CORRECTION OF HIGH-ORDER BDF CONVOLUTION QUADRATURE FOR FRACTIONAL EVOLUTION EQUATIONS *

BANGTI JIN†, BUYANG LI‡, AND ZHI ZHOU§

Abstract. We develop proper correction formulas at the starting $k-1$ steps to restore the desired k^{th}-order convergence rate of the k-step BDF convolution quadrature for discretizing evolution equations involving a fractional-order derivative in time. The desired k^{th}-order convergence rate can be achieved even if the source term is not compatible with the initial data, which is allowed to be nonsmooth. We provide complete error estimates for the subdiffusion case $\alpha \in (0,1)$, and sketch the proof for the diffusion-wave case $\alpha \in (1,2)$. Extensive numerical examples are provided to illustrate the effectiveness of the proposed scheme.

Key words. fractional evolution equation, convolution quadrature, initial correction, backward differentiation formulas, nonsmooth and incompatible data, error estimates

AMS subject classifications. 65M60, 65N30, 65N15, 35R11

1. Introduction. We are interested in the convolution quadrature (CQ) generated by high-order backward differentiation formulas (BDFs) for solving the fractional-order evolution equation (with $0 < \alpha < 1$)

$$\begin{cases} \partial_t^\alpha (u(t) - v) - Au(t) = f(t), & 0 < t < T, \\ u(0) = v, \end{cases} \tag{1.1}$$

where f is a given function, and $\partial_t^\alpha u$ denotes the left-sided Riemann-Liouville fractional time derivative of order α, defined by (cf. [20])

$$\partial_t^\alpha u(t) := \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t (t-s)^{-\alpha} u(s)\, ds, \tag{1.2}$$

where $\Gamma(z) := \int_0^\infty s^{z-1} e^{-s} ds$ is the Gamma function. Under the initial condition $u(0) = v$, the Riemann-Liouville fractional derivative $\partial_t^\alpha (u - v)$ in the model (1.1) is identical with the usual Caputo fractional derivative [20, pp. 91].

In the model (1.1), the operator A denotes either the Laplacian Δ on a polyhedral domain $\Omega \subset \mathbb{R}^d$ $(d = 1, 2, 3)$ with a homogenous Dirichlet boundary condition, or its Galerkin finite element approximation Δ_h. Thus A satisfies the following resolvent estimate (cf. [1, Example 3.7.5 and Theorem 3.7.11] and [37])

$$\|(z - A)^{-1}\|_{L^2(\Omega)\to L^2(\Omega)} \le c_\phi |z|^{-1}, \quad \forall z \in \Sigma_\phi, \tag{1.3}$$

for all $\phi \in (\pi/2, \pi)$, where $\Sigma_\theta := \{z \in \mathbb{C} \setminus \{0\} : |\arg z| < \theta\}$ is a sector of the complex plane $\mathbb{C}$. The model (1.1) covers a broad range of applications related to anomalous diffusion, e.g., dynamics of protein molecules, contaminant transport in complex geological formations and relaxation in polymer systems [35].

*The work of B. Jin is supported by UK EPSRC grant EP/M025160/1. The work of B. Li is partially supported by the Hong Kong RGC grant 15300817. The work of Z. Zhou was supported in part by the AFOSR MURI center for Material Failure Prediction through peridynamics and the ARO MURI Grant W911NF-15-1-0562.

†Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK. Email address: b.jin@ucl.ac.uk

‡Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong. Email address: buyang.li@polyu.edu.hk

§Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong. Email address: zhizhou@polyu.edu.hk

There has been much recent interest in developing high-order schemes for problem (1.1), especially spectral methods [3, 4, 21, 41] and discontinuous Galerkin [8, 28–30]. In this work, we develop robust high-order schemes based on CQs generated by high-order BDFs. The CQ developed by Lubich [23–25] provides a flexible framework for constructing high-order methods to discretize the fractional derivative $\partial_t^\alpha u$. By its very construction, it inherits the stability property of linear multistep methods, which greatly facilitates the analysis of the resulting numerical scheme, in a way often strikingly opposed to standard quadrature formulas [25, pp. 504]. Hence, it has been widely applied to discretize the model (1.1), especially the CQ generated by BDF1 and BDF2 (with BDFk denoting BDF of order k). In the literature, the CQ generated by BDF1 is commonly known as the Grünwald-Letnikov formula (see [2] for related discussions and its high-order variants).

By assuming that the solution is sufficiently smooth, which is equivalent to assuming smoothness of the initial data v and imposing certain *compatibility* conditions on the source term f at $t = 0$, the stability and convergence of the numerical solutions of fractional evolution equations have been investigated in [7, 11, 38, 40, 42]. In general, if the source term f is not compatible with the given initial data, the solution u of the model (1.1) will exhibit weak singularity at $t = 0$, which will deteriorate the convergence rate of the numerical solutions. This has been widely recognized in fractional ODEs [9, 10] and PDEs [6, 17, 34]. In particular, direct implementation of the CQ generated by high-order BDFs for discretizing the fractional evolution equations generally only yields first-order accuracy. To restore the theoretical $O(\tau^k)$ rate of BDFk, two different strategies have been proposed.

For fractional ODEs, one idea is to use starting weights [23] to correct the CQ in discretizing the fractional time derivative $\partial_t^\alpha \varphi(t_n)$, cf. (2.1) below, by

$$\bar\partial_\tau^\alpha \varphi^n = \frac{1}{\tau^\alpha}\sum_{j=0}^{n} b_{n-j}\varphi^j + \sum_{j=0}^{M} w_{n,j}\varphi^j,$$

where $M \in \mathbb{N}$ and the weights $w_{n,j}$ depend on α and k. The starting term $\sum_{j=0}^{M} w_{n,j}\varphi^j$ is to capture all leading singularities so as to recover a *uniform* $O(\tau^k)$ rate of the scheme. This approach works well for fractional ODEs, however, its extension to fractional PDEs relies on expanding the solution into power series of t, which requires imposing certain compatibility conditions on the source f.

The second idea is to split f into $f(t) = f(0) + (f(t) - f(0))$ and to approximate $f(0)$ by $\bar\partial_\tau \partial_t^{-1} f(0)$, with a similar treatment of the initial data v. This leads to a corrected BDF2 at the first step and restores the $O(\tau^2)$ rate for any $t_n > 0$. The idea was introduced in [26] for solving a variant of (1.1) in the diffusion-wave case and then systematically developed in [6] for BDF2, and was recently extended to (1.1) in [17] for both subdiffusion and diffusion-wave cases. Higher-order extension of this idea is possible, but is still unavailable in the literature.

The goal of this work is to develop robust high-order BDFs for fractional evolution equations along the second strategy [6, 17]. Instead of extending this strategy to each high-order BDF method, separately, we develop a systematic strategy for correcting initial steps for high-order BDFs, based on a few simple criteria, cf. (2.13) and (2.14) for the model (1.1). These criteria emerge naturally from solution representations, and are purely algebraic and straightforward to construct. The explicit correction coefficients will be given for BDFs up to order 6. For BDFk, the correction is only needed at the starting $k-1$ steps and thus the resulting scheme is easy to implement.

We develop proper corrections for high-order BDFs for both subdiffusion, i.e., $\alpha \in (0, 1)$, and diffusion wave, i.e., $\alpha \in (1, 2)$. It is noteworthy that for $\alpha \in (1, 2)$, high-order BDFs can be either unconditionally or conditionally stable, depending on the fractional order α, and in the latter case, an explicit CFL condition on the time step size τ is given. Theoretically, the corrected BDFk achieves the k^{th}-order accuracy at any fixed time $t = t_n$ (when t_n is bounded from below), and the error bound depends only on data regularity, without assuming any compatibility conditions on the source term or extra regularity on the solution (cf. Theorems 2.2 and 3.2). These results are supported by the numerical experiments in Section 4.

The rest of the paper is organized as follows. In Section 2 we develop the correction for the subdiffusion case, including the motivations of the algebraic criteria for choosing the correction coefficients. The extension of the approach to the diffusion wave case is given in Section 3. Numerical results are presented in Section 4 to illustrate the efficiency and robustness of the corrected schemes. Appendix A gives an alternative interpretation of our correction method in terms of Lubich's convolution quadrature for operator-valued convolution integrals. Some lengthy proofs are given in Appendices B–E. Throughout, the notation c denotes a generic positive constant, whose value may differ at each occurrence, but it is always independent of the time step size τ and the solution u.

2. BDFs for Subdiffusion and its Correction. Let $\{t_n = n\tau\}_{n=0}^N$ be a uniform partition of the interval $[0, T]$, with a time step size $\tau = T/N$. The CQ generated by BDFk, $k = 1, \ldots, 6$, approximates the fractional derivative $\partial_t^\alpha \varphi(t_n)$ by

$$\bar\partial_\tau^\alpha \varphi^n := \frac{1}{\tau^\alpha} \sum_{j=0}^{n} b_j \varphi^{n-j}, \tag{2.1}$$

with $\varphi^n = \varphi(t_n)$, where the weights $\{b_j\}_{j=0}^\infty$ are the coefficients in the series expansion

$$\delta_\tau(\zeta)^\alpha = \frac{1}{\tau^\alpha} \sum_{j=0}^{\infty} b_j \zeta^j \quad \text{with} \quad \delta_\tau(\zeta) := \frac{1}{\tau} \sum_{j=1}^{k} \frac{1}{j}(1-\zeta)^j. \tag{2.2}$$

Below we often write $\delta(\zeta) = \delta_1(\zeta)$, i.e., with $\tau = 1$. The coefficients b_j can be computed efficiently by the fast Fourier transform [32,36] or recursion [39]. Correspondingly, the BDF for solving (1.1) seeks approximations U^n, $n = 1, \ldots, N$, to the exact solution $u(t_n)$ by

$$\bar\partial_\tau^\alpha (U - v)^n - AU^n = f(t_n). \tag{2.3}$$

If the solution u is smooth and has sufficiently many vanishing derivatives at 0, then U^n converges at a rate $O(\tau^k)$ [23,25]. However, it generally only exhibits a first-order accuracy when solving fractional PDEs, due to the weak solution singularity at 0, even if the initial data v and source term f are smooth [33], as observed numerically [6,17]. For $\alpha = 1$, BDFk is known to be $A(\vartheta_k)$-stable with angle $\vartheta_k = 90°$, $90°$, $86.03°$, $73.35°$, $51.84°$, $17.84°$ for $k = 1, 2, 3, 4, 5, 6$, respectively [14, pp. 251].

To restore the k^{th}-order accuracy, we correct BDFk at the starting $k-1$ steps by (as usual, the summation disappears if the upper index is smaller than the lower one)

$$\begin{aligned} &\bar\partial_\tau^\alpha (U - v)^n - AU^n = a_n^{(k)}(Av + f(0)) + f(t_n) + \sum_{\ell=1}^{k-2} b_{\ell,n}^{(k)} \tau^\ell \partial_t^\ell f(0), && 1 \le n \le k-1, \\ &\bar\partial_\tau^\alpha (U - v)^n - AU^n = f(t_n), && k \le n \le N. \end{aligned} \tag{2.4}$$

where $a_n^{(k)}$ and $b_{\ell,n}^{(k)}$ are coefficients to be determined below. They are constructed so as to improve the accuracy of the overall scheme to $O(\tau^k)$ for a general $v \in D(A)$ and a possibly incompatible source f. The only difference between (2.4) and (2.3) lies in the correction terms at the starting $k-1$ steps. Hence, the scheme (2.4) is easy to implement.

Remark 2.1. *In the scheme* (2.4)*, the derivative* $\partial_t^\ell f(0)$ *may be replaced by its* $(k-\ell-1)$-

order finite difference approximation $f^{(\ell)}$, without sacrificing its accuracy.

