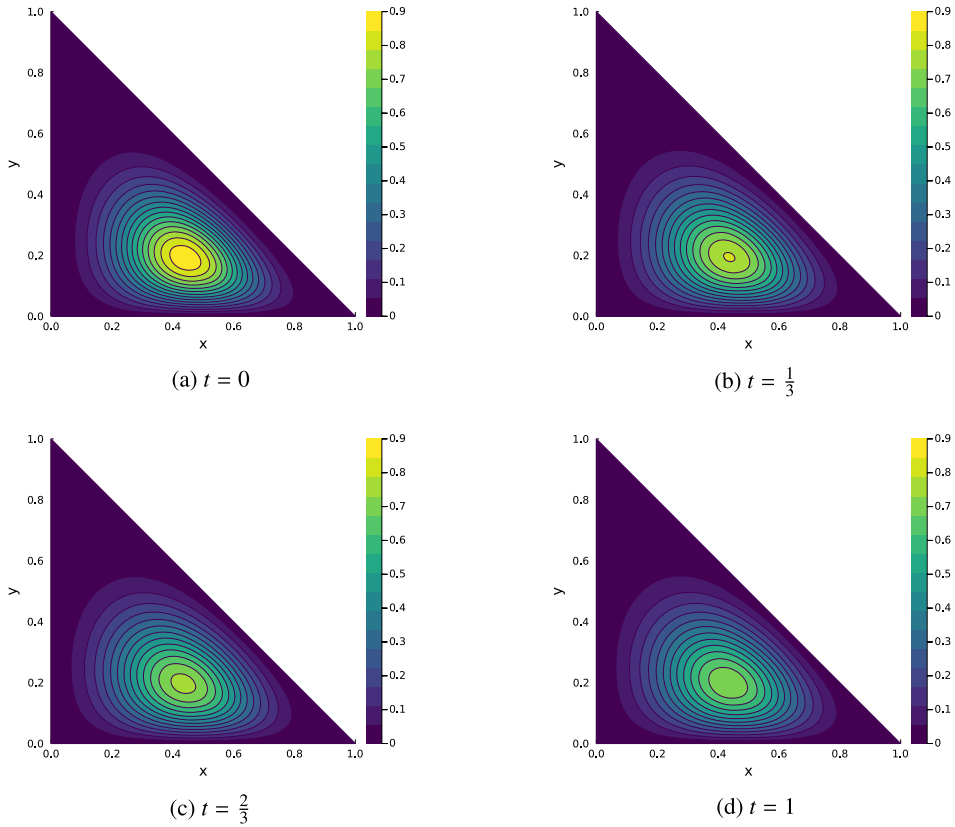


Fig. 9. (b), (c) and (d) show numerical solutions at different times t to the problem in Eq. (21) with $\alpha = \frac{3}{4}$ and $k = \frac{1}{500}$, starting from the initial condition in Eq. (24) which is shown in (a). The num. solutions are degree $K = 400$ approximations computed with $L = 45$ quadrature points and $\Delta t = 2^{-20}$. To make comparison easier the scale on all plots is kept consistent.

posed on the three dimensional ball surrounded by multiple matching spherical shells of varying thickness, cf. the sphere head models described in e.g. [82]. The presence of the term $\frac{\partial^\alpha}{\partial t^\alpha}(\Delta f)$ as well as the requirement to decompose the domain into matching segments add complications to this problem which we intend to address in a separate paper. In this section we discuss the similar but slightly simplified equation

$$\frac{1}{c_0^2} \frac{\partial^2}{\partial t^2} f - \Delta f + \tau \frac{\partial^\alpha}{\partial t^\alpha} f = 0, \quad (25)$$

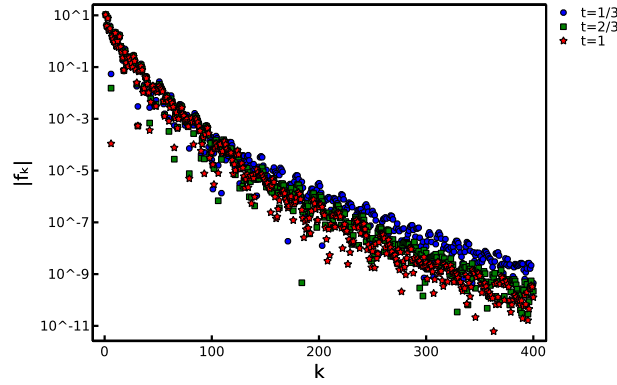


Fig. 10. Semilogarithmic plot of absolute values of the k -th coefficients of the weighted Prorior polynomial approximations corresponding to the numerical solutions shown in Fig. 9.

posed on a two-dimensional disk with zero Dirichlet conditions and given initial conditions $f(0, \mathbf{x})$ and $\left(\frac{\partial}{\partial t} f\right)(0, \mathbf{x})$. Note that this stand-in equation also describes power law attenuation in viscous media – it is in fact a modified version of Szabo’s wave equation, see [83,84] which has also been suggested for use in applications, cf. [85]. To our knowledge no non-trivial explicit solutions for this equation are known.

