Noname manuscript No.
(will be inserted by the editor)

Numerical Caputo differentiation by radial basis functions

Ming Li · Yujiao Wang · Leevan Ling

Received: date / Accepted: date

Abstract Previously, based on the method of (radial powers) radial basis functions, we proposed a procedure for approximating derivative values from one-dimensional scattered noisy data. In this work, we show that the same approach also allows us to approximate the values of (Caputo) fractional derivatives (for orders between 0 and 1). With either an *a priori* or *a posteriori* strategy of choosing the regularization parameter, our convergence analysis shows that the approximated fractional derivative values converge at the same rate as in the case of integer order 1.

Keywords Fractional derivatives · inverse problem · convergence analysis · noisy data · regularization.

Mathematics Subject Classification (2000) 62M40 · 65D25 · 65F22

1 Introduction

Fractional derivatives are generally used in both direct [7] and inverse problems [15,21,28]; for example, applications of fractional derivatives can be found in physics [20], finance [24], and hydrology [4]. Therefore, we recently can see a lot of studies on numerical methods [3,13,14,22] and numerical analysis [10,11,17,29] for fractional partial differential equations. As soon as one shifts from integer to fractional order models which involve approximating derivative values, the need for approximating integer order derivatives will turn into the fractional cases. Hence, finding the value of fractional derivatives from scattered noisy data is a topic that is worth investigating. The problems of

Ming Li · Yujiao Wang
Department of Mathematics, Taiyuan University of Technology, Taiyuan, China.

Leevan Ling
Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong
E-mail: lling@hkbu.edu.hk

approximating (integer [2,8] and fractional) derivatives are well-known to be ill-posed in the sense that small errors in the data might induce large errors in the computed derivative. In this work, we are going to generalize our previously proposed method to approximate derivative values of fractional orders.

Let $f : [a, b] \to \mathbb{R}$ be the unknown function of interest. A set of n data sites $\Xi_n = \{a = x_1 < x_2 < ... < x_n = b\} \subset \mathbb{R}$ is given inside the interval $[a, b]$. Let $h_{\max}$ and $h_{\min}$ denote respectively the maximum and minimum separating distances defined by

$$h := h_{\max} = \max_{2\leq j\leq n} h_j, \quad h_{\min} = \min_{2\leq i\leq n} h_i, \quad \text{where } h_j = x_j - x_{j-1}, \ 2 \leq j \leq n. \tag{1}$$

For the sake of convergence, we assume that the sets of data sites $\{\Xi_n\}$ satisfy a quasi-uniformity condition $h_{\max}/h_{\min} \leq \gamma$ for some $\gamma > 0$ independent of n. Now, the ill-posed problem we consider is as follows:

> Given a set of noisy data $\{(x_j, y_j^\delta)\}_1^n$, where $\frac{1}{n}\|y_j^\delta - f(x_j)\|_{\ell_2} \leq \delta^2$, we aim to seek a function to approximate the Caputo fractional derivative of order $0 < \alpha < 1$, ${}_0^cD^\alpha f$, of the unknown function f^*.

After the details of our numerical procedure are presented in the coming section, our theoretical goal is to prove that the resulting numerical approximation will in fact converge to ${}_0^cD^\alpha f$ as the data gets dense and the noise reduces (i.e. as $h \searrow 0$ and $\delta \searrow 0$). This work will then be concluded by some numerical verifications for the proven convergence rate.

2 Numerical fractional differentiation

To begin, the Caputo fractional derivative of order $0 < \alpha < 1$ is defined as

$${}_0^cD^\alpha f(x) = \frac{1}{\Gamma(1-\alpha)} \int_0^x \frac{d}{ds} f(s) \frac{ds}{(x-s)^\alpha}; \tag{2}$$

see monograph by Podlubny [19]. As $\alpha \nearrow 1$, the Caputo fractional derivative in (2) coincides with the (integer) first-order derivative.

The objective here is to provide a numerical procedure that allows us to find a numerical approximation of the Caputo fractional derivative of an unknown function from its noisy function values. The method of radial basis functions (RBF) is a renowned approach for finding interpolants from scattered data and was used for numerical differentiation [12]. Throughout the paper, we focus on the radial power RBF $\phi(x) = \|x\|^{2\beta-1}$, $\beta \in \mathbb{N}$, which is known (see [6]) to be strictly conditionally positive definite[1] of order β.

Suppose noisy data $\{(x_j, y_j^\delta)\}_1^n$ is available for some noise level δ such that

$$\frac{1}{n}\|y_j^\delta - f(x_j)\|_{\ell_2} \leq \delta^2. \tag{3}$$

[1] More precisely, radial power RBF is $(-1)^\beta\|x\|^{2\beta-1}$. As the interpolation matrix is not required in this work, we drop the term $(-1)^k$ for the sake of simplicity.

Having an interpolant going through all the noisy data is almost meaningless for numerical differentiation; we therefore seek a regularized fit in the Tihkonov sense [23]:

$$y_{\beta,\sigma} := \arg\min_{g\in X_\beta} J(g) \text{ where } J(g) := \frac{1}{n}\sum_{j=1}^{n}\left(g(x_j) - y_j^\delta\right)^2 + \sigma\, |g|_\beta, \quad (4)$$

where β is specified by the employed radial power RBF. The trial space is defined as $X_\beta = \{H^\beta \cap C^{\beta-1}\}[a,b]$, $-\infty < a \le 0 < b < \infty$, and the seminorm is defined as

$$|g|_\beta = \left(\int_a^b \left|\frac{d^\beta}{dx^\beta} g(x)\right|^2 dx\right)^{\frac{1}{2}}, \quad (5)$$

For any regularization parameter σ, we seek the optimizer of functional in (4) in the form of

$$y_{\beta,\sigma}(x) = \sum_{j=1}^{n} c_j\, \phi(x - x_j) + \sum_{j=1}^{\beta} d_j\, x^{j-1}. \quad (6)$$