REMARK 2.2. *The correction in* (2.4) *is minimal in the sense that there is no other correction scheme which modifies only the $k-1$ starting steps while retaining the $O(\tau^k)$ rate. This does not rule out corrections with more starting steps. We give an alternative correction closely related to* (2.4) *in Appendix A.*

2.1. Derivation of the correction criteria. Now we derive the criteria for choosing the coefficients $a_j^{(k)}$ and $b_{\ell,j}^{(k)}$, cf. (2.13) and (2.14) below, using Laplace transform and its discrete analogue, the generating function [26, 37]. We denote by $\widehat{\ }$ taking Laplace transform, and for a given sequence $(f^n)_{n=0}^\infty$, denote by $\widetilde{f}(\zeta) := \sum_{n=0}^\infty f^n\zeta^n$ its generating function. First we split the right hand side f into

$$f(t) = f(0) + \sum_{\ell=1}^{k-2} \frac{t^\ell}{\ell!}\partial_t^\ell f(0) + R_k, \tag{2.5}$$

and R_k is the corresponding local truncation error, given by

$$R_k = f(t) - f(0) - \sum_{\ell=1}^{k-2} \frac{t^\ell}{\ell!}\partial_t^\ell f(0) = \frac{t^{k-1}}{(k-1)!}\partial_t^{k-1} f(0) + \frac{t^{k-1}}{(k-1)!} * \partial_t^k f, \tag{2.6}$$

where $*$ denotes Laplace convolution. Thus the function $w(t) := u(t) - v$ satisfies

$$\partial_t^\alpha w - Aw = Av + f(0) + \sum_{\ell=1}^{k-2} \frac{t^\ell}{\ell!}\partial_t^\ell f(0) + R_k, \tag{2.7}$$

with $w(0) = 0$. Since $w(0) = 0$, the identity $\widehat{\partial_t^\alpha w}(z) = z^\alpha \widehat{w}(z)$ holds [20, Remark 2.8, pp. 84], and thus by Laplace transform, we obtain

$$z^\alpha \widehat{w}(z) - A\widehat{w}(z) = z^{-1}(Av + f(0)) + \sum_{\ell=1}^{k-2} \frac{1}{z^{\ell+1}}\partial_t^\ell f(0) + \widehat{R}_k(z).$$

By inverse Laplace transform, the function $w(t)$ can be readily represented by

$$\begin{aligned} w(t) =& \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_{\theta,\delta}} e^{zt} K(z)\big(Av + f(0)\big)dz \\ &+ \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_{\theta,\delta}} e^{zt} zK(z)\Big(\sum_{\ell=1}^{k-2} \frac{1}{z^{\ell+1}}\partial_t^\ell f(0) + \widehat{R}_k(z)\Big)dz, \end{aligned} \tag{2.8}$$

with the kernel function

$$K(z) = z^{-1}(z^\alpha - A)^{-1}. \tag{2.9}$$

In the representation (2.8), the contour $\Gamma_{\theta,\delta}$ is defined by

$$\Gamma_{\theta,\delta} = \{z \in \mathbb{C} : |z| = \delta, |\arg z| \le \theta\} \cup \{z \in \mathbb{C} : z = \rho e^{\pm \mathrm{i}\theta}, \rho \ge \delta\},$$

oriented with an increasing imaginary part. Throughout, we choose the angle θ in $\Gamma_{\theta,\delta}$ such that $\pi/2 < \theta < \min(\pi, \pi/\alpha)$ and hence, $z^\alpha \in \Sigma_{\theta'}$ with $\theta' = \alpha\theta < \pi$ for all $z \in \Sigma_\theta$. By the resolvent estimate (1.3), there exists a constant c which depends only on θ and α such that

$$\|(z^\alpha - A)^{-1}\| \le c|z|^{-\alpha} \quad \text{and} \quad \|K(z)\| \le c|z|^{-1-\alpha}, \quad \forall z \in \Sigma_\theta. \tag{2.10}$$

Next, we give a representation of the discrete solution $W^n := U^n - v$, which follows from lengthy but simple computation, cf. Appendix B.

THEOREM 2.1. *Let $f \in C^{k-1}([0,T]; L^2(\Omega))$ and $\int_0^t (t-s)^{\alpha-1}\|\partial_s^k f(s)\|_{L^2(\Omega)}ds < \infty$. The* *discrete solution* $W^n := U^n - v$ *is represented by*

$$
\begin{aligned}
W^n = {} & \frac{1}{2\pi \mathrm{i}}\int_{\Gamma^\tau_{\theta,\delta}} e^{zt_n}\mu(e^{-z\tau})K(\delta_\tau(e^{-z\tau}))(Av+f(0))\,dz \\
& + \frac{1}{2\pi \mathrm{i}}\int_{\Gamma^\tau_{\theta,\delta}} e^{zt_n}\delta_\tau(e^{-z\tau})K(\delta_\tau(e^{-z\tau}))\sum_{\ell=1}^{k-2}\Big(\frac{\gamma_\ell(e^{-z\tau})}{\ell!}+\sum_{j=1}^{k-1}b_{\ell,j}^{(k)}e^{-zt_j}\Big)\tau^{\ell+1}\partial_t^\ell f(0)\,dz \\
& + \frac{1}{2\pi \mathrm{i}}\int_{\Gamma^\tau_{\theta,\delta}} e^{zt_n}\delta_\tau(e^{-z\tau})K(\delta_\tau(e^{-z\tau}))\tau\widetilde{R}_k(e^{-z\tau})\,dz,
\end{aligned}
\tag{2.11}
$$

with the contour $\Gamma^\tau_{\theta,\delta} := \{z \in \Gamma_{\theta,\delta} : |\Im(z)| \le \pi/\tau\}$ *(oriented with an increasing imaginary part), where the functions* $\mu(\zeta)$ *and* $\gamma_\ell(\zeta)$ *are respectively defined by*

$$
\mu(\zeta) = \delta(\zeta)\Big(\frac{\zeta}{1-\zeta}+\sum_{j=1}^{k-1}a_j^{(k)}\zeta^j\Big) \quad \textit{and} \quad \gamma_\ell(\zeta) = \Big(\zeta\frac{d}{d\zeta}\Big)^\ell\frac{1}{1-\zeta}. \tag{2.12}
$$

By comparing the kernel functions in (2.8) and (2.11), we deduce that in order to have $O(\tau^k)$ accuracy, the following three conditions should be satisfied for $z \in \Gamma^\tau_{\theta,\delta}$:

$$
\begin{aligned}
& |\delta_\tau(e^{-z\tau}) - z| \le c|z|^{k+1}\tau^k, \qquad |\mu(e^{-z\tau}) - 1| \le c|z|^k\tau^k, \\
& \Big|\Big(\frac{\gamma_\ell(e^{-z\tau})}{\ell!}+\sum_{j=1}^{k-1}b_{\ell,j}^{(k)}e^{-zt_j}\Big)\tau^{\ell+1} - \frac{1}{z^{\ell+1}}\Big| \le c|z|^{k-\ell-1}\tau^k.
\end{aligned}
$$

Note that for BDFk, the estimate $|\delta_\tau(e^{-z\tau})-z| \le c|z|^{k+1}\tau^k$ holds automatically (cf. Lemma B.1). It suffices to impose the following two criteria (by changing $e^{-z\tau}$ to ζ and $z\tau$ to $1-\zeta$): for BDFk, choose the coefficients $\{a_j^{(k)}\}_{j=1}^{k-1}$ and $\{b_{\ell,j}^{(k)}\}_{j=1}^{k-1}$ such that

$$
|\mu(\zeta) - 1| \le c|1-\zeta|^k, \tag{2.13}
$$

$$
\Big|\frac{\gamma_\ell(\zeta)}{\ell!}+\sum_{j=1}^{k-1}b_{\ell,j}^{(k)}\zeta^j - \frac{1}{\delta(\zeta)^{\ell+1}}\Big| \le c|1-\zeta|^{k-\ell-1}, \quad \ell = 1,\ldots,k-2, \tag{2.14}
$$

where the functions $\mu(\zeta)$ and $\gamma_\ell(\zeta)$ are defined in (2.12). It can be verified that for BDFk, $k = 3,\ldots,6$, the leading singularities on the left hand side of (2.14) do cancel out, and thus the criterion can be satisfied.

2.2. Computation of the coefficients $a_j^{(k)}$ and $b_{\ell,j}^{(k)}$. First we compute the coefficients $a_j^{(k)}$. To this end, we rewrite $\sum_{j=1}^{k-1}a_j^{(k)}\zeta^j$ as

$$
\sum_{j=1}^{k-1}a_j^{(k)}\zeta^j = \zeta\sum_{j=0}^{k-2}c_j(1-\zeta)^j. \tag{2.15}
$$

Thus, by writing $\zeta = 1-(1-\zeta)$, expanding the summation and collecting terms, we obtain (with the convention $c_{-2} = c_{-1} = 0$)

$$
\begin{aligned}
\mu(\zeta) &= \sum_{j=1}^{k}\frac{1}{j}(1-\zeta)^j\Big(\frac{\zeta}{1-\zeta}+\zeta\sum_{j=0}^{k-2}c_j(1-\zeta)^j\Big) \\
&= \sum_{j=0}^{k-1}\frac{1}{j+1}(1-\zeta)^j\Big(1-(1-\zeta)-\sum_{j=0}^{k}c_{j-2}(1-\zeta)^j+\sum_{j=0}^{k-1}c_{j-1}(1-\zeta)^j\Big)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{j=0}^{k-1} \frac{1}{j+1}(1-\zeta)^j - \sum_{j=1}^{k} \frac{1}{j}(1-\zeta)^j - \sum_{j=2}^{k-1}\Big(\sum_{\ell=0}^{j} \frac{1}{j-\ell+1} c_{\ell-2}\Big)(1-\zeta)^j \\
&\quad + \sum_{j=1}^{k-1}\Big(\sum_{\ell=0}^{j} \frac{1}{j-\ell+1} c_{\ell-1}\Big)(1-\zeta)^j + O\big((1-\zeta)^k\big) \\
&= 1 + \sum_{j=1}^{k-1}\Big(\frac{1}{j+1} - \frac{1}{j} - \sum_{\ell=0}^{j} \frac{1}{j-\ell+1} c_{\ell-2} + \sum_{\ell=0}^{j} \frac{1}{j-\ell+1} c_{\ell-1}\Big)(1-\zeta)^j + O\big((1-\zeta)^k\big) \\
&= 1 + \sum_{j=1}^{k-1}\Big(- \frac{1}{j(j+1)} - \sum_{\ell=1}^{j-1} \frac{1}{j-\ell} c_{\ell-1} + \sum_{\ell=0}^{j-1} \frac{1}{j-\ell} c_{\ell}\Big)(1-\zeta)^j + O\big((1-\zeta)^k\big).
\end{aligned}
$$

Thus by choosing c_ℓ, $\ell = 0, \ldots, k-2$, such that

$$
\sum_{\ell=0}^{j-1} \frac{1}{j-\ell} c_\ell = \frac{1}{j(j+1)} + \sum_{\ell=1}^{j-1} \frac{1}{j-\ell} c_{\ell-1}, \quad j = 1, \ldots, k-1, \tag{2.16}
$$

Criterion (2.13) follows. The coefficients $a_j^{(k)}$ can be computed recursively from (2.16) and (2.15), and are given in Table 1. The result for $k=2$ recovers exactly the correction in [17], and thus our algebraic construction generalizes the approach in [17].

TABLE 1
The coefficients $a_j^{(k)}$ computed by (2.15)

order of BDF	$a_1^{(k)}$	$a_2^{(k)}$	$a_3^{(k)}$	$a_4^{(k)}$	$a_5^{(k)}$
$k=2$	$\frac{1}{2}$				
$k=3$	$\frac{11}{12}$	$-\frac{5}{12}$			
$k=4$	$\frac{31}{24}$	$-\frac{7}{6}$	$\frac{3}{8}$		
$k=5$	$\frac{1181}{720}$	$-\frac{177}{80}$	$\frac{341}{240}$	$-\frac{251}{720}$	
$k=6$	$\frac{2837}{1440}$	$-\frac{2543}{720}$	$\frac{17}{5}$	$-\frac{1201}{720}$	$\frac{95}{288}$

Next we compute the coefficients $b_{\ell,j}^{(k)}$. First we expand $\frac{\gamma_\ell(\zeta)}{\ell!} - \frac{1}{\delta(\zeta)^{\ell+1}}$ in $1-\zeta$ as

$$
\frac{\gamma_\ell(\zeta)}{\ell!} - \frac{1}{\delta(\zeta)^{\ell+1}} = \sum_{j=0}^{k-\ell-2} g_{\ell,j}^{(k)}(1-\zeta)^j + O(|1-\zeta|^{k-\ell-1}), \tag{2.17}
$$

and then choose the coefficients $b_{\ell,j}^{(k)}$, $j = 1, \ldots, k-1$ to satisfy (2.14). To this end, we rewrite $\sum_{j=1}^{k-1} b_{\ell,j}^{(k)}\zeta^j$ into the following form:

$$
\sum_{j=1}^{k-1} b_{\ell,j}^{(k)}\zeta^j = \zeta \sum_{j=0}^{k-2} d_{\ell,j}^{(k)}(1-\zeta)^j = \sum_{j=0}^{k-2} d_{\ell,j}^{(k)}(1-\zeta)^j - \sum_{j=1}^{k-1} d_{\ell,j-1}^{(k)}(1-\zeta)^j. \tag{2.18}
$$

Then it suffices to choose

$$
d_{\ell,0}^{(k)} = -g_{\ell,0}^{(k)}, \tag{2.19a}
$$

$$
d_{\ell,j}^{(k)} = d_{\ell,j-1}^{(k)} - g_{\ell,j}^{(k)} \qquad \text{for } j = 1, \ldots, k-\ell-2, \tag{2.19b}
$$

$$
d_{\ell,j}^{(k)} = 0 \qquad \text{for } j = k-\ell-1, \ldots, k-2. \tag{2.19c}
$$

Now the coefficients $b_{\ell,j}^{(k)}$ can be computed recursively using (2.17), (2.19) and (2.18), and the results are given in Table 2. Note that for $k=4$ and 6, the coefficients $b_{k-2,j}^{(k)}, j=1,2\ldots,k-1$ vanish identically.

TABLE 2
The coefficients $b_{\ell,j}^{(k)}$.

order of BDF		$b_{\ell,1}^{(k)}$	$b_{\ell,2}^{(k)}$	$b_{\ell,3}^{(k)}$	$b_{\ell,4}^{(k)}$	$b_{\ell,5}^{(k)}$
$k=3$	$\ell=1$	$\frac{1}{12}$	0			
$k=4$	$\ell=1$	$\frac{1}{6}$	$-\frac{1}{12}$	0		
	$\ell=2$	0	0	0		
$k=5$	$\ell=1$	$\frac{59}{240}$	$-\frac{29}{120}$	$\frac{19}{240}$	0	
	$\ell=2$	$\frac{1}{240}$	$-\frac{1}{240}$	0	0	
	$\ell=3$	$\frac{1}{720}$	0	0	0	
$k=6$	$\ell=1$	$\frac{77}{240}$	$-\frac{7}{15}$	$\frac{73}{240}$	$-\frac{3}{40}$	0
	$\ell=2$	$\frac{1}{96}$	$-\frac{1}{60}$	$\frac{1}{160}$	0	0
	$\ell=3$	$-\frac{1}{360}$	$\frac{1}{720}$	0	0	0
	$\ell=4$	0	0	0	0	0

2.3. Error estimates. We state the error estimate for (2.4) in the following theorem, whose proof can be found in Appendix C.