As the problem is posed on a disk domain, we use *generalized* Zernike polynomials $\mathbf{Z}^{(b)}(r, \theta)$. Zernike polynomials have a long-standing history [86,87] not just in mathematics research but also in natural science and engineering applications, especially in the field of optics [88–92], and as result a handful of different and conflicting conventions exist for them. The convention we follow for the generalized Zernike polynomials is that $\mathbf{Z}^{(b)}(r, \theta)$ are orthonormal polynomials on the unit disk with respect to the weight function $w(r) = (1 - r^2)^b$, whose radial components are given by orthonormal Jacobi polynomials $\tilde{\mathbf{P}}^{(\alpha, \beta)}(2r^2 - 1)$, i.e.:

$$\mathbf{Z}_{\ell, m}^{(b)}(r, \theta) = 2^{\frac{|m|+a+b+2}{2}} r^{|m|} \tilde{\mathbf{P}}_{\frac{\ell-|m|}{2}}^{(b, |m|)}(2r^2 - 1) \sqrt{\frac{2-\delta_{m,0}}{2\pi}} \begin{cases} \cos(m\theta) & \text{for } m \geq 0, \\ \sin(|m|\theta) & \text{for } m < 0. \end{cases}$$

The Zernike polynomials are traditionally given in terms of polar coordinates but they are polynomials in x and y and generally not polynomials in r and θ . For a discussion of sparse spectral methods using Zernike polynomials, see e.g. [34]. For open source numerical implementations we refer to FastTransforms [93,94] as well as MultivariateOrthogonalPolynomials.jl [71].

As in the triangle example, zero Dirichlet boundary conditions can be enforced by choosing an appropriate weighted basis

$$\mathbf{W}^{(b)}(r, \theta) = (1 - r^2)^b \mathbf{Z}^{(b)}(r, \theta),$$

to resolve the solution function at each time step $f(n\Delta t, r, \theta)$. Using a simple backward finite difference approximation for the second derivative we then find the following scheme for Equation (25):

$$\begin{aligned} \mathbf{Z}^{(1)}(r, \theta) \left(\left(\frac{1}{c_0^2 \Delta t} + \tau \left(\sum_{j=1}^L A_j \frac{1 - e^{-s_j^2 \Delta t}}{s_j^2} \right) \right) C - (\Delta t) \Delta \right) f^n = \\ \mathbf{Z}^{(1)}(r, \theta) C \left(\left(\frac{2}{c_0^2 \Delta t} + \tau \sum_{j=1}^L A_j \frac{1 - e^{-s_j^2 \Delta t}}{s_j^2} \right) f^{n-1} - \frac{f^{n-2}}{c_0^2 \Delta t} - \tau \sum_{j=1}^L A_j e^{-s_j^2 \Delta t} \psi_j^{n-1} \right), \end{aligned} \quad (26)$$

with approximate solution given by

$$f(N\Delta t, r, \theta) = \mathbf{W}^{(1)}(r, \theta) f^N.$$

Higher order schemes may be obtained in the straightforward way. As in the triangle experiment, we plot a sparsity pattern plot for C in Fig. 11 but omit the spy plot for the Laplacian Δ on these bases since it is in fact simply diagonal when mapping from $\mathbf{W}^{(1)}(r, \theta)$ to $\mathbf{Z}^{(1)}(r, \theta)$.

Remark 4. The second derivative in t requires access to at least two previous points in time, one of which can be re-used in the Caputo part of the equation. As a result, while the minimum total memory requirement discussed in Section 5 still holds, the memory *increase* from adding a Caputo term to a wave equation is K less than said minimum since f^{n-1} is a vector of length K and need not be stored twice.

We plot a numerical solution to Equation (25) with parameters $\tau = 1$, $c = 100$, $\alpha = \frac{1}{2}$ and initial conditions

$$f(0, x, y) = 4y(1 - x^2 - y^2)^2, \quad (27)$$

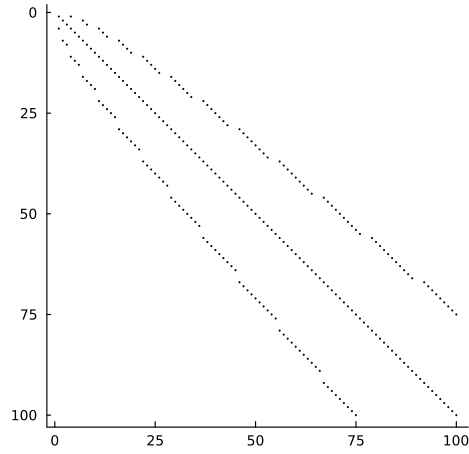


Fig. 11. Sparsity pattern (spy) plot for 100×100 block of the conversion operator C appearing in Equation (26). See e.g. [34] for how to compute its entries.