In our previous work, see [26], it is shown that the following two conditions

$$\sum_{j=1}^{n} c_j\, x_j^i = 0, \quad \text{and} \quad y_{\beta,\sigma}(x_i) + 2(2\beta-1)!(-1)^\beta \sigma n c_j = y_j^\delta, \quad (7)$$

for all $i = 1, \ldots, k-1$, which allows us to analytically express all required derivatives in the proof, are sufficient to ensure

$$J(g) - J(y_{\beta,\sigma}) = \frac{1}{n}\sum_{i=1}^{n}[g(x_i) - y_{\beta,\sigma}(x_i)]^2 + \sigma\|g^{(\beta)} - y_{\beta,\sigma}^{(\beta)}\|^2_{L^2(R)} \ge 0,$$

for all $g \in X_\beta$. It is straightforward to show that this minimizer is unique. This provides a way for finding the minimizer of the functional J in (4). The interested reader is referred to the original article for further information.

With the analytical work done, solving (4) now turns out to be a linear problem. Conditions in (7) allow us to find the *unique* minimizer to functional in (4), i.e. the coefficients c_j and d_j in (6), can *always* be obtained via solving the following $(n+\beta)\times(n+\beta)$ matrix system:

$$\begin{bmatrix} \Phi_n + 2\sigma\, n\, (-1)^\beta (2\beta-1)!\, I_n & P \\ P^T & 0 \end{bmatrix} \begin{pmatrix} \vec{c} \\ \vec{d} \end{pmatrix} = \begin{pmatrix} \vec{y^\delta} \\ \vec{0} \end{pmatrix}, \quad (8)$$

where Φ_n is the interpolation matrix of our RBF evaluated at points $\{x_j\}_1^n$, i.e., $\left[\Phi_n\right]_{j,k} = \phi(x_j,\, x_k)$ $(1 \le j,k \le n)$, and the matrix P arises from the appended polynomial whose elements are $\left[P\right]_{j,k} = (x_j)^{k-1}$ $(1 \le j \le n,\ 1 \le k \le \beta)$.

Note that the matrix in (8) becomes the standard interpolation matrix if we set the regularization parameter $\sigma = 0$. The regularization effect can also be obtained by using RBF with higher smoothness if we assume the unknown function is as smooth. The linkage between regularization and smoothness is specified in the following theorem proven previously in [26].

Theorem 1 *Let $f \in X_\beta$ be the exact function and $y_{\beta,\sigma}$ obtained by* (6)–(8) *either with $\sigma = \delta^2$ (a prior) or the Morozov's discrepancy principle*[2] *(a posteriori), then we have a convergence estimate*

$$\|y_{\beta,\sigma} - f\|_{L^2(a,b)} \le Ph^{\beta-\frac{1}{2}} + Q\delta, \tag{9}$$

where the constant P and Q depend on which σ is used and also on β, γ, and the β-seminorm $|f|_\beta$.

Obtaining the regularized fit $y_{\beta,\sigma}$ from noisy data (x_j, y_j^δ) is the most difficult part (besides convergence theories, see [26]) for finding the integer order numerical derivatives of f; one can simply differentiate each basis in (6) to obtain approximated derivatives. Here we quote another theorem that will become handy later in our analysis.

Theorem 2 *Assume all in Theorem 1 holds, then for $j = 1, \dots, \beta - 1$, we have*

$$\left\| \frac{d^j}{dx^j}\big(y_{\beta,\sigma} - f\big)(x) \right\|_{L^2(a,b)} \le Ch^{\beta-j-\frac{1}{2}+\frac{j}{2\beta}} + D\delta^{\frac{\beta-j}{\beta}}, \tag{10}$$

and the errors in the β-derivative are bounded by

$$\left\| \frac{d^\beta}{dx^\beta}\big(y_{\beta,\sigma} - f\big)(x) \right\|_{L^2(a,b)} \le \left(2 + 4 \left\| \frac{d^\beta}{dx^\beta} f(x) \right\|^2_{L^2(a,b)} \right)^{\frac{1}{2}}. \tag{11}$$

For our problem, the fractional derivatives of each basis must be worked out. First, we compute the Caputo fractional derivatives of the appended polynomials from definition (2) and the binomial identity.

$$\begin{aligned}
{}_0^cD^\alpha x^{j-1} &= \frac{1}{\Gamma(1-\alpha)} \int_0^x \frac{d}{ds} s^{j-1} \frac{ds}{(x-s)^\alpha}, \quad (j \ge 2) \\
&= \frac{j-1}{\Gamma(1-\alpha)} \int_0^x \frac{s^{j-2}}{(x-s)^\alpha} ds \\
&= \frac{j-1}{\Gamma(1-\alpha)} \int_0^x \frac{\big(x-(x-s)\big)^{j-2}}{(x-s)^\alpha} ds \\
&= \frac{j-1}{\Gamma(1-\alpha)} \int_0^x \frac{\big(x-\zeta\big)^{j-2}}{\zeta^\alpha} d\zeta.
\end{aligned}$$

[2] By the Morozov's discrepancy principle, we select σ that satisfies $\frac{1}{n}\sum_{i=1}^{n}\big(y_{\beta,\sigma}(x_i) - y_i^\delta\big)^2 = \delta^2$.