THEOREM 2.2. *Let Criteria* (2.13) *and* (2.14) *hold, and* $f\in C^{k-1}([0,T];L^2(\Omega))$ *and* $\int_0^t(t-s)^{\alpha-1}\|\partial_s^k f(s)\|_{L^2(\Omega)}ds<\infty$. *Then for the solution* U^n *to* (2.4), *the following error estimate holds for any* $t_n>0$

$$\|U^n-u(t_n)\|_{L^2(\Omega)}\le c\tau^k\Bigg(t_n^{\alpha-k}\|f(0)+Av\|_{L^2(\Omega)}+\sum_{\ell=1}^{k-1}t_n^{\alpha+\ell-k}\|\partial_t^\ell f(0)\|_{L^2(\Omega)}$$
$$+\int_0^{t_n}(t_n-s)^{\alpha-1}\|\partial_s^k f(s)\|_{L^2(\Omega)}ds\Bigg).$$

REMARK 2.3. *Note that the estimate depends only on the regularity of* f *and* v, *rather than the regularity of* u. *Theorem 2.2 implies that for any fixed* $t_n=\text{const}>0$, *the rate is* $O(\tau^k)$ *for BDFk. In order to have a* uniform $O(\tau^k)$ *rate, the following compatibility conditions are needed:*

$$f(0)+Av=0,\quad \textit{and}\quad \partial_t^{(\ell)}f(0)=0,\quad \ell=1,\ldots,k-1.$$

Otherwise, the estimate deteriorates as $t\to0$, *in accordance with the regularity theory: the solution* (*and its derivatives*) *exhibits weak singularity at* $t=0$ *[33].*

REMARK 2.4. *The estimate in Theorem 2.2 requires* $Av\in L^2(\Omega)$, *i.e.,* v *is reasonably smooth. Upon minor modifications of the proof in Appendix C, one can derive a similar error estimate for* $v\in L^2(\Omega)$:

$$\|U^n-u(t_n)\|_{L^2(\Omega)}\le c\tau^k\Bigg(t_n^{-k}\|v\|_{L^2(\Omega)}+\sum_{\ell=0}^{k-1}t_n^{\alpha+\ell-k}\|\partial_t^\ell f(0)\|_{L^2(\Omega)}+\int_0^{t_n}(t_n-s)^{\alpha-1}\|\partial_s^k f(s)\|_{L^2(\Omega)}ds\Bigg).$$

3. Corrected BDF for diffusion-wave problem. Now we extend the strategy in Section 2 to the diffusion-wave problem, i.e., $1 < \alpha < 2$:

$$\partial_t^\alpha (u(t) - v - tb) - Au(t) = f(t), \tag{3.1}$$

with the initial conditions $u(0) = v$ and $u'(0) = b$, where

$$\partial_t^\alpha u(t) := \frac{1}{\Gamma(2-\alpha)} \frac{d^2}{dt^2} \int_0^t (t-s)^{1-\alpha} u(s)\, ds. \tag{3.2}$$

The main differences from the subdiffusion case lie in the extra initial condition b and better temporal smoothing property [17]. A direct implementation of BDFk can fail to yield the $O(\tau^k)$ rate, as the subdiffusion case, and further, requires unnecessarily high regularity on f. We shall develop a corrected scheme to tackle both issues. First, to exploit the extra smoothing, we rewrite f as $f = \partial_t g$ with $g = \partial_t^{-1} f$. Then (3.1) can be rewritten as

$$\partial_t^\alpha (u - v - tb) - Au = \partial_t g, \tag{3.3}$$

Next we correct the first $k-1$ steps, and seek approximations U^n, $n = 1, \ldots, N$, by

(3.4)

$$\begin{aligned} &\bar\partial_\tau^\alpha (U - v - tb)^n - AU^n = a_n^{(k)} Av + c_n^{(k)} \tau Ab + \bar\partial_\tau g^n + \sum_{\ell=1}^{k-2} b_{\ell,n}^{(k)} \tau^{\ell-1} \partial_t^{\ell-1} f(0), \quad 1 \le n \le k-1, \\ &\bar\partial_\tau^\alpha (U - v - tb)^n - AU^n = \bar\partial_\tau g^n, \qquad\qquad k \le n \le N. \end{aligned}$$

The scheme involves $\bar\partial_\tau g^n$, instead of f^n, which enables one to relax the regularity requirement on f. The correction terms are to ensure the desired $O(\tau^k)$ rate.

Now we derive algebraic criteria for choosing the coefficients in (3.4) using Laplace transform and generating function. First, since $g(0) = 0$, $g(t)$ can be split into

$$g(t) = \sum_{\ell=1}^{k-2} \frac{t^\ell}{\ell!} \partial_t^\ell g(0) + R_k = \sum_{\ell=1}^{k-2} \frac{t^\ell}{\ell!} \partial_t^{\ell-1} f(0) + R_k, \tag{3.5}$$

where R_k is the local truncation error $R_k = \frac{t^{k-1}}{(k-1)!} \partial_t^{k-1} g(0) + \frac{t^{k-1}}{(k-1)!} * \partial_t^k g(t)$. Thus, the function $w = u - v - tb$ satisfies

$$\partial_t^\alpha w - Aw = Av + tAb + \sum_{\ell=1}^{k-2} \partial_t \frac{t^\ell}{\ell!} \partial_t^{\ell-1} f(0) + \partial_t R_k.$$

By Laplace transform and the kernel $K(z)$ from (2.9), we derive a representation of $w(t)$:

$$\begin{aligned} w(t) =& \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_{\theta,\delta}} e^{zt} K(z)(Av + z^{-1} Ab) dz \\ &+ \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_{\theta,\delta}} e^{zt} z K(z) \Big(\sum_{\ell=1}^{k-2} \frac{1}{z^\ell} \partial_t^{\ell-1} f(0) + z \widehat{R}_k(z) \Big) dz, \end{aligned} \tag{3.6}$$

where the angle $\theta \in (\pi/2, \pi)$ is close to $\pi/2$ such that $\alpha\theta < \pi$, and δ is small.

Since BDFk is $A(\vartheta_k)$-stable, the scheme (3.4) is unconditionally stable for any $\alpha < \alpha^*(k) := \pi/(\pi - \vartheta_k)$. The critical value $\alpha^*(k)$ is 1.91, 1.68, 1.40 and 1.11 for $k = 3, \ldots, 6$. In contrast, for $\alpha \ge \alpha^*(k)$, it is only conditionally stable. Note that for any $\alpha \in (1, 2)$, the curve $\delta(e^{-\mathrm{i}\theta})^\alpha$ is not tangent to the real axis at the origin (i.e., θ close to zero). This naturally gives rise to the following condition.

Condition 3.1. *Let $r(A)$ be the numerical radius of A, and the following condition holds:* (i) *$\alpha < \alpha^*(k)$ or* (ii) *$\alpha \ge \alpha^*(k)$ and $\tau^\alpha r(A) \le c(\alpha, k) - \gamma$ for some $\gamma > 0$, where the* *constant* $c(\alpha,k)$ *is given by the intersection point (closest to the origin) of* $\{\delta(\zeta)^\alpha : |\zeta|=1\}$ *with the negative real axis.*

REMARK 3.1. *Condition 3.1(ii) specifies the CFL condition on the time step size* τ *(so it holds only if* $r(A)<\infty$*). The CFL constant* $c(\alpha,k)$ *is not available in closed form, but can be determined numerically; see Fig. 1 for the values.*

It is interesting to observe the qualitative differences of BDFs of different order. For example, the CFL constant $c(\alpha,6)$ of BDF6 does not approach zero even for α tends to 2; and there is an interval of α values for which the CFL constant $c(\alpha,4)$ for BDF4 is larger than $c(\alpha,3)$ for BDF3, i.e., BDF4 is less stringent in time step size.

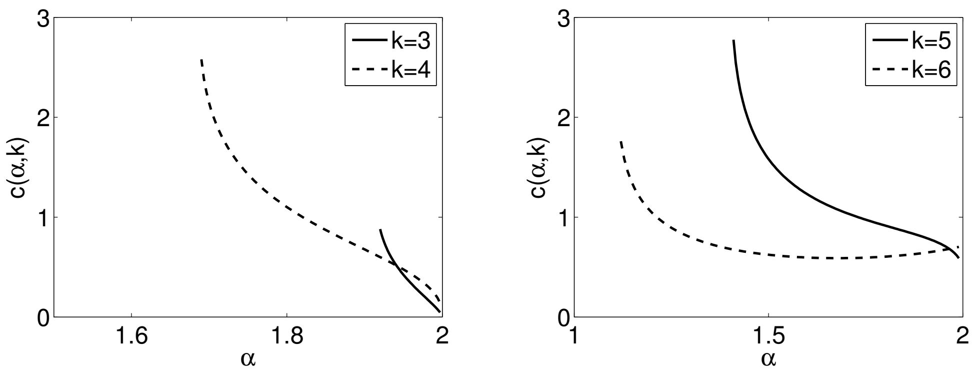


FIG. 1. *The CFL constant* $c(\alpha,k)$ *for BDFk,* $k=3,4,5,6$*, at different* α *values.*

The next result gives the representation of the solution $W^n = U^n - v - t_n b$, which follows from simple yet lengthy computations, cf. Appendix D.

THEOREM 3.1. *Let* $f\in C^{k-2}([0,T];L^2(\Omega))$ *and* $\int_0^t (t-s)^{\alpha-2}\|\partial_s^{k-1}f(s)\|_{L^2(\Omega)}ds<\infty$. *Under Condition 3.1,* $W^n := U^n - v - t_n b$ *is represented by*

$$
\begin{aligned}
W^n =& \frac{1}{2\pi\mathrm{i}}\int_{\Gamma^\tau_{\theta,\delta}} e^{zt_n}\mu(e^{-z\tau})K(\delta_\tau(e^{-z\tau}))Av\,dz \\
&+\frac{1}{2\pi\mathrm{i}}\int_{\Gamma^\tau_{\theta,\delta}} e^{zt_n}K(\delta_\tau(e^{-z\tau}))\delta_\tau(e^{-z\tau})\Big(\gamma_1(e^{-z\tau})+\sum_{j=1}^{k-1}c_j^{(k)}e^{-zt_j}\Big)\tau^2Ab\,dz \\
&+\frac{1}{2\pi\mathrm{i}}\int_{\Gamma^\tau_{\theta,\delta}} e^{zt_n}\delta_\tau(e^{-z\tau})K(\delta_\tau(e^{-z\tau}))\sum_{\ell=1}^{k-2}\Big(\delta(e^{-z\tau})\frac{\gamma_\ell(e^{-z\tau})}{\ell!}+\sum_{j=1}^{k-1}b_{\ell,j}^{(k)}e^{-zt_j}\Big)\tau^\ell\partial_t^{\ell-1}f(0)\,dz \\
&+\frac{1}{2\pi\mathrm{i}}\int_{\Gamma^\tau_{\theta,\delta}} e^{zt_n}\delta_\tau(e^{-z\tau})^2K(\delta_\tau(e^{-z\tau}))\tau\widetilde{R}_k(e^{-z\tau})\,dz,
\end{aligned}
\tag{3.7}
$$

with $\Gamma^\tau_{\theta,\delta} := \{z\in\Gamma_{\theta,\delta} : |\Im(z)|\le\pi/\tau\}$ *(oriented with an increasing imaginary part), for some* θ *sufficiently close to* $\pi/2$*, where* $\mu(\zeta)$ *and* $\gamma_\ell(\zeta)$ *are defined in* (2.12).

Like before, from the representations (3.6) and (3.7), we arrive at the following algebraic criteria for choosing the coefficients $a_j^{(k)}$, $c_j^{(k)}$ and $b_{\ell,n}^{(k)}$:

$$|\mu(\zeta)-1|\le c|1-\zeta|^k, \tag{3.8}$$

$$\Big|\gamma_1(\zeta)+\sum_{j=1}^{k-1}c_j^{(k)}\zeta^j-\frac{1}{\delta(\zeta)^2}\Big|\le c|1-\zeta|^{k-2}, \tag{3.9}$$

$$
\left|\delta(\zeta)\frac{\gamma_\ell(\zeta)}{\ell!}+\sum_{j=1}^{k-1} b_{\ell,j}^{(k)}\zeta^j-\frac{1}{\delta(\zeta)^\ell}\right|\le c|1-\zeta|^{k-\ell},\quad \ell=1,2,\ldots,k-2, \tag{3.10}
$$

where the functions $\mu(\zeta)$ and $\gamma_\ell(\zeta)$ are defined in (2.12).

By comparing Criterion (3.8) with (2.13), and respectively Criterion (3.9) with (2.14), the coefficients $a_j^{(k)}$ are identical with that for $\alpha \in (0,1)$, and respectively $c_j^{(k)}$ with $b_{1,j}^{(k)}$ for $\alpha \in (0,1)$. However, due to the presence of the extra factor $\delta(\zeta)$, the coefficients $b_{\ell,j}^{(k)}$ are different from that of the case $0<\alpha<1$, and have to be determined. The procedure for computing $b_{\ell,j}^{(k)}$ is similar to that in Section 2.2, and the coefficients $b_{\ell,j}^{(k)}$ are given in Table 3.