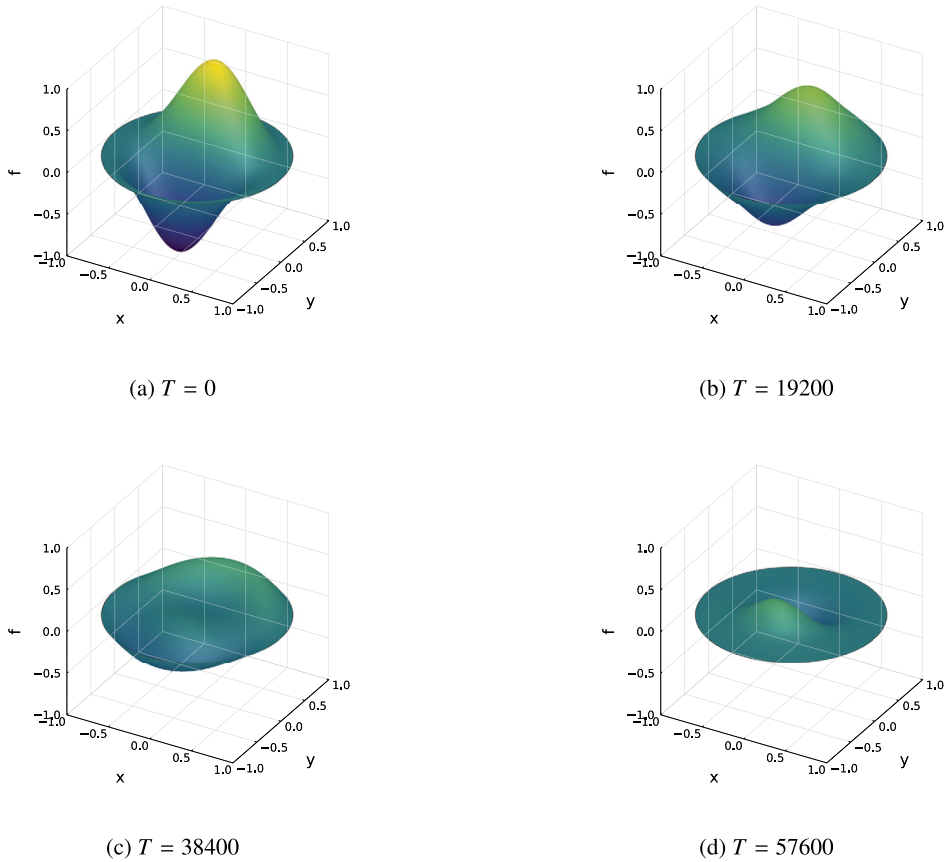


Fig. 12. (b), (c) and (d) show numerical solutions to the fractionally dampened wave equation in Equation (25) after indicated amount of time steps starting with the initial conditions in (27) pictured in (a). The numerical parameters for the pictured simulation were $(\Delta t, K, L) = (2^{-20}, 500, 50)$. Animations for deeper time as well as other initial conditions are available on FigShare, see [95].

$$\frac{\partial f}{\partial t}(0, x, y) = 0,$$

on the disk in Fig. 12. Four animations of solutions to this problem for various parameters and initial conditions, including the one pictured in Fig. 12, have been made available via FigShare, see [95].

In Fig. 13 we plot an example of an initial condition along with a circular array of 70 sensors and the data these sensors read out over time. State-of-the-art software such as the Matlab package k-wave, cf. [96–99], can use data from such sensor arrays to approximately reconstruct source images with natural applications in medical ultrasound as well as photo-acoustic imaging. This