For the radial power RBF, with a similar approach, a similar integral will be yielded. For $0 \le x \le x_j$, we have

$$
\begin{aligned}
{}_0^cD^\alpha \phi(x - x_j) &= \frac{1}{\Gamma(1-\alpha)} \int_0^x \frac{d}{ds}|s - x_j|^{2\beta-1} \frac{ds}{(x-s)^\alpha} \\
&= \frac{1-2\beta}{\Gamma(1-\alpha)} \int_0^x \frac{(s-x_j)^{2\beta-2}}{(x-s)^\alpha}\, ds, \\
&= \frac{1-2\beta}{\Gamma(1-\alpha)} \int_0^x \frac{\big((x-x_j)-(x-s)\big)^{2\beta-2}}{(x-s)^\alpha}\, ds \\
&= \frac{1-2\beta}{\Gamma(1-\alpha)} \int_0^x \frac{\big((x-x_j)-\zeta\big)^{2\beta-2}}{\zeta^\alpha}\, d\zeta,
\end{aligned}
$$

and, similarly, for $0 \le x_j \le x$,

$$
\begin{aligned}
{}_0^cD^\alpha \phi(x - x_j) &= \frac{1}{\Gamma(1-\alpha)} \int_0^x \frac{d}{ds}|s - x_j|^{2\beta-1} \frac{ds}{(x-s)^\alpha} \\
&= \frac{1-2\beta}{\Gamma(1-\alpha)} \left(\int_{x-x_j}^x + \int_{x-x_j}^0 \right) \frac{\big((x-x_j)-\zeta\big)^{2\beta-2}}{\zeta^\alpha}\, d\zeta.
\end{aligned}
$$

To complete the calculations, we simply need to apply the binomial formula to evaluate the integrals of the form

$$
\begin{aligned}
\int_\ell^u \frac{(\theta-\zeta)^\wp}{\zeta^\alpha}\, d\zeta &= \int_\ell^u \sum_{k=0}^{\wp} {}_\wp C_k\, (-1)^k\, \theta^{\wp-k}\, \zeta^{k-\alpha} d\zeta \\
&= \sum_{k=0}^{\wp} \frac{{}_\wp C_k\, (-1)^k\, \theta^{\wp-k}}{k-\alpha+1} \big(u^{k-\alpha+1} - \ell^{k-\alpha+1}\big),
\end{aligned}
$$

with the appropriate u, ℓ, θ, and $\wp$. For completeness, we will give the long expression for evaluating approximated Caputo fractional derivatives:

$$
\begin{aligned}
{}_0^cD^\alpha y_{\beta,\sigma}(x) = \sum_{j=1}^{n} \frac{(1-2\beta)c_j}{\Gamma(1-\alpha)} \sum_{k=0}^{2\beta-2} \frac{{}_{2\beta-2}C_k\, (-1)^k}{k-\alpha+1} (x-x_j)^{2\beta-2-k} \Big(x^{k-\alpha+1} \\
- 2\max(x-x_j, 0)^{k-\alpha+1} \Big) + \sum_{j=2}^{\beta} \frac{(j-1)d_j}{\Gamma(1-\alpha)} x^{j-1-\alpha} \sum_{k=0}^{j-2} {}_{j-2}C_k\, \frac{(-1)^k}{k-\alpha+1}.
\end{aligned}
\tag{12}
$$

Recall that the coefficients c_j and d_j were obtained by solving (6)–(8). Although the regularization parameter σ does not appear on the right-hand side of (12), its influence is implicit in the values of the coefficients. Numerically, the approximate Caputo derivative (12) can be evaluated with any σ. In the next section, we shall prove that the two choices, used in Theorem 1, will guarantee convergence ${}_0^cD^\alpha y_{\beta,\sigma} \to {}_0^cD^\alpha f$.

3 Convergence analysis

This section contains a few lengthy and technical proofs for our main theoretical result given in the following theorem.

Theorem 3 *For arbitrary $f \in X_\beta$, let $y_{\beta,\sigma}$ be the approximate fit given by (6)–(8). Suppose the regularization parameter is chosen either by $\sigma = \delta^2$ or by the Morozov's discrepancy principle, then for sufficiently small noisy level δ and data site spacing h, the following convergence estimate holds:*

$$\| {}^c_0D^\alpha y_{\beta,\sigma} - {}^c_0D^\alpha f\|_{L^2(a,b)} \le Ch^{\beta-\frac{3}{2}+\frac{1}{2\beta}} + D\delta^{\frac{\beta-1}{\beta}}, \tag{13}$$

where the constant C and D depend on β, $b-a$, $|y|_\beta$, and the selection of σ.

Note that inequality (13) has the same form as that for the first derivative in [26]. If one can be sure about the smoothness of the unknown function f, as both exponents of the h and δ terms are monotone increasing for $\beta \ge 1$, using the largest possible β will yield the highest convergence rate in theory. This claim, of course, does not take the problem of ill-conditioning into account; we will investigate that in the next section.

The proof of Theorem 3 relies heavily on the conditional stability of functions within the trial space H_β. The following theorem, which requires another lemma (Lemma 5) to be proved, makes this specific.

Theorem 4 *For any arbitrary $f \in X_\beta$, $0 < \alpha < 1$, and $0 < \varepsilon \le 1$, the following inequality holds:*

$$|f|^2_\alpha \le K\varepsilon|f|^2_\beta + K\varepsilon^{-1/(\beta-1)}|f|^2_0, \tag{14}$$

for some constant K that depends on α, β, $b-a$, and $|y|^2_\beta$.

Lemma 5 *Suppose $f \in C^2[a,b]$, $0 < \alpha < 1$, and $0 < \varepsilon_0 < \infty$. Then for any $0 < \varepsilon < \varepsilon_0$, the following inequality holds:*

$$|f|^2_\alpha \le K\varepsilon|f|^2_2 + K\varepsilon^{-1}|f|^2_0, \tag{15}$$

for some constant K that depends on α, ε_0, and $b-a$.

Proof. A technique for proofing interpolation inequalities [1] will be used here. For simplicity, we begin by assuming $a = 0$ and $b = 1$. Without loss of generality, we assume $\varepsilon_0 = 1$; if (15) is proved under this condition, due to $0 < \frac{\varepsilon}{\varepsilon_0} < 1$, a simple scaling, i.e. $K \leftarrow K\max(\varepsilon_0, 1)$, allows us to prove the original inequality.