TABLE 3
The coefficients $b_{\ell,j}^{(k)}$ according to Criterion (3.10).

order of BDF		$b_{\ell,1}^{(k)}$	$b_{\ell,2}^{(k)}$	$b_{\ell,3}^{(k)}$	$b_{\ell,4}^{(k)}$	$b_{\ell,5}^{(k)}$
$k=3$	$\ell=1$	$\frac{1}{12}$	$-\frac{1}{12}$			
$k=4$	$\ell=1$	$\frac{5}{24}$	$-\frac{1}{3}$	$\frac{1}{8}$		
	$\ell=2$	0	0	0		
$k=5$	$\ell=1$	$\frac{257}{720}$	$-\frac{187}{240}$	$\frac{137}{240}$	$-\frac{107}{240}$	
	$\ell=2$	$\frac{1}{240}$	$-\frac{1}{120}$	$\frac{1}{240}$	0	
	$\ell=3$	$-\frac{1}{720}$	$\frac{1}{720}$	0	0	
$k=6$	$\ell=1$	$\frac{749}{1440}$	$-\frac{1031}{720}$	$\frac{31}{20}$	$-\frac{577}{720}$	$\frac{47}{288}$
	$\ell=2$	$\frac{1}{80}$	$-\frac{1}{30}$	$\frac{7}{240}$	$-\frac{1}{120}$	0
	$\ell=3$	$-\frac{1}{288}$	$\frac{1}{180}$	$-\frac{1}{480}$	0	0
	$\ell=4$	0	0	0	0	0

Last, we state an error estimate for the approximation U^n. The proof is similar to that of Theorem 3.2, but with $g=\partial_t^{-1}f$ in place of f; see Appendix E for a sketch.

THEOREM 3.2. *Let Criteria* (3.8)–(3.10) *hold, and Condition 3.1 be fulfilled, and* $f\in C^{k-2}([0,T];L^2(\Omega))$ *and* $\int_0^t (t-s)^{\alpha-2}\|\partial_s^{k-1}f(s)\|_{L^2(\Omega)}ds<\infty$. *Then for the solution* U^n *to* (3.4)*, the following error estimate holds for any* $t_n>0$

$$
\begin{aligned}
\|U^n-u(t_n)\|_{L^2(\Omega)} \le & c\tau^k\bigg(t_n^{\alpha-k}\|f(0)+Av\|_{L^2(\Omega)}+t_n^{\alpha+1-k}\|Ab\|_{L^2(\Omega)}\\
&+\sum_{\ell=1}^{k-2} t_n^{\alpha+\ell-k}\|\partial_t^\ell f(0)\|_{L^2(\Omega)}+\int_0^{t_n}(t_n-s)^{\alpha-2}\|\partial_s^{k-1}f(s)\|_{L^2(\Omega)}ds\bigg).
\end{aligned}
$$

REMARK 3.2. *Theorem 3.2 only requires* $(k-1)^{\text{th}}$ *derivative of* f *in time, instead of* k^{th} *derivative of* f *as in Theorem 2.2. Thus it relaxes the regularity condition.*

4. Numerical experiments and discussions. Now we present numerical results to show the efficiency and accuracy of the schemes (2.4) and (3.4) in one-spatial dimension, on the unit interval $\Omega=(0,1)$. In space, it is discretized with the piecewise linear Galerkin finite element method [15]: we divide Ω into M equally spaced subintervals with a mesh size $h=1/M$. Since the convergence behavior of the spatial discretization is well understood, we focus on the temporal convergence. In the computation, we fix the time step size τ at $\tau=t/N$, where t is the time of interest. We measure the accuracy by the normalized errors $e^N=\|u(t_N)-U^N\|_{L^2(\Omega)}/\|u(t_N)\|_{L^2(\Omega)}$, where the reference solution $u(t_N)$ is computed using a much finer mesh. All the computations are carried out in MATLAB R2015a on a personal laptop, and further, to observe error beyond double precision, we employ the Multiprecision Computing Toolbox[1] for MATLAB.

4.1. Numerical results for subdiffusion. Consider the following examples:
(a) $v = x(1-x) \in H^2(\Omega) \cap H_0^1(\Omega)$ and $f \equiv 0$;
(b) $v \equiv 0$ and $f(x,t) = \cos(t)(1 + \chi_{(0,1/2)}(x))$.
(c) $v \equiv 0$ and $f(x,t) = 2t^\alpha + \Gamma(1+\alpha)x(1-x)$.

The numerical results for case (a) by the corrected scheme (2.4) are presented in Table 4, where the numbers in the bracket denote the theoretical rate predicted by Theorem 2.2. It converges steadily at an $O(\tau^k)$ rate for all BDFs, which agrees well with the theory, showing clearly its robustness. Surprisingly, the asymptotic convergence of BDF6 kicks in only at a relatively small time step size, at $N = 50$, which contrasts sharply with other BDF schemes. Thus in the preasymptotic regime, BDF5 is preferred over BDF6. To further illustrate Theorem 2.2, in Fig. 2, we plot the numerical solution by BDF5 and its error profile. The solution decays first rapidly and then slowly, resulting in an initial layer. This layer shows clearly the limited temporal regularity of the solution at 0 and as a result, the approximation error near 0 is predominant, partly confirming the prefactor $t_n^{\alpha-k}$ in Theorem 2.2.

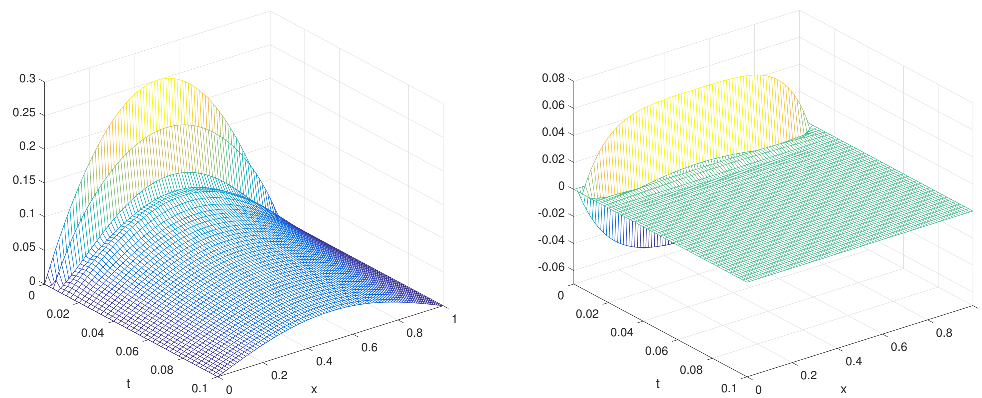
Table 4
The L^2-norm error e^N for case (a) at $t_N = 1$, by the corrected scheme (2.4) with $h = 1/100$.

α	$k\backslash N$	50	100	200	400	800	rate
0.25	2	5.66e-5	1.39e-5	3.46e-6	8.64e-7	2.16e-7	$\approx$ 2.00 (2.00)
	3	2.29e-6	2.76e-7	3.39e-8	4.20e-9	5.23e-10	$\approx$ 3.01 (3.00)
	4	1.42e-7	8.33e-9	5.04e-10	3.10e-11	1.91e-12	$\approx$ 4.02 (4.00)
	5	1.26e-8	3.41e-10	1.01e-11	3.07e-13	9.45e-15	$\approx$ 5.03 (5.00)
	6	1.09e-5	1.60e-9	2.55e-13	3.82e-15	5.83e-17	$\approx$ 6.04 (6.00)
0.5	2	1.74e-4	4.30e-5	1.07e-5	2.65e-6	6.62e-7	$\approx$ 2.00 (2.00)
	3	7.73e-6	9.29e-7	1.14e-7	1.41e-8	1.76e-9	$\approx$ 3.01 (3.00)
	4	5.12e-7	2.98e-8	1.80e-9	1.10e-10	6.83e-12	$\approx$ 4.02 (4.00)
	5	4.75e-8	1.27e-9	3.76e-11	1.14e-12	3.52e-14	$\approx$ 5.03 (5.00)
	6	3.01e-5	2.79e-9	9.85e-13	1.47e-14	2.25e-16	$\approx$ 6.05 (6.00)
0.75	2	4.84e-4	1.19e-4	2.93e-5	7.30e-6	1.82e-6	$\approx$ 2.00 (2.00)
	3	2.55e-5	3.04e-6	3.72e-7	4.60e-8	5.71e-9	$\approx$ 3.01 (3.00)
	4	1.94e-6	1.11e-7	6.68e-9	4.09e-10	2.53e-11	$\approx$ 4.02 (4.00)
	5	2.95e-7	5.30e-9	1.55e-10	4.70e-12	1.45e-13	$\approx$ 5.03 (5.00)
	6	1.67e-3	3.01e-7	4.53e-12	6.61e-14	1.01e-15	$\approx$ 6.07 (6.00)

To illustrate the impact of initial correction, we present in Table 5 the numerical results by the uncorrected BDF scheme (2.3), and two popular finite difference schemes, i.e., L1 [22] and L1-2 [12,27]. The uncorrected BDFk scheme can only achieve an $O(\tau)$ rate, and they all have almost identical accuracy, irrespective of the order k. This low-order convergence is due to the poor approximation at the initial steps, which persists in the numerical solutions at later steps. Meanwhile, for sufficiently smooth solutions, the L1 and L1-2 schemes converge at a rate $O(\tau^{2-\alpha})$ and $O(\tau^{3-\alpha})$, respectively. For general problem data, the L1 scheme converges at an $O(\tau)$ rate [16,19]. The L1 and L1-2 schemes delivers only an $O(\tau)$ rate for case (a), due to insufficient solution regularity. Although not presented, it is noted that the

[1]`http://www.advanpix.com/`, last accessed on January 11, 2017.

numerical results for other fractional orders are similarly. Thus, the correction is necessary in order to retain the desired rate, even for smooth initial data.



(a) numerical solution (b) error profile

FIG. 2. *Numerical solution and error profile for case (a), with* $\alpha = 0.5$*,* $h = 1/100$*,* $\tau = 0.002$ *and BDF5.*

TABLE 5
The L^2*-norm error* e^N *for case (a) at* $t_N = 1$*, by the uncorrected scheme* (2.3) *with* $h = 1/100$*.*

α	N	50	100	200	400	800	rate
	BDF3	4.98e-3	2.48e-3	1.24e-3	6.19e-4	3.09e-4	$\approx$ 1.00 (1.00)
	BDF4	4.97e-3	2.48e-3	1.24e-3	6.19e-4	3.09e-4	$\approx$ 1.00 (1.00)
0.5	BDF5	4.97e-3	2.48e-3	1.24e-3	6.19e-4	3.09e-4	$\approx$ 1.00 (1.00)
	BDF6	4.94e-3	2.48e-3	1.24e-3	6.19e-4	3.09e-4	$\approx$ 1.00 (1.00)
	L1	5.10e-3	2.52e-3	1.25e-3	6.24e-4	3.11e-4	$\approx$ 1.04 (1.00)
	L1-2	3.57e-3	1.73e-3	8.40e-4	4.08e-4	1.99e-4	$\approx$ 1.04 ($--$)

Next we consider the inhomogeneous problem in case (b). Since the source f is smooth in time, Theorem 2.2 is applicable, which predicts an $O(\tau^k)$ rate for the corrected BDFk scheme (2.4). This is fully supported by the numerical results in Table 6. Like before, the uncorrected scheme (2.3) and the L1 and L1-2 schemes can only achieve an $O(\tau)$ rate, despite the smoothness of the problem data, cf. Table 7.

Last, we consider case (c), where the exact solution is given by $u = t^\alpha x(1-x)$. The source $f = 2t^\alpha + \Gamma(1+\alpha)x(1-x)$ is not regular enough in time. It can be verified that $f \in W^{1+\alpha-\epsilon,1}(0,T/2;L^2(\Omega)) \cap W^{1+\alpha-\epsilon,\infty}(T/2,T;L^2(\Omega))$ for any small $\epsilon > 0$. The proof of Theorem 2.2 indicates that the numerical solutions converge at a rate $O(\tau^m)$ at $t = T$ if $f \in W^{m,1}(0,T/2;L^2(\Omega)) \cap W^{m,\infty}(T/2,T;L^2(\Omega))$ for $m \le k$, and the interpolation between the two cases $m = 1$ and $m = 2$ yields a rate $O(\tau^{1+\alpha-\epsilon})$ for this example. The numerical results in Table 8 are consistent with this theoretical prediction, and illustrate the sharpness of Theorem 2.2 with respect to the regularity assumption on f.

4.2. Numerical results for diffusion-wave. Consider the following example:

(d) $v(x) = x(1-x)$, $b(x) = \sin(2\pi x)$ and $f = e^t(1 + \chi_{(0,1/2)}(x))$

For the diffusion-wave model, the scheme (3.4) is only conditionally stable for $\alpha \ge \alpha^*(k) = \pi/(\pi - \vartheta_k)$, with a stability threshold $\tau_0 = (c(\alpha,k)/r(A))^{1/\alpha}$, according to Condition

TABLE 6
The L^2-norm error e^N for case (b) at $t_N = 1$, by the corrected scheme (2.4) with $h = 1/100$.