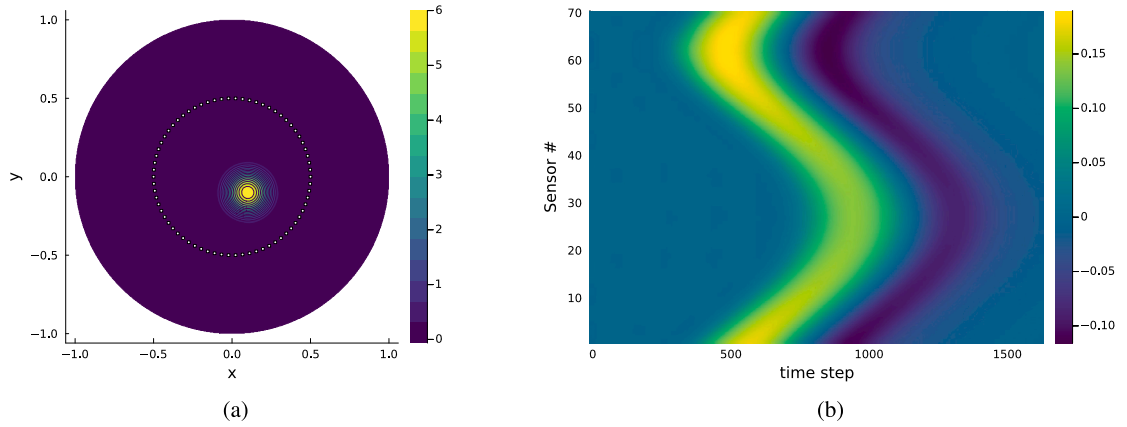


Fig. 13. (b) shows the read-out as a function of time of the 70 sensors placed around the initial state plotted in (a) – the circular sensor array with radius $r_s = \frac{1}{2}$ is marked by white squares in (a). The underlying equation is of the type in (25) with $\alpha = \frac{1}{10}$. Inputs of the form of (b) can be used by the Matlab package k-wave to approximately reconstruct the original source images.

section serves as a proof of concept that a recursive sparse spectral element method may be applicable to such problems in the future – we discuss additional research required to reach that point in the next section.

8. Discussion & future research

In this paper we have suggested a nonclassical sparse spectral approach to solving PDEs involving Caputo fractional derivatives. Among the method's core strengths are its recursive and history-free design, allowing computations that would otherwise be extremely challenging, and the reliance on sparse spectral methods in space resulting in banded operators which can be efficiently solved in each time step. The generic, almost 'plug-and-play', nature of the approach means that it can easily be adapted to any domains on which sparse spectral methods are viable, i.e. all domains on which suitable sets of orthogonal polynomials and derivative operators are available. We demonstrated this by presenting numerical solutions to a time-fractional heat equation on the triangle and a time-fractional modification of the wave equation on the disk. As seen in the numerical experiments, the method does not require excessively many quadrature points to achieve good accuracy, with time step size Δt becoming the primary error contributor once convergence in the number of quadrature points is achieved. This suggests that the approach outlined in this paper may in practice benefit significantly from replacing the various time stepping aspects of the method with higher order in time methods in a straightforward manner.

In future research we intend to adapt this method to work for the full equation of interest

$$\frac{1}{c_0^2} \frac{\partial^2}{\partial t^2} f - \Delta f - \tau \frac{\partial^\alpha}{\partial t^\alpha} \Delta f = 0,$$

which will require memory and computationally efficient extensions of this method for computing terms like $\frac{\partial^\alpha}{\partial t^\alpha} (\Delta f(t, x))$ as well as a discussion of how to join up different domains with matching boundary conditions – in this case combining a core ball cell with several layers of surrounding spherical shells of varying thickness. Applying such methods for the time reversed problem from finite sensor array readouts, as suggested with Fig. 13 and the Matlab package k-wave, is of significant interest in applications. This functionality would require the use of theoretical results on the time-reversal of the discussed equations as well as an implementation of absorbing boundary conditions to prevent noise coming back from the domain boundary. To sensibly use absorbing boundary conditions with the present method, it is natural to e.g. use a spectral elements approach and enforce rapid decay in the outer-most element. We intend to discuss image reconstruction using a spectral element variant of this method in detail in a future paper.

While the above points remain a challenge for now, the introduced method's minimal theoretical overhead when changing domains suggest it is a good candidate for such multi-domain problems. Advances in this direction should be compared with the current state-of-the-art methods for solving these equations, which involve replacing the temporal nonlocality with a spatial nonlocality as discussed in [3] (as implemented in k-wave).

CRediT authorship contribution statement

Timon S. Gutleb: Conceptualization, Software, Visualization, Writing – original draft, Writing – review & editing. **José A. Carrillo:** Funding acquisition, Project administration, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

TSG and JAC were supported by EPSRC grant number EP/T022132/1. JAC was supported by the Advanced Grant Nonlocal-CPD (Nonlocal PDEs for Complex Particle Dynamics: Phase Transitions, Patterns and Synchronization) of the European Research Council Executive Agency (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 883363). JAC was also partially supported by the EPSRC grant number EP/V051121/1. The authors wish to thank Ioannis P. A. Papadopoulos and Sheehan Olver for providing helpful feedback on an early draft of the paper and Bradley Treeby for helpful conversations about medical ultrasound, k-wave's functionality and fractionally dampened wave equations.