Let $\xi \in (0, 1/3)$ and $\eta \in (2/3, 1)$. For any $f \in C^2[0,1]$, the mean value theorem ensures that there exists a $\lambda \in (\xi,\eta)$ such that

$$\begin{aligned}\left|\frac{d}{dx}f(\lambda)\right| &= \left|\frac{f(\xi)-f(\eta)}{\xi-\eta}\right| \\ &\le \frac{1}{|\xi-\eta|}\big(|f(\xi)|+|f(\eta)|\big) \le 3\big(|f(\xi)|+|f(\eta)|\big).\end{aligned}$$

Now, for any $x \in (0,1)$, we have

$$\begin{aligned}\left|\frac{d}{dx}f(x)\right| &= \left|\frac{d}{dx}f(\lambda) + \int_\lambda^x \frac{d^2}{dt^2}f(t)dt\right| \\ &\leq 3|f(\xi)| + 3|f(\eta)| + \int_0^1 \left|\frac{d^2}{dt^2}f(t)\right| dt.\end{aligned}$$

We now consider the magnitude of the Caputo fractional derivative of f at any $x \in (0,1)$:

$$\begin{aligned}|{}_0^cD^\alpha f(x)| &:= \frac{1}{\Gamma(1-\alpha)}\left|\int_0^x \frac{df(s)}{ds}(x-s)^{-\alpha}ds\right| \\ &\leq \frac{1}{\Gamma(1-\alpha)}\left(3|f(\xi)| + 3|f(\eta)| + \int_0^1 \left|\frac{d^2}{dt^2}f(t)\right| dt\right)\int_0^x (x-s)^{-\alpha}ds \\ &= \frac{1}{\Gamma(1-\alpha)}\left(3|f(\xi)| + 3|f(\eta)| + \int_0^1 \left|\frac{d^2}{dt^2}f(t)\right| dt\right)\frac{x^{1-\alpha}}{1-\alpha}.\end{aligned}$$

Note that the improper integral can be evaluated because $0 < \alpha < 1$. We integrate the former inequality with respect to the ξ and η variables from 0 to 1/3 and 2/3 to 1 respectively to obtain

$$\begin{aligned}\frac{1}{9}|{}_0^cD^\alpha f(x)| &\leq \frac{1}{\Gamma(1-\alpha)}\left(\int_0^{\frac{1}{3}} |f(\xi)|d\xi + \int_{\frac{2}{3}}^1 |f(\eta)|d\eta\right. \\ &\qquad \left. + \frac{1}{9}\int_0^1 \left|\frac{d^2}{dt^2}f(t)\right| dt\right)\frac{x^{1-\alpha}}{1-\alpha} \\ &\leq \frac{1}{\Gamma(1-\alpha)}\left(\int_0^1 |f(t)|dt + \frac{1}{9}\int_0^1 \left|\frac{d^2}{dt^2}f(t)\right| dt\right)\frac{x^{1-\alpha}}{1-\alpha}.\end{aligned}$$

By the Hölder's inequality, we have

$$\begin{aligned}\int_0^1 |{}_0^cD^\alpha f(x)|^2 dx &\leq \left(\frac{1}{\Gamma(1-\alpha)}\int_0^1 \frac{t^{1-\alpha}}{1-\alpha}dt\right)^2 \\ &\qquad \cdot \left(2\cdot 9^2 \int_0^1 |f(t)|^2 dt + 2\int_0^1 \left|\frac{d^2}{dt^2}f(t)\right|^2 dt\right),\end{aligned}$$

and hence, by taking $K := 2\cdot 9^2\left(\frac{1}{\Gamma(1-\alpha)}\int_0^1 \frac{t^{1-\alpha}}{1-\alpha}dt\right)^2$, we have an inequality for the α-seminorm, see (5), for the function f:

$$|f|_\alpha^2 := \int_0^1 |{}_0^cD^\alpha f(t)|^2 dt \leq \Big(K \underbrace{\int_0^1 \left|\frac{d^2}{dt^2}f(t)\right|^2 dt}_{=|f|_2^2} + K \underbrace{\int_0^1 |f(t)|^2 dt}_{=|f|_0^2}\Big). \tag{16}$$

The general case of arbitrary intervals (a,b) can be handled similarly with variables $\xi \in (a, a+(b-a)/3)$ and $\eta \in (a+2(b-a)/3, b)$. We can generalize (16) to

$$|f|_\alpha^2 \le K(b-a)^2 |f|_2^2 + K(b-a)^{-2}|f|_0^2. \tag{17}$$

Since $0 < \varepsilon \le 1$, there exits a positive integer n such that $2^{-1}\varepsilon^{1/2} \le n^{-1} \le \varepsilon^{1/2}$.

Now, we define a uniform partition $a_j = a + j\,(b-a)/n$, $j = 0, 1, \ldots, n$, for the interval (a,b). The fractional seminorm can be handled by applying (17) to each subinterval as follows:

$$\begin{aligned} |f|_\alpha^2 &:= \int_a^b |\,{}_0^c D^\alpha f(t)|^2\,dt = \sum_{j=1}^n \int_{a_{j-1}}^{a_j} |\,{}_0^c D^\alpha f(t)|^2\,dt \\ &\le K \sum_{j=1}^n \left\{ (a_j - a_{j-1})^2 \int_{a_{j-1}}^{a_j} \left| \frac{d^2}{dt^2} f(t) \right|^2 + (a_j - a_{j-1})^{-2} \int_{a_{j-1}}^{a_j} |f(t)|^2 dt \right\} \\ &\le K \sum_{j=1}^n \left\{ \left(\frac{b-a}{n} \right)^2 \int_{a_{j-1}}^{a_j} \left| \frac{d^2}{dt^2} f(t) \right|^2 dt + \left(\frac{n}{b-a} \right)^2 \int_{a_{j-1}}^{a_j} |f(t)|^2 dt \right\}. \end{aligned}$$