α	$k\backslash N$	50	100	200	400	800	rate
0.25	2	6.67e-6	1.65e-6	4.10e-7	1.02e-7	2.55e-8	$\approx$ 2.00 (2.00)
	3	2.68e-7	3.20e-8	3.91e-9	4.83e-10	6.00e-11	$\approx$ 3.01 (3.00)
	4	2.14e-8	1.25e-9	7.57e-11	4.65e-12	2.88e-13	$\approx$ 4.02 (4.00)
	5	1.90e-9	5.11e-11	1.51e-12	4.61e-14	1.42e-15	$\approx$ 5.03 (5.00)
	6	1.63e-6	2.40e-10	3.79e-14	5.68e-16	8.67e-18	$\approx$ 6.05 (6.00)
0.5	2	1.76e-5	4.35e-6	1.08e-6	2.70e-7	6.62e-8	$\approx$ 2.00 (2.00)
	3	6.35e-7	7.56e-8	9.22e-9	1.14e-9	1.42e-10	$\approx$ 3.01 (3.00)
	4	5.23e-8	3.03e-9	1.83e-10	1.12e-11	6.95e-13	$\approx$ 4.02 (4.00)
	5	4.94e-9	1.33e-10	3.91e-12	1.19e-13	3.66e-15	$\approx$ 5.03 (5.00)
	6	3.14e-6	2.91e-10	1.02e-13	1.52e-15	2.32e-17	$\approx$ 6.05 (6.00)
0.75	2	3.03e-5	7.47e-6	1.86e-6	4.63e-7	1.16e-7	$\approx$ 2.00 (2.00)
	3	1.10e-6	1.31e-7	1.59e-8	1.96e-9	2.43e-10	$\approx$ 3.01 (3.00)
	4	9.98e-8	5.72e-9	3.43e-10	2.10e-11	1.30e-12	$\approx$ 4.02 (4.00)
	5	1.57e-8	2.81e-10	8.24e-12	2.50e-13	7.68e-15	$\approx$ 5.03 (5.00)
	6	8.95e-5	1.61e-8	2.40e-13	3.50e-15	5.33e-17	$\approx$ 6.07 (6.00)

TABLE 7
The L^2-norm error e^N for case (b) at $t_N = 1$, by the uncorrected scheme (2.3) with $h = 1/100$.

α	N	50	100	200	400	800	rate
0.5	BDF2	5.14e-4	2.57e-4	1.29e-4	6.45e-5	3.22e-5	$\approx$ 1.00 (1.00)
	BDF3	5.19e-4	2.59e-4	1.29e-4	6.45e-5	3.23e-5	$\approx$ 1.00 (1.00)
	BDF4	5.18e-4	2.59e-4	1.29e-4	6.45e-5	3.23e-5	$\approx$ 1.00 (1.00)
	BDF5	5.19e-4	2.59e-4	1.29e-4	6.45e-5	3.23e-5	$\approx$ 1.00 (1.00)
	BDF6	5.15e-4	2.59e-4	1.29e-4	6.45e-5	3.23e-5	$\approx$ 1.00 (1.00)
	L1	5.98e-4	2.86e-4	1.39e-4	6.80e-5	3.35e-5	$\approx$ 1.02 (1.00)
	L1-2	3.71e-4	1.80e-4	8.76e-5	4.25e-5	2.07e-5	$\approx$ 1.04 (−−)

3.1. To illustrate the sharpness of the threshold τ_0 or equivalently the CFL constant $c(\alpha, k)$, we consider case (d) with $k = 5$, $\alpha = 1.5$, $h = 1/M = 1/100$. The eigenvalues of the discrete Laplacian A are available in closed form [15]:

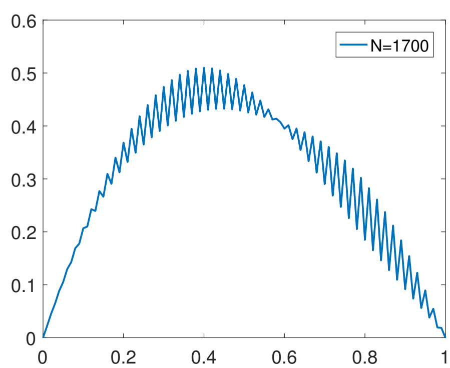
$$\lambda_j^h = \bar{\lambda}_j^h/(1 - \tfrac{h^2}{6}\bar{\lambda}_j^h), \qquad \text{with } \bar{\lambda}_j^h = -\frac{4}{h^2}\sin^2\frac{\pi j}{2(N+1)}, \quad j = 1, 2, \ldots, M-1.$$

Thus the numerical radius $r(A) = \max_j(\lambda_j^h) \approx 1.2 \times 10^5$, and with the value $c(\alpha, k) = 1.58$ from Fig. 1, it gives a stability threshold $\tau_0 \approx 5.60 \times 10^{-4}$. In Fig. 3, we plot the numerical solutions computed by the corrected scheme (3.4) with $N = 1700$ (i.e., $\tau = 5.88 \times 10^{-4}$) and $N = 1800$ (i.e., $\tau = 5.55 \times 10^{-4}$). The scheme (3.4) is unstable for $N = 1700$ but stable for $N = 1800$. This confirms the sharpness of the CFL constant $c(\alpha, k)$ in Condition 3.1. In Table 9, we present the L^2 error for $\alpha > \alpha^*$ and small τ (such that it satisfies the CFL condition). The numerical results indicate the desired $O(\tau^k)$ rate, supporting the theory.

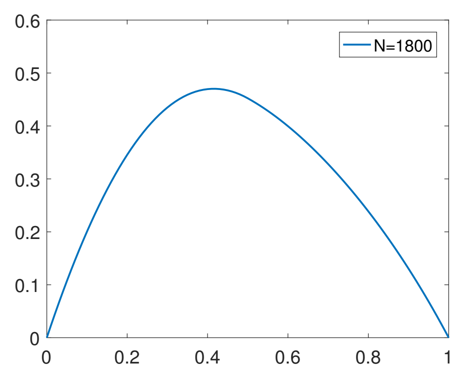
For $\alpha < \alpha^*(k) = \pi/(\pi - \vartheta_k)$, the corrected scheme (3.4) based on BDFk is unconditionally stable. Numerically, the corrected scheme (3.4) converges at an $O(\tau^k)$ rate steadily, cf. Table 10, which agrees well with Theorem 3.2.

TABLE 8
The L^2-norm error e^N for case (c) at $t_N = 1$, by the corrected scheme (2.4) with $h = 1/100$.

α	$k\backslash N$	50	100	200	400	800	rate
0.25	2	4.31e-5	1.81e-5	7.59e-6	3.16e-6	1.30e-6	$\approx$ 1.27 (1.25)
	3	3.19e-5	1.34e-5	5.61e-6	2.33e-6	9.45e-7	$\approx$ 1.28 (1.25)
	4	2.76e-5	1.15e-5	4.81e-6	1.99e-6	8.03e-7	$\approx$ 1.28 (1.25)
	5	2.46e-5	1.03e-5	4.29e-6	1.77e-6	7.11e-7	$\approx$ 1.29 (1.25)
	6	2.31e-5	9.64e-6	4.02e-6	1.65e-6	6.63e-7	$\approx$ 1.29 (1.25)
0.5	2	1.62e-5	5.76e-6	2.04e-6	7.21e-7	2.52e-7	$\approx$ 1.51 (1.50)
	3	9.78e-6	3.48e-6	1.23e-6	4.32e-7	1.50e-7	$\approx$ 1.52 (1.50)
	4	8.03e-6	2.82e-6	9.93e-7	3.48e-7	1.20e-7	$\approx$ 1.53 (1.50)
	5	6.87e-6	2.42e-6	8.51e-7	2.93e-7	1.02e-7	$\approx$ 1.53 (1.50)
	6	6.46e-6	2.22e-6	7.81e-7	2.73e-7	9.33e-8	$\approx$ 1.53 (1.50)
0.75	2	3.37e-6	1.02e-6	3.05e-7	9.15e-8	2.73e-8	$\approx$ 1.74 (1.75)
	3	1.32e-6	4.02e-7	1.21e-7	3.59e-8	1.06e-8	$\approx$ 1.75 (1.75)
	4	1.03e-6	3.04e-7	9.00e-8	2.65e-8	7.76e-9	$\approx$ 1.77 (1.75)
	5	8.37e-7	2.48e-7	7.33e-8	2.16e-8	6.30e-9	$\approx$ 1.77 (1.75)
	6	5.52e-6	2.21e-7	6.59e-8	1.94e-8	5.64e-9	$\approx$ 1.77 (1.75)



(a) $N = 1700$



(b) $N = 1800$

FIG. 3. *The numerical solutions for case (d) at $t = 1$, with $N = 1700$ ($\tau = 5.88 \times 10^{-4}$) and $N = 1800$ ($\tau = 5.56 \times 10^{-4}$), $h = 1/100$. The theoretical stability threshold is $\tau_0 = 5.60 \times 10^{-4}$.*

TABLE 9
The L^2-norm error e^N for case (d) at $t_N = 1$, by the corrected scheme (3.4), with $h = 1/10$.

k (α^*)	$\alpha\backslash N$	100	200	400	800	1600	rate
3 (1.91)	1.95	2.96e-5	3.84e-6	5.00e-7	6.40e-8	8.27e-9	$\approx$ 2.96 (3.00)
4 (1.68)	1.75	2.08e-6	1.43e-7	9.29e-9	5.92e-10	3.74e-11	$\approx$ 3.98 (4.00)
5 (1.40)	1.5	7.29e-8	2.49e-10	6.22e-12	1.72e-13	5.05e-15	$\approx$ 5.14 (5.00)
6 (1.11)	1.5	5.67e-2	2.56e-10	6.88e-13	1.05e-14	1.62e-16	$\approx$ 6.03 (6.00)

Acknowledgements. The authors are grateful to Professor Christian Lubich for his valuable comments on an earlier version of the paper. The authors thank the anonymous

TABLE 10
The L^2-norm error e^N for case (d) at $t_N = 1$, by the corrected scheme (3.4), $h = 1/100$.

k (α^*)	N	100	200	400	800	1600	rate
2 (2.00)	1.25	2.34e-5	5.85e-6	1.46e-6	3.65e-7	9.14e-8	$\approx$ 2.00 (2.00)
	1.5	6.87e-5	1.69e-5	4.18e-6	1.04e-6	2.59e-7	$\approx$ 2.00 (2.00)
	1.75	3.15e-4	8.55e-5	2.21e-5	5.62e-6	1.42e-6	$\approx$ 1.98 (2.00)
3 (1.91)	1.25	1.54e-8	1.66e-9	3.20e-10	4.80e-11	6.33e-12	$\approx$ 3.00 (3.00)
	1.5	4.22e-6	5.12e-7	6.30e-8	7.82e-9	9.74e-10	$\approx$ 3.00 (3.00)
	1.75	5.27e-5	6.43e-6	7.93e-7	9.78e-8	1.15e-8	$\approx$ 3.02 (3.00)
4 (1.68)	1.25	2.74e-8	1.64e-9	1.00e-10	6.20e-12	3.63e-13	$\approx$ 4.00 (4.00)
	1.5	1.88e-7	1.27e-8	8.22e-10	5.19e-11	3.07e-12	$\approx$ 4.00 (4.00)
5 (1.40)	1.1	3.32e-10	9.52e-12	2.85e-13	8.71e-15	2.69e-16	$\approx$ 5.00 (5.00)
	1.3	2.38e-7	1.28e-10	1.08e-12	3.40e-14	1.06e-15	$\approx$ 5.00 (5.00)
6 (1.11)	1.05	3.31e-5	1.94e-7	1.28e-10	7.58e-17	7.39e-19	$\approx$ 6.68 (6.00)

referees for their constructive comments.

Appendix A. An alternative view on the correction scheme (2.4)**.** In this appendix, we discuss the connection between our approach and the one in [25]. The observation of this connection is due to Professor Christian Lubich.

For the following integral and its convolution quadrature approximation

$$u(t) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_{\theta,\delta}} F(z) e^{tz} dz \quad \text{and} \quad U^n = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma^\tau_{\theta,\delta}} F(\delta_\tau(e^{-\tau z})) e^{t_n z} dz, \tag{A.1}$$

Lubich [25, Theorem 2.1] showed the following error estimate away from $t = 0$:

$$|U^n - u(t_n)| \leq c t_n^{\nu-k-1} \tau^k, \tag{A.2}$$

where $\nu \in \mathbb{R}$ is a parameter in the kernel estimate $|\frac{d^m}{dz^m} F(z)| \leq c|z|^{-\nu-m}$, $m \geq 0$. If we choose $F(z) = (z^\alpha - A)^{-1} z^{-\ell-1} \partial_t^\ell f(0)$ in (A.1), then

$$u(t) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma_{\theta,\delta}} (z^\alpha - A)^{-1} z^{-\ell-1} \partial_t^\ell f(0) e^{tz} dz$$

and

$$U^n = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma^\tau_{\theta,\delta}} (\delta_\tau(e^{-\tau z})^\alpha - A)^{-1} \delta_\tau(e^{-\tau z})^{-\ell-1} \partial_t^\ell f(0) e^{t_n z} dz$$

are integral representations of the solutions of

$$\partial_t^\alpha u(t) - Au(t) = \frac{t^\ell}{\ell!} \partial_t^\ell f(0), \qquad \text{with} \quad u(0) = 0, \tag{A.3}$$

$$\bar\partial_\tau^\alpha U^n - AU^n = \tau^\ell \omega_n^{(\ell)} \partial_t^\ell f(0), \qquad \text{with} \quad U^0 = 0, \tag{A.4}$$

respectively, which are the solutions and approximations of (1.1) corresponding to a single component in the splitting (2.5). The weights $\{\omega_n^{(\ell)}\}_{n=0}^\infty$ are generated by the expansion $\delta(\zeta)^{-\ell-1} = \sum_{n=0}^\infty \omega_n^{(\ell)} \zeta^n$. By [25, Theorem 2.1], $\{U^n\}$ has the desired accuracy (A.2). Our scheme (2.4) is connected with (A.4) as follows: we replace $\delta(\zeta)^{-\ell-1}$ by an $O(|\zeta-1|^{k-\ell-1})$ accurate approximation $\frac{\gamma_\ell(\zeta)}{\ell!} + \sum_{j=1}^{k-1} b_{\ell,j}^{(k)} \zeta^j$, cf. (2.14). Our choice of the kernel leads to

$$\bar\partial_\tau^\alpha U^n - AU^n = \frac{t_n^\ell}{\ell!} \partial_t^\ell f(0) + b_{\ell,n}^{(k)} \tau^\ell \partial_t^\ell f(0), \tag{A.5}$$

which differs from (A.4) only in the starting $k-1$ steps, since $b_{\ell,n}^{(k)}=0$ for $n \geq k$. Further, (A.5) is minimal (or optimal) in the sense that it is the unique scheme that only modifies the starting $k-1$ steps while retaining an accuracy of $O(\tau^k)$.