References

- [1] M. Caputo, Linear models of dissipation whose Q is almost frequency independent—II, *Geophys. J. Int.* 13 (5) (1967) 529–539.
- [2] B.E. Treeby, B.T. Cox, Modeling power law absorption and dispersion for acoustic propagation using the fractional Laplacian, *J. Acoust. Soc. Am.* 127 (5) (2010) 2741–2748.
- [3] B.E. Treeby, B.T. Cox, Modeling power law absorption and dispersion in viscoelastic solids using a split-field and the fractional Laplacian, *J. Acoust. Soc. Am.* 136 (4) (2014) 1499–1510.
- [4] R. Magin, Fractional calculus in bioengineering, part 1, *Crit. Rev. Biomed. Eng.* 32 (1) (2004).
- [5] A. Goulart, M. Lazo, J. Suarez, D. Moreira, Fractional derivative models for atmospheric dispersion of pollutants, *Physica A* 477 (2017) 9–19.
- [6] Ü. Dikmen, Modeling of seismic wave attenuation in soil structures using fractional derivative scheme, *J. Balkan Geophys. Soc.* 8 (4) (2005) 175–188.
- [7] L.S. Taylor, A.L. Lerner, D.J. Rubens, K.J. Parker, A Kelvin-Voigt fractional derivative model for viscoelastic characterization of liver tissue, in: *ASME International Mechanical Engineering Congress and Exposition*, vol. 36509, 2002, pp. 447–448.
- [8] F. Cesarone, M. Caputo, C. Cametti, Memory formalism in the passive diffusion across a biological membrane, *J. Membr. Sci.* 250 (2004) 79–84.
- [9] A. Bounaim, W. Chen, Computations for a breast ultrasonic imaging technique and finite element approach for a fractional derivative modeling the breast tissue acoustic attenuation, *Int. J. Tomogr. Stat.* 10 (8) (2008) 31–43.
- [10] B. Jin, Z. Zhou, *Numerical Treatment and Analysis of Time-Fractional Evolution Equations*, vol. 214, Springer Nature, 2023.
- [11] M. Harker, Fractional differential equations: numerical methods for applications, *Stud. Syst. Decis. Control* 258 (2020).
- [12] D. Baleanu, K. Diethelm, E. Scalas, J.J. Trujillo, *Fractional Calculus: Models and Numerical Methods*, vol. 3, World Scientific, 2012.
- [13] I. Podlubny, *Fractional Differential Equations*, Elsevier, 1998.
- [14] N.J. Ford, A.C. Simpson, The numerical solution of fractional differential equations: speed versus accuracy, *Numer. Algorithms* 26 (2001) 333–346.
- [15] L. Yuan, O.P. Agrawal, A numerical scheme for dynamic systems containing fractional derivatives, *J. Vib. Acoust.* 124 (2) (2002) 321–324.
- [16] N. Hale, S. Olver, A fast and spectrally convergent algorithm for rational-order fractional integral and differential equations, *SIAM J. Sci. Comput.* 40 (4) (2018) A2456–A2491.
- [17] T. Pu, M. Fasoldini, The numerical solution of fractional integral equations via orthogonal polynomials in fractional powers, *Adv. Comput. Math.* 49 (1) (2023) 7.
- [18] M.J. Colbrook, L.J. Ayton, A contour method for time-fractional pdes and an application to fractional viscoelastic beam equations, *J. Comput. Phys.* 454 (2022) 110995.
- [19] W. Chen, S. Holm, Fractional Laplacian time-space models for linear and nonlinear lossy media exhibiting arbitrary frequency power-law dependency, *J. Acoust. Soc. Am.* 115 (4) (2004) 1424–1430.
- [20] M. Kwaśnicki, Ten equivalent definitions of the fractional laplace operator, *Fract. Calc. Appl. Anal.* 20 (1) (2017) 7–51.
- [21] E. Di Nezza, G. Palatucci, E. Valdinoci, Hitchhiker's guide to the fractional sobolev spaces, *Bull. Sci. Math.* 136 (5) (2012) 521–573.
- [22] S. Olver, A. Townsend, A fast and well-conditioned spectral method, *SIAM Rev.* 55 (3) (2013) 462–489.
- [23] S. Olver, R.M. Slevinsky, A. Townsend, Fast algorithms using orthogonal polynomials, *Acta Numer.* 29 (2020) 573–699.
- [24] F. Olver, A. Daalhuis, D. Lozier, B. Schneider, R. Boisvert, C. Clark, B. Miller (Eds.), *NIST Digital Library of Mathematical Functions*, 2022, <http://dlmf.nist.gov>.
- [25] T.S. Gutleb, A fast sparse spectral method for nonlinear integro-differential Volterra equations with general kernels, *Adv. Comput. Math.* 47 (3) (2021) 42.
- [26] S. Olver, A. Townsend, G. Vasil, A sparse spectral method on triangles, *SIAM J. Sci. Comput.* 41 (6) (2019) A3728–A3756, <https://doi.org/10.1137/19M1245888>.
- [27] R.M. Slevinsky, S. Olver, A fast and well-conditioned spectral method for singular integral equations, *J. Comput. Phys.* 332 (2017) 290–315.
- [28] T.S. Gutleb, Sparse spectral methods for integral equations and equilibrium measures, Ph.D. thesis, Imperial College London, 2022.
- [29] T.S. Gutleb, S. Olver, A sparse spectral method for Volterra integral equations using orthogonal polynomials on the triangle, *SIAM J. Numer. Anal.* 58 (3) (2020) 1993–2018.
- [30] T.S. Gutleb, J.A. Carrillo, S. Olver, Computation of power law equilibrium measures on balls of arbitrary dimension, *Constr. Approx.* (2022) 1–46.
- [31] I. Papadopoulos, S. Olver, A sparse spectral method for fractional differential equations in one-spacial dimension, *arXiv preprint arXiv:2210.08247*, 2022.
- [32] N. Hale, An ultraspherical spectral method for linear Fredholm and Volterra integro-differential equations of convolution type, *IMA J. Numer. Anal.* 39 (4) (2019) 1727–1746.
- [33] T.S. Gutleb, J. Carrillo, S. Olver, Computing equilibrium measures with power law kernels, *Math. Comput.* 91 (337) (2022) 2247–2281.
- [34] G. Vasil, K.J. Burns, D. Lecoanet, S. Olver, B.P. Brown, J.S. Oishi, Tensor calculus in polar coordinates using Jacobi polynomials, *J. Comput. Phys.* 325 (2016) 53–73, <https://doi.org/10.1016/j.jcp.2016.08.013>.
- [35] S. Olver, A. Townsend, G.M. Vasil, Recurrence relations for a family of orthogonal polynomials on a triangle, in: *Spectral and High Order Methods for Partial Differential Equations (ICOSAHOM 2018)*, Springer, 2020, pp. 79–92.
- [36] R. Aktas, I. Area, E. Güldoğan, A new family of orthogonal polynomials in three variables, *J. Inequal. Appl.* 1 (2020) 170.
- [37] B. Snowball, S. Olver, Sparse spectral and -finite element methods for partial differential equations on disk slices and trapeziums, *Stud. Appl. Math.* 145 (1) (2020) 3–35, <https://doi.org/10.1111/sapm.12303>.