By recalling that $n^2 \le 4\varepsilon^{-1}$ and $n^{-2} \le \varepsilon$ and if we define $\tilde{K} = K \max\big((b-a)^2, 4(b-a)^{-2}\big)$, we obtain the following inequality (15) for $f \in C^2[a,b]$. □

Proof of Theorem 4. By assumptions, $f \in X_\beta$ allows us to apply Lemma 5. Hence, there are constants $0 < \delta < \delta_0 = 1$ and K_1 such that

$$|f|_\alpha^2 \le K_1 \delta |f|_2^2 + K_1 \delta^{-1} |f|_0^2, \tag{18}$$

Since the right-hand side of (18) contains only integer order seminorms, we are now in the right setting to apply the Sobolev imbedding theorems, see [1]. In particular, $|f|_j^2 \le K_j\,\delta |f|_{j+1}^2 + K_j\,\delta^{-j}|f|_0^2$, $j \ge 0$, allows us to bring the seminorm $|f|_2$ in (18) to $|f|_\beta$. First, with $j = 2$, putting $|f|_2^2 \le K_2\delta|f|_3^2 + K_2\delta^{-2}|f|_0^2$ into (18) yields

$$\begin{aligned} |f|_\alpha^2 &\le K_1\delta\big(K_2\delta|f|_3^2 + K_2\delta^{-2}|f|_0^2\big) + K_1\delta^{-1}|f|_0^2 \\ &= K_1K_2\delta^2|f|_3^2 + (K_1K_2 + K_1)\delta^{-1}|f|_0^2. \end{aligned}$$

Then, with $|f|_3^2 \le K_3\delta|f|_4^2 + K_3\delta^{-3}|f|_0^2$, we have

$$|f|_\alpha^2 = K_1K_2K_3\delta^3|f|_3^2 + (K_1K_2K_3 + K_1K_2 + K_1)\delta^{-1}|f|_0^2.$$

The process can be repeated until the case $j = \beta$, and we will have

$$|f|_\alpha^2 \le K\delta^{\beta-1}|f|_\beta^2 + K\delta^{-1}|f|_0^2,$$

with

$$K := \max\left(\prod_{j=1}^{\beta} K_j,\ \sum_{k=1}^{\beta} \prod_{j=k}^{n} K_j \right).$$

Now, let $\varepsilon := \delta^{\beta-1}$, then we obtain the desired inequality (14). □

Proof of Theorem 3. Denote the difference function $e(x) := y_{\beta,\sigma}(x) - f(x) \in X_\beta$. According to Theorem 1, we can get $|e|_0 \leq 1$ for sufficiently small δ and h. Now, by setting $\varepsilon = |e|_0^{\frac{2(\beta-1)}{\beta}}$ in Theorem 4 and the fact that $\sqrt{r^2+1} < r+1$ for $0 < r < 1$, we get

$$|e|_\alpha \leq \sqrt{K}\big(|e|_\beta + 1\big)\, |e|_0^{\frac{(\beta-1)}{\beta}}.$$

Note that $|e|_0 = \|y_{\beta,\sigma} - f\|_{L^2(a,b)}$. We know from Theorem 1 that $|e|_0 \leq Ph^{\beta-\frac{1}{2}} + Q\delta$, and from Theorem 2 that $|e|_\beta \leq (2 + |f|_\beta^2)^{1/2}$. Hence,

$$\begin{aligned}
|e|_\alpha &\leq \sqrt{K}\left(\sqrt{2+|f|_\beta^2}+1\right)\big(Ph^{\beta-\frac{1}{2}} + Q\delta\big)^{\frac{\beta-1}{\beta}} \\
&\leq \sqrt{K}\left(\sqrt{2}+2|f|_\beta+1\right)\left(P^{\frac{\beta-1}{\beta}} h^{(\beta-\frac{1}{2})\cdot\frac{\beta-1}{\beta}} + \big(Q\delta\big)^{\frac{\beta-1}{\beta}}\right) \\
&= \sqrt{K}\left(\sqrt{2}+2|f|_\beta+1\right)\left(P^{\frac{\beta-1}{\beta}} h^{\beta-\frac{3}{2}+\frac{1}{2\beta}} + Q^{\frac{\beta-1}{\beta}}\delta^{\frac{\beta-1}{\beta}}\right) \\
&=: \sqrt{K}\left(\sqrt{2}+2|f|_\beta+1\right)\left(Ch^{\beta-\frac{3}{2}+\frac{1}{2\beta}} + D\delta^{\frac{\beta-1}{\beta}}\right).
\end{aligned}$$

By noting that $|e|_\alpha = \|{}_0^cD^\alpha e\|_{L^2(a,b)} = \|{}_0^cD^\alpha y_{\beta,\sigma} - {}_0^cD^\alpha f\|_{L^2(a,b)}$, the theorem is proved. The constants in this theorem and those in Theorem 1 and Theorem 4 are related by $C = \sqrt{K}\left(\sqrt{2}+2|f|_\beta+1\right)P^{\frac{\beta-1}{\beta}}$ and $D = \sqrt{K}\left(\sqrt{2}+2|f|_\beta+1\right)Q^{\frac{\beta-1}{\beta}}$. □

4 Numerical examples

To demonstrate the effectiveness and stability of the proposed numerical differentiation scheme, we consider the smooth exact function $f(x) = \cos(x) + x^3$ on $[0,1]$. The Caputo derivatives for various orders α are shown in Figure 1.

As the proven theories in the previous sections do not have numerical consideration included, in Figure 2 we display the condition numbers of the matrix system (8) for various smoothness β and regularization parameters σ as the number of uniformly distributed data n increases. When $\sigma = 0$, we are looking at the standard RBF interpolation matrices; it is well known that higher smoothness (hence faster theoretical convergence rate) and/or larger n yields a more ill-conditioned matrix system. The matrix system (8), however, becomes better conditioned once we start regularizing (even with a tiny parameter $\sigma = 10^{-6}$).