Appendix B. Proof of Theorem 2.1. We need a few estimates on $\delta_\tau(e^{-z\tau})$.

LEMMA B.1. *Let $\alpha \in (0,2)$. For any ε, there exists $\theta_\varepsilon \in (\pi/2,\pi)$ such that for any $\theta \in (\pi/2,\theta_\varepsilon)$, there exist positive constants c, c_1, c_2 (independent of τ) such that*

$$\begin{aligned} &c_1|z| \leq |\delta_\tau(e^{-z\tau})| \leq c_2|z|, \qquad \delta_\tau(e^{-z\tau}) \in \Sigma_{\pi-\vartheta_k+\varepsilon}, \\ &|\delta_\tau(e^{-z\tau}) - z| \leq c\tau^k|z|^{k+1}, \quad |\delta_\tau(e^{-z\tau})^\alpha - z^\alpha| \leq c\tau^k|z|^{k+\alpha}, \quad \forall\, z \in \Gamma_{\theta,\delta}^\tau. \end{aligned}$$

Proof. Since the function $\delta(\zeta)/(1-\zeta)$ has no zero in a neighborhood $\mathcal{N}$ of the unit circle [5, Proof of Lemma 2] and for θ sufficiently close to $\pi/2$, $e^{-z\tau}$ lies in the neighborhood $\mathcal{N}$, there are positive constants c_1' and c_2' such that

$$c_1' \leq \frac{|\delta(e^{-z\tau})|}{|1-e^{-z\tau}|} = \frac{|\delta_\tau(e^{-z\tau})|}{|(1-e^{-z\tau})/\tau|} \leq c_2', \qquad \forall\, z \in \Gamma_{\theta,\delta}^\tau.$$

Since $\widetilde{c}_1|z\tau| \leq |1-e^{-z\tau}| \leq \widetilde{c}_2|z\tau|$ for $z \in \Gamma_{\theta,\delta}^\tau$, the first estimate follows.

When $|\zeta| \leq 1$ and $\zeta \neq 0$, we have $\delta_\tau(\zeta) \in \Sigma_{\pi-\vartheta_k}$ for the $A(\vartheta_k)$ stable BDFk [14]. Hence, by expressing $e^{-z\tau}$ as $e^{-|z|\tau\cos(\theta)}e^{-\mathrm{i}|z|\tau\sin(\theta)}$, we have

$$\begin{aligned} |\delta_\tau(e^{-z\tau}) - \delta_\tau(e^{-\mathrm{i}|z|\tau\sin(\theta)})| &= |\delta_\tau(e^{-|z|\tau\cos(\theta)}e^{-\mathrm{i}|z|\tau\sin(\theta)}) - \delta_\tau(e^{-\mathrm{i}|z|\tau\sin(\theta)})| \\ &\leq ce^{-\sigma|z|\tau\cos(\theta)}\left|\delta_\tau'(e^{-\sigma|z|\tau\cos(\theta)}e^{-\mathrm{i}|z|\tau\sin(\theta)})z\tau\cos(\theta)\right| \end{aligned}$$

for some $\sigma \in (0,1)$, by the mean value theorem. For θ close to $\pi/2$ and $z \in \Gamma_{\theta,\delta}^\tau$, by Taylor expansion, $|z|\tau \leq \pi/\sin\theta$ and the first estimate, we have

$$\tau|\delta_\tau'(e^{-\sigma|z|\tau\cos(\theta)}e^{-\mathrm{i}|z|\tau\sin(\theta)})| \leq c \quad \text{and} \quad |\delta_\tau(e^{-\mathrm{i}|z|\tau\sin(\theta)})| \geq c|z|.$$

Consequently, we deduce

$$\begin{aligned} |\delta_\tau(e^{-|z|\tau\cos(\theta)}e^{-\mathrm{i}|z|\tau\sin(\theta)}) - \delta_\tau(e^{-\mathrm{i}|z|\tau\sin(\theta)})| &\leq c|\cos(\theta)||\delta_\tau(e^{-\mathrm{i}|z|\tau\sin(\theta)})| \\ &\leq c|\theta-\pi/2||\delta_\tau(e^{-\mathrm{i}|z|\tau\sin(\theta)})|. \end{aligned} \tag{B.1}$$

Hence, $\delta_\tau(e^{-\tau z})$ lies in a sector $\Sigma_{\pi-\vartheta_k+c|\theta-\pi/2|}$. If $\theta > \pi/2$ is sufficiently close to $\pi/2$, then $c|\theta-\pi/2| < \varepsilon$. This proves the second assertion.

The third estimate is given in [37, eq. (10.6)]. The last estimate follows from

$$|\delta_\tau(e^{-z\tau})^\alpha - z^\alpha| = \alpha\left|\int_z^{\delta_\tau(e^{-z\tau})}\xi^{\alpha-1}d\xi\right| \leq \max_\xi|\xi|^{\alpha-1}|\delta_\tau(e^{-z\tau}) - z|, \tag{B.2}$$

where ξ lies in the line segment with end points $\delta_\tau(e^{-z\tau})$ and z. Since $\delta(e^{-\mathrm{i}\theta}) > 0$ for $\theta \in (0,\pi)$ (see, e.g., [13, pp. 214–216] or [14, pp. 246] for the plot), it follows from the continuity estimate (B.1) and by choosing θ_ε sufficiently close to $\pi/2$ that $\Im\delta_\tau(e^{-z\tau}) > 0$ for $z \in \Gamma_{\theta,\delta}^\tau$ with $\Im z > 0$, from which and the first estimate we deduce

$$|\xi|^{\alpha-1} \leq \max(|z|^{\alpha-1}, |\delta_\tau(e^{-z\tau})|^{\alpha-1}) \leq c|z|^{\alpha-1}.$$

This inequality and (B.2) yield the last estimate. ∎

Proof of Theorem 2.1. The functions W^n, $n = 1, \ldots, N$, satisfy (with $W^0 = 0$):

$$\bar\partial_\tau^\alpha W^n - AW^n = (1 + a_n^{(k)})(Av + f(0)) + \sum_{\ell=1}^{k-2}\left(\frac{t_n^\ell}{\ell!} + b_{\ell,n}^{(k)}\tau^\ell\right)\partial_t^\ell f(0) + R_k(t_n), \quad 1 \le n \le k-1,$$

$$\bar\partial_\tau^\alpha W^n - AW^n = Av + f(0) + \sum_{\ell=1}^{k-2}\frac{t_n^\ell}{\ell!}\partial_t^\ell f(0) + R_k(t_n), \qquad k \le n \le N.$$

By multiplying both sides by ζ^n, summing over n and collecting terms, we obtain

$$\begin{aligned}
&\sum_{n=1}^\infty \zeta^n \bar\partial_\tau^\alpha W^n - \sum_{n=1}^\infty AW^n\zeta^n \\
&= \left(\sum_{n=1}^\infty \zeta^n + \sum_{j=1}^{k-1} a_j^{(k)}\zeta^j\right)(Av + f(0)) + \sum_{\ell=1}^{k-2}\left(\sum_{n=1}^\infty \frac{t_n^\ell}{\ell!}\zeta^n + \sum_{j=1}^{k-1} b_{\ell,j}^{(k)}\tau^\ell\zeta^j\right)\partial_t^\ell f(0) + \widetilde{R}_k(\zeta) \\
&= \left(\frac{\zeta}{1-\zeta} + \sum_{j=1}^{k-1} a_j^{(k)}\zeta^j\right)(Av + f(0)) + \sum_{\ell=1}^{k-2}\left(\frac{\gamma_\ell(\zeta)}{\ell!} + \sum_{j=1}^{k-1} b_{\ell,j}^{(k)}\zeta^j\right)\tau^\ell\partial_t^\ell f(0) + \widetilde{R}_k(\zeta),
\end{aligned}$$

where $\widetilde{R}_k(\zeta) = \sum_{n=1}^\infty R_k(t_n)\zeta^n$ and we have used elementary identities

$$\text{(B.3)} \qquad \sum_{n=1}^\infty \zeta^n = \frac{\zeta}{1-\zeta} \quad \text{and} \quad \sum_{n=1}^\infty n^\ell\zeta^n = \left(\zeta\frac{d}{d\zeta}\right)^\ell \frac{1}{1-\zeta} := \gamma_\ell(\zeta).$$

Next we simplify the summations on both sides. Since $W^0 = 0$, by the convolution rule, $\sum_{n=1}^\infty \zeta^n\bar\partial_\tau^\alpha W^n = \delta_\tau(\zeta)^\alpha\widetilde{W}(\zeta)$, and consequently, we obtain

$$\widetilde{W}(\zeta) = K(\delta_\tau(\zeta))\left[\tau^{-1}\mu(\zeta)(Av + f(0)) + \sum_{\ell=1}^{k-2}\delta_\tau(\zeta)\left(\frac{\gamma_\ell(\zeta)}{\ell!} + \sum_{j=1}^{k-1} b_{\ell,j}^{(k)}\zeta^j\right)\tau^\ell\partial_t^\ell f(0) + \delta_\tau(\zeta)\widetilde{R}_k(\zeta)\right].$$

where K is given by (2.9), and $\mu(\zeta)$ and $\gamma_\ell(\zeta)$ are given by (2.12). Since $\widetilde{W}(\zeta)$ is analytic with respect to ζ in the unit disk on the complex plane, thus Cauchy's integral formula and the change of variables $\zeta = e^{-z\tau}$ give the following representation for arbitrary $\varrho \in (0,1)$

$$\text{(B.4)} \qquad W^n = \frac{1}{2\pi\mathrm{i}}\int_{|\zeta|=\varrho} \zeta^{-n-1}\widetilde{W}(\zeta)d\zeta = \frac{\tau}{2\pi\mathrm{i}}\int_{\Gamma^\tau} e^{zt_n}\widetilde{W}(e^{-z\tau})\,dz,$$

where Γ^τ is given by $\Gamma^\tau := \{z = -\ln(\varrho)/\tau + \mathrm{i}y : y \in \mathbb{R} \text{ and } |y| \le \pi/\tau\}$. Note that

(1) $\eta(\zeta) := \delta_\tau(\zeta)/(1-\zeta)$ is a polynomial without roots in a neighborhood $\mathcal{N}$ of the unit circle [5]. Thus, $\eta(\zeta)^\alpha$ is analytic in $\mathcal{N}$.
(2) By choosing θ and ϱ sufficiently close to $\pi/2$ and 1, and $0 < \delta < -\ln(\varrho/\tau)$, the function $e^{-\tau z}$ lies in $\mathcal{N}$ for
$$z \in \Sigma_{\theta,\delta}^\tau = \{z \in \Sigma_\theta : |z| \ge \delta,\ |\mathrm{Im}(z)| \le \tau/\pi,\ \mathrm{Re}(z) \le -\ln(\varrho)/\tau\};$$
(3) $(1 - e^{-\tau z})^\alpha$ is analytic for $z \in \mathbb{C}\backslash(-\infty, 0] \supset \Sigma_{\theta,\delta}^\tau$.

Hence, $\delta_\tau(e^{-\tau z})^\alpha = \tau^{-\alpha}(1 - e^{-\tau z})^\alpha\eta(e^{-\tau z})^\alpha$ is analytic for $z \in \Sigma_{\theta,\delta}^\tau$. By choosing ε small enough, Lemma B.1 implies $0 \ne \delta_\tau(e^{-\tau z})^\alpha \in \Sigma_{\alpha(\vartheta_k+\varepsilon)} \subset \Sigma_{\pi-\varepsilon}$ for $z \in \Sigma_{\theta,\delta}^\tau$. Thus $K(\delta_\tau(e^{-\tau z})) = \delta_\tau(e^{-\tau z})^{-1}(\delta_\tau(e^{-\tau z})^\alpha - A)^{-1}$ is analytic for $z \in \Sigma_{\theta,\delta}^\tau$, which is a region enclosed by Γ^τ, $\Gamma_{\theta,\delta}^\tau$ and the two lines $\Gamma_\pm^\tau := \mathbb{R} \pm \mathrm{i}\pi/\tau$ (oriented from left to right). Since the values of $e^{zt_n}\widetilde{W}(e^{-z\tau})$ on $\Gamma_\pm^\tau$ coincide, Cauchy's theorem allows deforming the contour Γ^τ to $\Gamma_{\theta,\delta}^\tau$ in the integral (B.4) to obtain the desired representation.

The regularity assumptions on f is needed to guarantee that the right-hand side is well defined; see Appendix C. ∎

Appendix C. Proof of Theorem 2.2.

The proof of Theorem 2.2 relies on the splitting $u(t_n) - U^n = w(t_n) - W^n$ and the representations (2.8) and (2.11), and then bounding each term using (2.10). The details are given below. First, we give some useful estimates.