- [38] B. Snowball, S. Olver, Sparse spectral methods for partial differential equations on spherical caps, *Trans. Math. Appl.* 5 (1) (2021).
- [39] S. Olver, Y. Xu, Orthogonal structure on a wedge and on the boundary of a square, *Found. Comput. Math.* 19 (3) (2019) 561–589, <https://doi.org/10.1007/s10208-018-9393-0>.
- [40] S. Olver, Y. Xu, Orthogonal polynomials in and on a quadratic surface of revolution, *Math. Comput.* 89 (326) (2020) 2847–2865, <https://doi.org/10.1090/mcom/3544>.
- [41] M. Fasondini, S. Olver, Y. Xu, Orthogonal polynomials on planar cubic curves, *Found. Comput. Math.* (2021) 1–31.
- [42] S. Olver, Y. Xu, Orthogonal structure on a quadratic curve, *IMA J. Numer. Anal.* 41 (1) (2021) 206–246, <https://doi.org/10.1093/imanum/draa001>.
- [43] C.F. Dunkl, Y. Xu, *Orthogonal Polynomials of Several Variables*, second edition, *Encyclopedia of Mathematics and Its Applications*, vol. 155, Cambridge University Press, Cambridge, 2014.
- [44] T.S. Gutleb, S. Olver, R.M. Slevinsky, Polynomial and rational measure modifications of orthogonal polynomials via infinite-dimensional banded matrix factorizations, <https://doi.org/10.48550/arXiv.2302.08448>, Feb. 2023.
- [45] C. Trinks, P. Ruge, Treatment of dynamic systems with fractional derivatives without evaluating memory-integrals, *Comput. Mech.* 29 (6) (2002) 471–476.
- [46] K. Diethelm, An investigation of some nonclassical methods for the numerical approximation of Caputo-type fractional derivatives, *Numer. Algorithms* 47 (4) (2008) 361–390.
- [47] A. Chatterjee, Statistical origins of fractional derivatives in viscoelasticity, *J. Sound Vib.* 284 (3–5) (2005) 1239–1245.
- [48] A. Schmidt, L. Gaul, On a critique of a numerical scheme for the calculation of fractionally damped dynamical systems, *Mech. Res. Commun.* 33 (1) (2006) 99–107.
- [49] J.-F. Lu, A. Hanyga, Wave field simulation for heterogeneous porous media with singular memory drag force, *J. Comput. Phys.* 208 (2) (2005) 651–674.
- [50] A. Hanyga, J.-F. Lu, Wave field simulation for heterogeneous transversely isotropic porous media with the JKD dynamic permeability, *Comput. Mech.* 36 (3) (2005) 196–208.
- [51] J.-F. Lu, A. Hanyga, Numerical modelling method for wave propagation in a linear viscoelastic medium with singular memory, *Geophys. J. Int.* 159 (2) (2004) 688–702.
- [52] Q. Liu, J. Liu, Y. Chen, On the theoretical basis of memory-free approaches for fractional differential equations, *Eng. Comput.* (2019).
- [53] W. Gautschi, Quadrature formulae on half-infinite intervals, *BIT Numer. Math.* 31 (3) (1991) 437–446.
- [54] A. Bultheel, A. Cuyt, W. Van Assche, M. Van Barel, B. Verdonk, Generalizations of orthogonal polynomials, *J. Comput. Appl. Math.* 179 (1–2) (2005) 57–95.
- [55] C. Birk, C. Song, An improved non-classical method for the solution of fractional differential equations, *Comput. Mech.* 46 (5) (2010) 721–734.
- [56] S.G. Samko, A.A. Kilbas, O.I. Marichev, *Fractional Integrals and Derivatives: Theory and Applications*, Gordon and Breach Science Publishers, Switzerland; Philadelphia, Pa., USA, 1993.
- [57] C.-C. Tseng, S.-C. Pei, S.-C. Hsia, Computation of fractional derivatives using Fourier transform and digital FIR differentiator, *Signal Process.* 80 (1) (2000) 151–159.
- [58] A. Townsend, S. Olver, The automatic solution of partial differential equations using a global spectral method, *J. Comput. Phys.* 299 (2015) 106–123.
- [59] F. Olver, D. Sookne, Note on backward recurrence algorithms, *Math. Comput.* 26 (120) (1972) 941–947.
- [60] J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, *Texts in Applied Mathematics*, vol. 12, Springer, New York, New York, 1996.
- [61] L.N. Trefethen, *Approximation Theory and Approximation Practice*, extended edition, SIAM, 2019.
- [62] J.P. Boyd, *Chebyshev and Fourier Spectral Methods*, second edition, Courier Corporation, 2001.
- [63] M. Hamzehnejad, M.M. Hosseini, A. Salemi, An improved upper bound for ultraspherical coefficients, *J. Math. Model.* 10 (3) (2022) 421–431.
- [64] H. Wang, A new and sharper bound for Legendre expansion of differentiable functions, *Appl. Math. Lett.* 85 (2018) 95–102.
- [65] H. Wang, How much faster does the best polynomial approximation converge than Legendre projection? *Numer. Math.* 147 (2) (2021) 481–503.