Next, we want to verify the h-convergence behaviour implied from Theorem 3. In Figure 3, the L_2-errors of the approximate Caputo derivatives $\|{}_0^cD^\alpha y_{\beta,\sigma} - {}_0^cD^\alpha f\|_{L^2(0,1)}$ are based on noise-free data. As $\delta = 0$, the regularization parameter σ is also 0 for both selection strategies. The most obvious observation is that the convergence stagnates for small space distances h; this

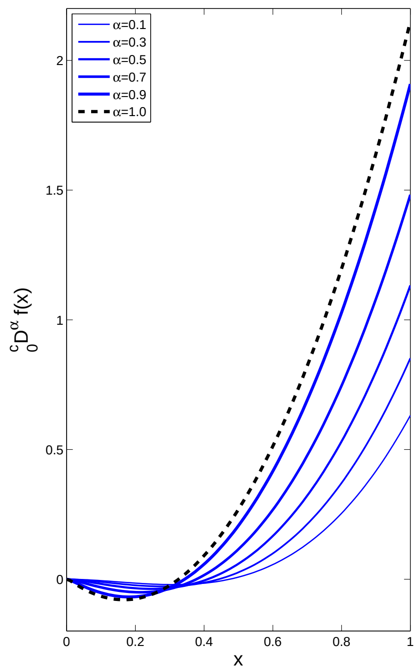


Fig. 1 Caputo derivatives of $f(x) = \cos(x) + x^3$ for various orders α.

is a behaviour also observable in the integer order cases. Before stagnation, we can see that higher smoothness does yield faster convergence. Moreover, the convergence behaviour is independent of differentiation order α. We remark that the radial powers RBF basis will not achieve exponential convergence. For direct problem, when one wants to discretize or evaluates Caputo derivatives from *noise-free data*, the recent work of Piret and Hanert [18] using Gaussian basis will be a much better choice.

The next test aims to study the δ-convergence behaviour. We use a rather large number of data points, $n = 2^7$, to suppress the error due to the space distance. Errors for both σ selection strategies are shown in Figure 4; both strategies yield similar convergence behaviour (as long as they work) and accuracy. Without a doubt, determining σ by the Morozov's discrepancy principle is numerically more expensive in comparison with the *a priori* strategy, $\sigma = \delta^2$. Moreover, it gets harder and harder to determine the correct Morozov's discrepancy principle regularization parameter for small noise. The failure of $\sigma = DP$ for small δ in Figure 4 could be corrected if one invests more effort to improve the subroutine of discrepancy principle; it is obvious that picking $\sigma = \delta^2$ is simple and works as well. Noise level, in many cases, is another unknown that requires some knowledgable guesses or approxima-

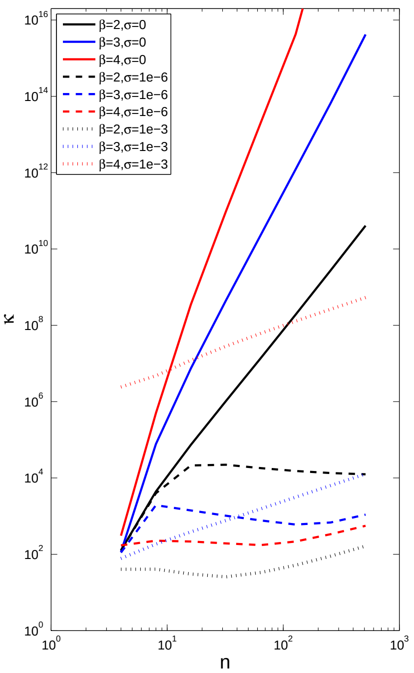


Fig. 2 Condition numbers of the resultant matrix in (8) with various kernels' smoothness β and regularization parameter σ.

tions. This means, in practice, one can only hope for having a value $\tilde{\delta} \approx \delta$ in the bound (3) for noise level. For comparison to other methods, the list of available methods for this inverse problem is rather short. In Figure 4, results obtained by the regularized finite difference scheme (with $N = 500$ points) in [9] are also shown. Convergence behavior is very much identical to our *a priori* strategy, but accuracy is off by a factor of 10. As the mollification technique [16] is not too accurate, even in the case of no noise, we hence did not show any comparison.

Instead of the Morozov's discrepancy principle, we suggest taking regularization parameter simply by $\sigma = \delta^2$; this should be simple as long as one knows what δ is. Figure 5, in which we show similar δ-convergence behaviours with various over- and under-estimated noise levels, shows that our scheme is rather forgiving even if one incorrectly picks a noise level ten times larger than the real value.

Edge detection in X-ray tomography [5] is an known application of fractional differentiation. We shall end the conclusion with an application of edge detection. The role of the Caputo differentiation order is to control the localization of data used in the reconstruction, which allows higher flexibility [27]

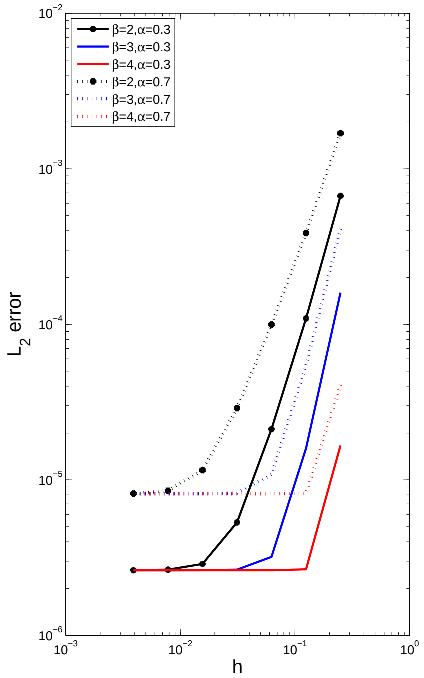


Fig. 3 h-Convergence behaviour ($\delta = 0$) of various kernel's smoothness β and differentiation order α.

by having one extra parameter (i.e., the order of differentiation) to alter the output images.