LEMMA C.1. *Let Criteria* (2.13) *and* (2.14) *hold. Then for* $z \in \Gamma^\tau_{\theta,\delta}$, *there hold*

$$\|\mu(e^{-z\tau})K(\delta_\tau(e^{-z\tau})) - K(z)\| \le c\tau^k|z|^{k-1-\alpha},$$

$$\left\|(\delta_\tau(e^{-z\tau})^\alpha - A)^{-1}\left(\frac{1}{\ell!}\gamma_\ell(e^{-z\tau}) + \sum_{j=1}^{k-1} b^{(k)}_{\ell,j}e^{-jz\tau}\right)\tau^{\ell+1} - z^{-\ell}K(z)\right\| \le c\tau^k|z|^{k-\ell-1-\alpha}.$$

Proof. Since $|1-e^{-z\tau}| \le c\tau|z|$ for $z \in \Gamma^\tau_{\theta,\delta}$, by Criterion (2.13), there holds $|\mu(e^{-z\tau})-1| \le c|1-e^{-z\tau}|^k \le c\tau^k|z|^k$. Meanwhile, by the triangle inequality, we have

$$\begin{aligned}\|K(\delta_\tau(e^{-z\tau})) - K(z)\| &= \|\delta_\tau(e^{-z\tau})^{-1}(\delta_\tau(e^{-z\tau})^\alpha - A)^{-1} - z^{-1}(z^\alpha - A)^{-1}\| \\ &\le |\delta_\tau(e^{-z\tau})^{-1} - z^{-1}|\|(\delta_\tau(e^{-z\tau})^\alpha - A)^{-1}\| \\ &\quad + |z|^{-1}\|(\delta_\tau(e^{-z\tau})^\alpha - A)^{-1} - (z^\alpha - A)^{-1}\|.\end{aligned}$$

The identity $(\delta_\tau(e^{-z\tau})^\alpha - A)^{-1} - (z^\alpha - A)^{-1} = (z^\alpha - \delta_\tau(e^{-z\tau})^\alpha)(\delta_\tau(e^{-z\tau})^\alpha - A)^{-1}(z^\alpha - A)^{-1}$, Lemma B.1 and the resolvent estimate (2.10) imply directly $\|K(\delta_\tau(e^{-z\tau})) - K(z)\| \le c|\tau|^k|z|^{k-1-\alpha}$. Thus, we obtain the first estimate by

$$\begin{aligned}\|\mu(e^{-z\tau})K(\delta_\tau(e^{-z\tau})) - K(z)\| &\le |\mu(e^{-z\tau}) - 1|\|K(\delta_\tau(e^{-z\tau}))\| \\ &\quad + \|K(\delta_\tau(e^{-z\tau})) - K(z)\| \le c\tau^k|z|^{k-1-\alpha} \quad \forall z \in \Gamma^\tau_{\theta,\delta}.\end{aligned}$$

Next we show the second estimate. By Lemma B.1, there holds

$$|\delta_\tau(e^{-z\tau})^{\ell+1} - z^{\ell+1}| \le c|\delta_\tau(e^{-z\tau}) - z||z|^\ell \le c\tau^k|z|^{k+\ell+1} \quad \forall z \in \Gamma^\tau_{\theta,\delta}.$$

By Criterion (2.14), there holds

$$\left|\frac{\gamma_\ell(e^{-z\tau})}{\ell!} + \sum_{j=1}^{k-1} b^{(k)}_{\ell,j}e^{-jz\tau} - \frac{1}{\delta(e^{-z\tau})^{\ell+1}}\right| \le c\tau^{k-\ell-1}|z|^{k-\ell-1} \quad \forall z \in \Gamma^\tau_{\theta,\delta}.$$

Hence, for any $z \in \Gamma^\tau_{\theta,\delta}$, we have

$$\begin{aligned}&\left\|(\delta_\tau(e^{-z\tau})^\alpha - A)^{-1}\left(\frac{1}{\ell!}\gamma_\ell(e^{-z\tau}) + \sum_{j=1}^{k-1} b^{(k)}_{\ell,j}e^{-jz\tau}\right)\tau^{\ell+1} - z^{-\ell}K(z)\right\| \\ \le&\left\|(\delta_\tau(e^{-z\tau})^\alpha - A)^{-1}\left[\left(\frac{1}{\ell!}\gamma_\ell(e^{-z\tau}) + \sum_{j=1}^{k-1} b^{(k)}_{\ell,j}e^{-jz\tau}\right)\tau^{\ell+1} - \delta_\tau(e^{-z\tau})^{-\ell-1}\right]\right\| \\ &+ \|\delta_\tau(e^{-z\tau})^{-\ell}K(\delta_\tau(e^{-z\tau})) - z^{-\ell}K(z)\| \le c\tau^k|z|^{k-\ell-1-\alpha}.\end{aligned}$$

This completes the proof of the lemma. ∎

Proof of Theorem 2.2. By (2.8) and (2.11), we appeal to the splitting

$$U^n - u(t_n) = W^n - w(t_n) = I_1 + \sum_{\ell=1}^{k-2} I_{2,\ell} - I_3 + I_4,$$

where the terms $I_1, \ldots, I_4$ are given by

$$I_1 = \frac{1}{2\pi\mathrm{i}}\int_{\Gamma^\tau_{\theta,\delta}} e^{zt_n}\Big(\mu(e^{-z\tau})K(\delta_\tau(e^{-z\tau})) - K(z)\Big)(Av + f(0))dz,$$

$$I_{2,\ell} =\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}^\tau} e^{zt_n}\Big[\delta_\tau(e^{-z\tau})\Big(\frac{\gamma_\ell(e^{-z\tau})}{\ell!}+\sum_{j=1}^{k-1} b_{\ell,j}^{(k)}e^{-z\tau j}\Big)\tau^{\ell+1}K(\delta_\tau(e^{-z\tau}))-z^{-\ell}K(z)\Big]\partial_t^\ell f(0)\,dz,$$

$$I_3 =\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}\backslash\Gamma_{\theta,\delta}^\tau} e^{zt_n}K(z)\big(Av+f(0)+\sum_{\ell=1}^{k-2}z^{-\ell}\partial_t^\ell f(0)\big)\,dz,$$

$$I_4 =\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}^\tau} e^{zt_n}(\delta_\tau(e^{-z\tau})^\alpha-A)^{-1}\tau\widetilde{R}_k(e^{-z\tau})dz-\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt_n}(z^\alpha-A)^{-1}\widehat{R}_k(z)dz.$$

It suffices to bound these terms separately. By Lemma C.1, and choosing $\delta = t_n^{-1}$ in the contour $\Gamma_{\theta,\delta}^\tau$, we bound the first term I_1 by

$$\begin{aligned}\|I_1\|_{L^2(\Omega)} &\le c\tau^k\|Av+f(0)\|_{L^2(\Omega)}\Big(\int_\delta^{\pi/(\tau\sin\theta)} e^{rt_n\cos\theta}r^{k-1-\alpha}dr+\int_{-\theta}^{\theta} e^{\delta t_n|\cos\psi|}\delta^{k-\alpha}d\psi\Big)\\ &\le c\tau^k(t_n^{\alpha-k}+\delta^{k-\alpha})\|Av+f(0)\|_{L^2(\Omega)} \le c\tau^k t_n^{\alpha-k}\|Av+f(0)\|_{L^2(\Omega)}.\end{aligned}$$

By Lemma C.1 and choosing $\delta = t_n^{-1}$ in $\Gamma_{\theta,\delta}^\tau$, we bound the terms $I_{2,\ell}$ by

$$\begin{aligned}\|I_{2,\ell}\|_{L^2(\Omega)} &\le c\tau^k\|\partial_t^\ell f(0)\|_{L^2(\Omega)}\Big(\int_\delta^{\pi/(\tau\sin\theta)} e^{rt_n\cos\theta}r^{k-\ell-1-\alpha}dr+\int_{-\theta}^{\theta} e^{\delta t_n|\cos\psi|}\delta^{k-\ell-\alpha}d\psi\Big)\\ &\le c\tau^k t_n^{\alpha+\ell-k}\|\partial_t^\ell f(0)\|_{L^2(\Omega)}, \quad \ell=1,2...,k-1.\end{aligned}$$

Direct computation yields the following estimate on I_3:

$$\|I_3\|_{L^2(\Omega)} \le c\tau^k\Big(t_n^{\alpha-k}\|Av+f(0)\|_{L^2(\Omega)}+\sum_{\ell=1}^{k-2}t_n^{\alpha+\ell-k}\|\partial_t^\ell f(0)\|_{L^2(\Omega)}\Big).$$

The term I_4 is the error of the numerical solution with a compatible right-hand side R_k. Upon recalling the definition of R_k in (2.6), we use the splitting $R_k = \frac{t^{k-1}}{(k-1)!}\partial_t^{k-1}f(0) + \frac{t^{k-1}}{(k-1)!} * \partial_t^k f(t) =: R_k^1 + R_k^2$. Then we have $I_4 = I_4^1 + I_4^2$ with

$$I_4^i = \frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}^\tau} e^{zt_n}(\delta_\tau(e^{-z\tau})^\alpha-A)^{-1}\tau\widetilde{R}_k^i(e^{-z\tau})dz-\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_{\theta,\delta}} e^{zt_n}(z^\alpha-A)^{-1}\widehat{R}_k^i(z)dz.$$

By repeating the preceding argument and (2.17), we have the estimate for I_4^1:

$$\|I_4^1\|_{L^2(\Omega)} \le c\tau^k t_n^{\alpha-1}\|\partial_t^{k-1}f(0)\|_{L^2(\Omega)},$$

and using the argument in [18, Lemma 3.7],

$$\|I_4^2\|_{L^2(\Omega)} \le c\tau^k\int_0^{t_n}(t_n-s)^{\alpha-1}\|\partial_s^k f(s)\|_{L^2(\Omega)}ds.$$

This completes the proof of the theorem. □

Appendix D. Proof of Theorem 3.1. Using the splitting (3.5), the functions W^n, $n = 1, \ldots, N$, satisfy (with $W^0 = 0$):

$$\bar\partial_\tau^\alpha W^n - AW^n = (1+a_n^{(k)})Av+(t_n+\tau c_n^{(k)})Ab+\sum_{\ell=1}^{k-2}\Big(\frac{\bar\partial_\tau t_n^\ell}{\ell!}+b_{\ell,n}^{(k)}\tau^{\ell-1}\Big)\partial_t^{\ell-1}f(0)+\bar\partial_\tau R_k(t_n), \quad 1\le n\le k-1,$$

$$\bar\partial_\tau^\alpha W^n - AW^n = Av+t_nAb+\sum_{\ell=1}^{k-2}\frac{\bar\partial_\tau t_n^\ell}{\ell!}\partial_t^{\ell-1}f(0)+\bar\partial_\tau R_k(t_n), \quad k\le n\le N.$$

By multiplying both sides by ζ^n and summing over n from 1 to ∞, we obtain

$$\sum_{n=1}^{\infty}\zeta^n\bar\partial_\tau^\alpha W^n-\sum_{n=1}^{\infty}AW^n\zeta^n=\Big(\sum_{n=1}^{\infty}\zeta^n+\sum_{j=1}^{k-1}a_j^{(k)}\zeta^j\Big)Av+\Big(\sum_{n=1}^{\infty}\tau n\zeta^n+\sum_{j=1}^{k-1}\tau c_j^{(k)}\zeta^j\Big)Ab$$
$$+\sum_{\ell=1}^{k-2}\Big(\sum_{n=1}^{\infty}\frac{\bar\partial_\tau t_n^\ell}{\ell!}\zeta_n^n+\sum_{j=1}^{k-1}b_{\ell,j}^{(k)}\tau^{\ell-1}\zeta^j\Big)\partial_t^{\ell-1}f(0)+\sum_{n=1}^{\infty}\bar\partial_\tau R_k(t_n)\zeta^n.$$

Using the identities in (B.3), the convolution rule $\sum_{n=1}^{\infty}\zeta^n\bar\partial_\tau^\alpha W^n=\delta_\tau(\zeta)^\alpha\widetilde{W}$, and

$$\sum_{n=1}^{\infty}\zeta^n\bar\partial_\tau\frac{t_n^\ell}{\ell!}=\delta_\tau(\zeta)\sum_{n=0}^{\infty}\frac{t_n^\ell}{\ell!}\zeta^n=\delta(\zeta)\frac{\tau^{\ell-1}}{\ell!}\gamma_\ell(\zeta),$$

we derive

$$\widetilde{W}(\zeta)=K(\delta_\tau(\zeta))\Bigg[\tau^{-1}\mu(\zeta)Av+\delta_\tau(\zeta)\Big(\gamma_1(\zeta)+\sum_{j=1}^{k-1}c_j^{(k)}\zeta^j\Big)\tau Ab$$
$$+\sum_{\ell=1}^{k-2}\delta_\tau(\zeta)\Big(\delta(\zeta)\frac{\gamma_\ell(\zeta)}{\ell!}+\sum_{j=1}^{k-1}b_{\ell,j}^{(k)}\zeta^j\Big)\tau^{\ell-1}\partial_t^\ell g(0)+\delta_\tau(\zeta)^2\widetilde{R}_k(\zeta)\Bigg].$$