- [66] X. Zhao, L.-L. Wang, Z. Xie, Sharp error bounds for Jacobi expansions and Gegenbauer–Gauss quadrature of analytic functions, *SIAM J. Numer. Anal.* 51 (3) (2013) 1443–1469.
- [67] S. Xiang, On error bounds for orthogonal polynomial expansions and Gauss-type quadrature, *SIAM J. Numer. Anal.* 50 (3) (2012) 1240–1263.
- [68] H. Wang, S. Xiang, On the convergence rates of Legendre approximation, *Math. Comput.* 81 (278) (2012) 861–877.
- [69] H. Wang, On the optimal estimates and comparison of Gegenbauer expansion coefficients, *SIAM J. Numer. Anal.* 54 (3) (2016) 1557–1581.
- [70] J. Bezanon, A. Edelman, S. Karpinski, V.B. Shah Julia, A fresh approach to numerical computing, *SIAM Rev.* 59 (1) (2017) 65–98, <https://doi.org/10.1137/141000671>.
- [71] *MultivariateOrthogonalPolynomials.jl*, <https://github.com/JuliaApproximation/MultivariateOrthogonalPolynomials.jl>, 2022. (Accessed 20 December 2022).
- [72] *FastGaussQuadratures.jl*, Julia Package, v0.5.0, <https://github.com/JuliaApproximation/FastGaussQuadratures.jl>, 2022. (Accessed 20 December 2022).
- [73] *ApproxFun.jl*, Julia Package, v0.13.0, <https://github.com/JuliaApproximation/ApproxFun.jl>, 2022. (Accessed 20 December 2022).
- [74] I. Bogaert, Iteration-free computation of Gauss–Legendre quadrature nodes and weights, *SIAM J. Sci. Comput.* 36 (3) (2014) A1008–A1026, <https://doi.org/10.1137/140954969>.
- [75] A. Glaser, X. Liu, V. Rokhlin, A fast algorithm for the calculation of the roots of special functions, *SIAM J. Sci. Comput.* 29 (4) (2007) 1420–1438, <https://doi.org/10.1137/06067016X>.
- [76] A. Townsend, T. Trogdon, S. Olver, Fast computation of Gauss quadrature nodes and weights on the whole real line, *IMA J. Numer. Anal.* 36 (1) (2016) 337–358.
- [77] A. Townsend, The race for high order Gauss–Legendre quadrature, *SIAM News* 48 (2015) 1–3.
- [78] R. Herrmann, *Fractional Calculus: An Introduction for Physicists*, 2nd edition, World Scientific, Hackensack, NJ, 2014.
- [79] G.M. Mittag-Leffler, Sur la nouvelle fonction $Ea(x)$, *C.R. Acad. Sci. Paris* 137 (2) (1903) 554–558.
- [80] A. Wiman, Über den Fundamentalsatz in der Theorie der Funktionen $Ea(x)$, *Acta Math.* 29 (1905) 191–201 (in German), spelling error in title present in original.
- [81] K. Diethelm, *The Analysis of Fractional Differential Equations*, *Lecture Notes in Mathematics*, vol. 2004, Springer, Berlin Heidelberg, Berlin, Heidelberg, 2010.
- [82] S. Näss, S. Chintaluri, T.V. Ness, A.M. Dale, G.T. Einevoll, D.K. Wójcik, Corrected four-sphere head model for EEG signals, *Front. Human Neurosci.* 11 (2017) 490.
- [83] T.L. Szabo, Time domain wave equations for lossy media obeying a frequency power law, *J. Acoust. Soc. Am.* 96 (1) (1994) 491–500.
- [84] W. Chen, S. Holm, Modified Szabo's wave equation models for lossy media obeying frequency power law, *J. Acoust. Soc. Am.* 114 (5) (2003) 2570–2574.
- [85] S. Holm, V. Pandey, Wave propagation in marine sediments expressed by fractional wave and diffusion equations, in: 2016 IEEE/OES China Ocean Acoustics (COA), IEEE, 2016, pp. 1–5.
- [86] F. Zernike, H. Brinkman, Hypersphärische Funktionen und die in sphärischen Bereichen orthogonalen Polynome, in: *Proc. Akad. Wet. Amsterdam*, vol. 38, 1935, pp. 161–170.
- [87] F. Zernike, Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode, *Physica* 1 (7) (1934) 689–704, [https://doi.org/10.1016/S0031-8914\(34\)80259-5](https://doi.org/10.1016/S0031-8914(34)80259-5).
- [88] L.N. Thibos, R.A. Applegate, J.T. Schwiegerling, R. Webb, VSIA standards taskforce members, standards for reporting the optical aberrations of eyes, in: *Vision Science and Its Applications*, OSA, Santa Fe, New Mexico, 2000.
- [89] N.A. Roddier, Atmospheric wavefront simulation using Zernike polynomials, *Opt. Eng.* 29 (10) (1990) 1174–1180, <https://doi.org/10.1117/12.55712>.
- [90] V.N. Mahajan, Zernike circle polynomials and optical aberrations of systems with circular pupils, *Appl. Opt.* 33 (34) (1994) 8121, <https://doi.org/10.1364/AO.33.008121>.
- [91] K.M. Rocha, L. Vabre, F. Harms, N. Chateau, R.R. Krueger, Effects of Zernike wavefront aberrations on visual acuity measured using electromagnetic adaptive optics technology, *J. Refract. Surg.* 23 (9) (2007) 953–959, <https://doi.org/10.3928/1081-597X-20071101-17>.