Similar to the numerical procedure given in [25], we apply the proposed scheme in both x- and y-direction to the standard test image—head *phantom*; see Figure 6. This results in the top images in Figure 7. To identify edges, cutting threshold is calculated by a simple average of the magnitude of the resulting differentiated image. The bottom images in Figure 7 show the pixels having values above the threshold. The results obviously look different to human eyes even though both reveal similar details. Using fractional order Caputo derivative results in images with the feel of pen sketching. The drawback is that the directional property of the Caputo derivative leaves some artifacts (in the left and bottom of the outer ellipse). We did not make any claim about one model (i.e., integer verse fractional order) being superior to the other; we believe the answer lies in the actual application itself.

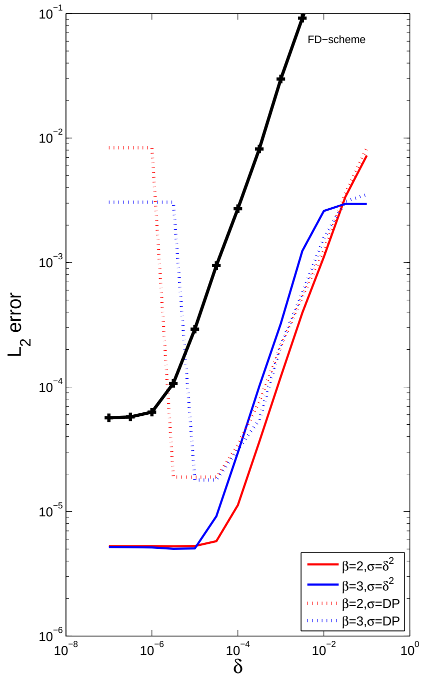
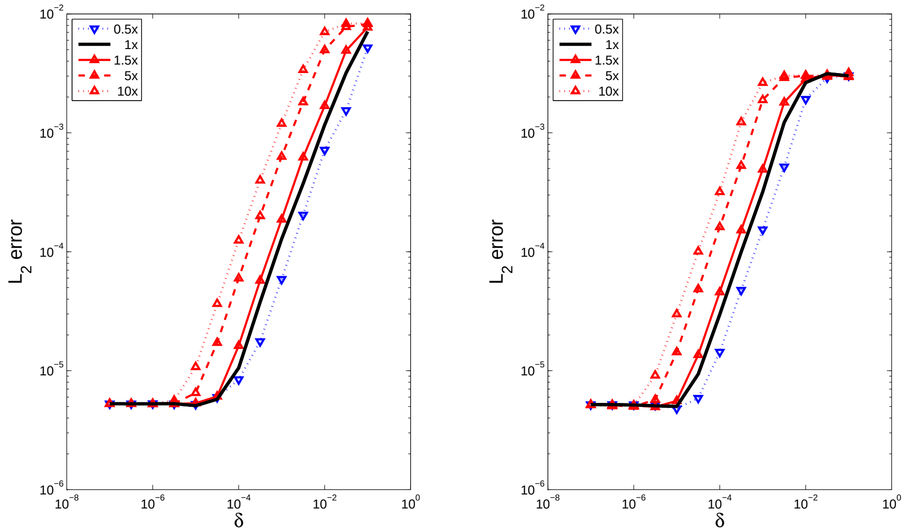


Fig. 4 δ-Convergence behaviours ($\alpha = 0.5$) of various kernels' smoothness β and differentiation order α.

5 Conclusion

A numerical scheme for approximating Caputo derivatives from noisy data is proposed. Our method starts with a regularized radial basis function interpolant and follows by taking Caputo differentiation of the basis. Convergence and error estimates are theoretically proven for proper regularization parameters. Some numerical examples are provided to demonstrate the convergence behaviours of the proposed scheme and its application to image processing. It is not the scope of this work to study edge detection algorithms, but combining the proposed algorithm of fractional differentiation with existing edge detection models could open up a new area of research for more real-life applications.

Acknowledgements This project was supported by the CERG Grant of the Hong Kong Research Grant Council and the FRG Grant of the Hong Kong Baptist University.



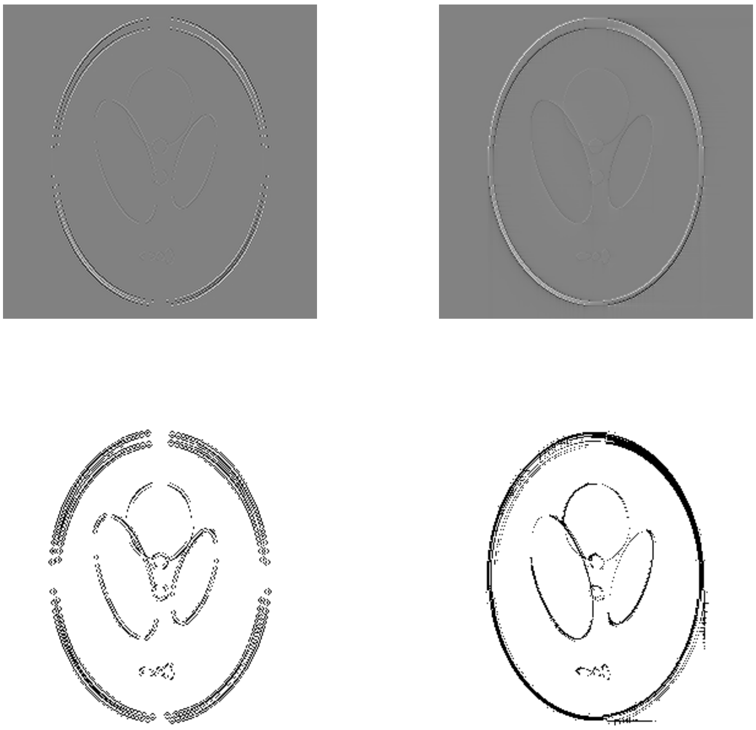
$\beta = 2$ $\beta = 3$

Fig. 5 δ-Convergence behaviour ($\alpha = 0.5$, $\sigma = \widetilde{\delta}^2$) of various over- and underestimated noise levels $\widetilde{\delta}$.

References

1. R. A. Adams. *Sobolev spaces*. Academic Press, New York-London, 1975. Pure and Applied Mathematics, Vol. 65.
2. R. S. Anderssen and P. Bloomfield. Numerical differentiation procedures for non-exact data. *Numerische Mathematik*, 22(3):157–182, 1974.
3. H. Brunner, L. Ling, and M. Yamamoto. Numerical simulations of 2D fractional subdiffusion problems. *J. Comput. Phys.*, 229(18):6613–6622, 2010.
4. S. R. Durrans. Distributions of fractional order statistics in hydrology. *Water Resources Research*, 28(6):1649–1655, 1992.
5. A. Faridani, L. Ritman, E, and K. T. Smith. Local tomography. *SIAM J. Appl. Math.*, 52(2):459484, 1992.
6. G. E. Fasshauer. *Meshfree approximation methods with Matlab*. Interdisciplinary Mathematical Sciences 6. Hackensack, NJ: World Scientific., 2007.
7. A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo. *Theory and applications of fractional differential equations*, volume 204 of *North-Holland Mathematics Studies*. Elsevier Science B.V., Amsterdam, 2006.
8. H. C. Kranzer. An error formula for numerical differentiation. *Numerische Mathematik*, 5(1):439–442, 1963.
9. M. Li, X. T. Xiong, and Y. J. Wang. A numerical evaluation and regularization of caputo fractional derivatives. *Journal of Physics: Conference Series*, 290(1):012011, 2011.
10. X. Li and C. Xu. A space-time spectral method for the time fractional diffusion equation. *SIAM J. Numer. Anal.*, 47(3):2108–2131, 2009.
11. X. Li and C. Xu. Existence and uniqueness of the weak solution of the space-time fractional diffusion equation and a spectral method approximation. *Commun. Comput. Phys.*, 8(5):1016–1051, 2010.
12. L. Ling. Finding numerical derivatives for unstructured and noisy data by multiscale kernels. *SIAM J. Numer. Anal.*, 44(4):1780–1800, 2006.

Fig. 6 The image of a head phantom used for edge detection.



$\alpha = 1$ $\alpha = 0.5$

Fig. 7 Results of edge detection by integer and Caputo fractional derivatives.

13. L. Ling and M. Yamamoto. Numerical simulations for space-time fractional diffusion equations. *International Journal of Computational Methods*, 10(2), 2013.
14. D. A. Murio. Stable numerical solution of a fractional-diffusion inverse heat conduction problem. *Comput. Math. Appl.*, 53(10):1492–1501, 2007.
15. D. A. Murio and C. E. Mejia. Generalized time fractional IHCP with Caputo fractional derivatives. *Journal of Physics: Conference Series*, 135(1):012074, 2008.
16. D.A. Murio. On the stable numerical evaluation of caputo fractional derivatives. *Comput. Math. Appl.*, 51(910):1539 – 1550, 2006.
17. P. Novati. Numerical approximation to the fractional derivative operator. *Numerische Mathematik*, pages 1–28, 2013.
18. C. Piret and E. Hanert. Fractional differential operator discretization using the radial basis functions method. *J. Comput. Phys.*, To appear.
19. I. Podlubny. *Fractional differential equations*, volume 198 of *Mathematics in Science and Engineering*. Academic Press Inc., San Diego, CA, 1999.
20. K. Sakamoto and M. Yamamoto. Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems. *Math. Anal. Appl.*, 382(1):426–447, 2011.
21. K. Sakamoto and M. Yamamoto. Inverse source problem with a final overdetermination for a fractional diffusion equation. *Math. Control Related Fields*, 1(4):509–518, 2011.
22. A. Shirzadi, L. Ling, and S. Abbasbandy. Meshless simulations of the two-dimensional fractional-time convection-diffusion-reaction equations. *Eng. Anal. Bound. Elem.*, 36(11):1522–1527, 2012.
23. A. N. Tikhonov and V. Y. Arsenin. *Solutions of ill-posed problems*. V. H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York, 1977. Translated from the Russian, Preface by translation editor Fritz John, Scripta Series in Mathematics.
24. R. Vilela Mendes. A fractional calculus interpretation of the fractional volatility model. *Nonlinear Dynamics*, 55(4):395–399, 2009.
25. X. Q. Wan, Y. B. Wang, and M Yamamoto. Detection of irregular points by regularization in numerical differentiation and application to edge detection. *Inverse Problems*, 22(3):1089–1103, 2006.
26. T. Wei and M. Li. High order numerical derivatives for one-dimensional scattered noisy data. *Appl. Math. Comput.*, 175(2):1744–1759, 2006.
27. J. Zhao, L. Zhang, W. Zheng, H. Tian, D.-M. Hao, and S.-H. Wu. Normalized cut segmentation of thyroid tumor image based on fractional derivatives. In Jing He, Xiaohui Liu, ElizabethA. Krupinski, and Guandong Xu, editors, *Health Information Science*, volume 7231 of *Lecture Notes in Computer Science*, pages 100–109. Springer Berlin Heidelberg, 2012.
28. G. H. Zheng and T. Wei. Spectral regularization method for a Cauchy problem of the time fractional advection-dispersion equation. *Journal of Computational and Applied Mathematics*, 233(10):2631–2640, 2010.
29. H. Zhou, W. Y. Tian, and W. Deng. Quasi-compact finite difference schemes for space fractional diffusion equations. *Journal of Scientific Computing*, 56(1):45–66, 2013.