- [92] R.J. Noll, Zernike polynomials and atmospheric turbulence, *J. Opt. Soc. Am.* 66 (3) (1976) 207–211, <https://doi.org/10.1364/JOSA.66.000207>.
- [93] FastTransforms, C library, v0.5.1, <https://github.com/MikaelSlevinsky/FastTransforms>, 2022. (Accessed 20 December 2022).
- [94] FastTransforms.jl, julia Package, v0.13.2, <https://github.com/JuliaApproximation/FastTransforms.jl>, 2022. (Accessed 20 December 2022).
- [95] T.S. Gutleb, J.A. Carrillo, Wave equation on the disk dampened by Caputo time-fractional term, *FigShare* (2023), <https://doi.org/10.6084/m9.figshare.22630744.v1>.
- [96] B.E. Treeby, B.T. Cox, k-wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields, *J. Biomed. Opt.* 15 (2) (2010) 021314.
- [97] B.E. Treeby, J. Jaros, A.P. Rendell, B. Cox, Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using a k-space pseudospectral method, *J. Acoust. Soc. Am.* 131 (6) (2012) 4324–4336.
- [98] B.E. Treeby, J. Budisky, E.S. Wise, J. Jaros, B. Cox, Rapid calculation of acoustic fields from arbitrary continuous-wave sources, *J. Acoust. Soc. Am.* 143 (1) (2018) 529–537.
- [99] B.E. Treeby, J. Jaros, D. Rohrbach, B. Cox, Modelling elastic wave propagation using the k-wave MATLAB toolbox, in: 2014 IEEE International Ultrasonics Symposium, IEEE, 2014, pp. 146–149.