Under Condition 3.1 (i), by choosing ε small enough, Lemma B.1 implies $0\neq\delta_\tau(e^{-\tau z})^\alpha\in\Sigma_{\alpha(\vartheta_k+\varepsilon)}\subset\Sigma_{\pi-\varepsilon}$ for $z\in\Sigma_{\theta,\delta}^\tau$. Under Condition 3.1 (ii), we have $\mathrm{dist}(\delta(e^{-z\tau})^\alpha,\tau^\alpha S(A))>0$ (cf. Appendix E), where $S(A)$ denotes the closure of the spectrum of A in the complex plane $\mathbb{C}$. In either case, the operator $K(\delta_\tau(e^{-\tau z}))=\delta_\tau(e^{-\tau z})^{-1}(\delta_\tau(e^{-\tau z})^\alpha-A)^{-1}$ is analytic for $z\in\Sigma_{\theta,\delta}^\tau$, which is the region enclosed by the four curves $\Gamma_{\theta,\delta}^\tau$, $-\ln(\varrho)/\tau+i\mathbb{R}$ and $\mathbb{R}\pm\mathrm{i}\pi/\tau$ (for θ and ϱ sufficiently close to $\pi/2$ and 1, respectively). Then, like in the proof of Theorem 2.1, the assertion follows from Cauchy's integral formula and the change of variables $\zeta=e^{-z\tau}$. □

Appendix E. Proof of Theorem 3.2. Under Condition 3.1(i), Theorem 3.2 can be proved as Theorem 2.2, using (3.6) and (3.7). Under Condition 3.1(ii), it can be proved analogously, if the following resolvent estimate holds:

$$\|(\delta_\tau(e^{-z\tau})^\alpha-A)^{-1}\|\le c|z|^{-\alpha},\quad\forall\, z\in\Gamma_{\theta,\delta}^\tau. \tag{E.1}$$

To prove (E.1), we use the following estimate [31, Theorem 3.9, Chapter 1, pp. 12]:

$$\begin{aligned}\|(\delta_\tau(e^{-z\tau})^\alpha-A)^{-1}\|&=\tau^\alpha\|(\delta(e^{-z\tau})^\alpha-\tau^\alpha A)^{-1}\|\\&\le c\tau^\alpha\,\mathrm{dist}(\delta(e^{-z\tau})^\alpha,\tau^\alpha S(A))^{-1},\quad\forall\, z\in\Gamma_{\theta,\delta}^\tau,\end{aligned} \tag{E.2}$$

where $S(A)$ denotes the closure of the spectrum of A in $\mathbb{C}$. For the discrete Laplacian $A=\Delta_h$, we have $S(A)=[-r(A),0]$. Note that in a small neighborhood of $\theta=0$, simple expansion shows that the contour $\delta(e^{-\mathrm{i}\theta})$ and the segment $[-r(A),0]$ intersects at $\theta=0$ only, with an angle $\pi/2$. Thus the angle between the contour $\delta(e^{-\mathrm{i}\xi\tau})^\alpha$, $\xi\in[-\frac{\pi}{\tau},\frac{\pi}{\tau}]$, and the segment $[-r(A),0]$ in the neighborhood of $\xi\tau=0$ is $(1-\alpha/2)\pi>0$, and it follows that, for small κ,

$$\mathrm{dist}(\delta(e^{-\mathrm{i}\xi\tau})^\alpha,\tau^\alpha S(A))\approx|\delta(e^{-\mathrm{i}\xi\tau})^\alpha|\sin[(1-\alpha/2)\pi]\ge c|\delta(e^{-\mathrm{i}\xi\tau})^\alpha|\quad\text{if }|\xi|\tau\le\kappa.$$

Furthermore, CFL condition 3.1(ii) implies

$$\mathrm{dist}(\delta(e^{-\mathrm{i}\xi\tau})^\alpha,\tau^\alpha S(A))\ge c\ge c|\xi\tau|^\alpha\quad\text{if }\kappa\le|\xi|\tau\le\pi.$$

Let $\Gamma_\theta^\tau=\{z\in\mathbb{C}:\arg(z)=\theta,\ -\frac{\pi}{\tau}\le|z|\sin(\theta)\le\frac{\pi}{\tau}\}$. Then the angle between the contour $\delta(e^{-z\tau})^\alpha$, $z\in\Gamma_\theta^\tau$, and the segment $[-r(A),0]$ is $\pi-\alpha\theta>0$ (if θ is close to $\pi/2$). For small

κ and $z \in \Gamma_\theta^\tau$, $|z|\tau\sin(\theta) \le \kappa$, we have

$$\operatorname{dist}(\delta(e^{-z\tau})^\alpha, \tau^\alpha S(A)) \approx |\delta(e^{-z\tau})^\alpha| \sin(\pi - \alpha\theta) \ge c|\delta(e^{-z\tau})^\alpha| \ge c|z\tau|^\alpha.$$

Estimate (B.1) implies $|\delta(e^{-z\tau}) - \delta(e^{-\mathrm{i}|z|\tau\sin(\theta)})| \le c|\theta - \pi/2|$, and thus

$$\begin{aligned} |\delta(e^{-z\tau})^\alpha - \delta(e^{-\mathrm{i}|z|\tau\sin(\theta)})^\alpha| &\le c|\theta-\pi/2| \min(|\delta(e^{-z\tau})|^{\alpha-1}, |\delta(e^{-\mathrm{i}|z|\tau\sin(\theta)}|^{\alpha-1}) \\ &\le c|\theta - \pi/2||z\tau|^{\alpha-1}, \qquad z \in \Gamma_\theta^\tau. \end{aligned}$$

Hence, if $z \in \Gamma_\theta^\tau$ and $\kappa \le |z|\tau\sin(\theta) \le \pi$, with θ close to $\pi/2$, we have

$$\begin{aligned} \operatorname{dist}(\delta(e^{-z\tau})^\alpha, \tau^\alpha S(A)) &\ge \operatorname{dist}(\delta(e^{-\mathrm{i}|z|\tau\sin(\theta)})^\alpha, \tau^\alpha S(A)) - |\delta(e^{-z\tau})^\alpha - \delta(e^{-\mathrm{i}|z|\tau\sin(\theta)})^\alpha| \\ &\ge c - c|\theta-\pi/2||z\tau|^{\alpha-1} \ge c - c|\theta - \pi/2||z\tau\sin(\theta)|^{\alpha-1} \\ &\ge c - c|\theta - \pi/2|\max(\kappa,\pi)^{\alpha-1} \ge c \ge c|z\tau|^\alpha. \end{aligned}$$

Thus we have $\operatorname{dist}(\delta(e^{-z\tau})^\alpha, \tau^\alpha S(A)) \ge c|z\tau|^\alpha$ for $z \in \Gamma_\theta^\tau$. This inequality and (E.2) yield (E.1) for $z \in \Gamma_{\theta,\delta}^\tau \cap \Gamma_\theta^\tau$. Further, if $z \in \Gamma_{\theta,\delta}^\tau \backslash \Gamma_\theta^\tau$, then $|z| = \delta$ and $-\theta < \arg(z) < \theta$, and Taylor expansion yields $|\delta(e^{-z\tau})|^\alpha \le |z\tau|^\alpha \le \delta^\alpha\tau^\alpha$. By choosing δ small, we have

$$\operatorname{dist}(\delta(e^{-z\tau})^\alpha, \tau^\alpha S(A)) \ge \lambda_{\min}\tau^\alpha - \delta^\alpha\tau^\alpha \ge c\tau^\alpha,$$

where $\lambda_{\min}$ is the smallest positive eigenvalue of the operator A (which can be made independent of h). This and (E.2) yield

$$\|(\delta_\tau(e^{-z\tau})^\alpha - A)^{-1}\| \le c \le c\delta^{-\alpha} = c|z|^{-\alpha}, \quad \forall\, z \in \Gamma_{\theta,\delta}^\tau \backslash \Gamma_\theta^\tau.$$

This completes the proof of (E.1). □

REFERENCES

[1] W. Arendt, C. J. Batty, M. Hieber, and F. Neubrander, *Vector-valued Laplace Transforms and Cauchy Problems*, Birkhäuser, second ed., 2011.

[2] B. Baeumer, M. Kovács, and H. Sankaranarayanan, *Higher order Grünwald approximations of fractional derivatives and fractional powers of operators*, Trans. Amer. Math. Soc., 367 (2015), pp. 813–834.

[3] F. Chen, Q. Xu, and J. S. Hesthaven, *A multi-domain spectral method for time-fractional differential equations*, J. Comput. Phys., 293 (2015), pp. 157–172.

[4] S. Chen, J. Shen, and L.-L. Wang, *Generalized Jacobi functions and their applications to fractional differential equations*, Math. Comp., 85 (2016), pp. 1603–1638.

[5] D. M. Creedon and J. J. H. Miller, *The stability properties of q-step backward difference schemes*, BIT, 15 (1975), pp. 244–249.

[6] E. Cuesta, C. Lubich, and C. Palencia, *Convolution quadrature time discretization of fractional diffusion-wave equations*, Math. Comp., 75 (2006), pp. 673–696.

[7] E. Cuesta and C. Palencia, *A fractional trapezoidal rule for integro-differential equations of fractional order in Banach spaces*, Appl. Numer. Math., 45 (2003), pp. 139–159.

[8] W. Deng and J. S. Hesthaven, *Local discontinuous Galerkin methods for fractional ordinary differential equations*, BIT, 55 (2015), pp. 967–985.

[9] K. Diethelm, N. J. Ford, and A. D. Freed, *Detailed error analysis for a fractional Adams method*, Numer. Algorithms, 36 (2004), pp. 31–52.

[10] A. C. Galucio, J.-F. Deü, S. Mengué, and F. Dubois, *An adaptation of the Gear scheme for fractional derivatives*, Comput. Methods Appl. Mech. Engrg., 195 (2006), pp. 6073–6085.

[11] G.-H. Gao, H.-W. Sun, and Z.-Z. Sun, *Stability and convergence of finite difference schemes for a class of time-fractional sub-diffusion equations based on certain superconvergence*, J. Comput. Phys., 280 (2015), pp. 510–528.

[12] G.-H. Gao, Z.-Z. Sun, and H.-W. Zhang, *A new fractional numerical differentiation formula to approximate the Caputo fractional derivative and its applications*, J. Comput. Phys., 259 (2014), pp. 33–50.

[13] C. W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.

[14] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations. II*, Springer-Verlag, Berlin, sec-

ond ed., 1996. Stiff and differential-algebraic problems.
[15] B. Jin, R. Lazarov, and Z. Zhou, *Error estimates for a semidiscrete finite element method for fractional order parabolic equations*, SIAM J. Numer. Anal., 51 (2013), pp. 445–466.
[16] ———, *An analysis of the L1 scheme for the subdiffusion equation with nonsmooth data*, IMA J. Numer. Anal., 36 (2016), pp. 197–221.
[17] ———, *Two fully discrete schemes for fractional diffusion and diffusion-wave equations with nonsmooth data*, SIAM J. Sci. Comput., 38 (2016), pp. A146–A170.
[18] B. Jin, B. Li, and Z. Zhou, *An analysis of the Crank-Nicolson method for subdiffusion*, IMA J. Numer. Anal., in press (2017, DOI: 10.1093/imanum/drx019).
[19] B. Jin and Z. Zhou, *An analysis of Galerkin proper orthogonal decomposition for subdiffusion*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 89–113.
[20] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and Applications of Fractional Differential Equations*, Elsevier Science B.V., Amsterdam, 2006.
[21] X. Li and C. Xu, *A space-time spectral method for the time fractional diffusion equation*, SIAM J. Numer. Anal., 47 (2009), pp. 2108–2131.
[22] Y. Lin and C. Xu, *Finite difference/spectral approximations for the time-fractional diffusion equation*, J. Comput. Phys., 225 (2007), pp. 1533–1552.
[23] C. Lubich, *Discretized fractional calculus*, SIAM J. Math. Anal., 17 (1986), pp. 704–719.
[24] ———, *Convolution quadrature and discretized operational calculus. I*, Numer. Math., 52 (1988), pp. 129–145.
[25] ———, *Convolution quadrature revisited*, BIT, 44 (2004), pp. 503–514.
[26] C. Lubich, I. H. Sloan, and V. Thomée, *Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term*, Math. Comp., 65 (1996), pp. 1–17.
[27] C. Lv and C. Xu, *Error analysis of a high order method for time-fractional diffusion equations*, SIAM J. Sci. Comput., 38 (2016), pp. A2699–A2724.
[28] K. Mustapha, *Time-stepping discontinuous Galerkin methods for fractional diffusion problems*, Numer. Math., 130 (2015), pp. 497–516.
[29] K. Mustapha, B. Abdallah, and K. M. Furati, *A discontinuous Petrov-Galerkin method for time-fractional diffusion equations*, SIAM J. Numer. Anal., 52 (2014), pp. 2512–2529.
[30] K. Mustapha and D. Schötzau, *Well-posedness of hp-version discontinuous Galerkin methods for fractional diffusion wave equations*, IMA J. Numer. Anal., 34 (2014), pp. 1426–1446.
[31] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.
[32] I. Podlubny, *Fractional Differential Equations*, Academic press, 1998.
[33] K. Sakamoto and M. Yamamoto, *Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems*, J. Math. Anal. Appl., 382 (2011), pp. 426–447.
[34] J. M. Sanz-Serna, *A numerical method for a partial integro-differential equation*, SIAM J. Numer. Anal., 25 (1988), pp. 319–327.
[35] I. M. Sokolov, J. Klafter, and A. Blumen, *Fractional kinetics*, Phys. Today, 55 (2002), pp. 48–54.
[36] E. Sousa, *How to approximate the fractional derivative of order* $1 < \alpha \leq 2$, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 22 (2012), pp. 1250075, 13.
[37] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer-Verlag, Berlin, second ed., 2006.
[38] Z. Wang and S. Vong, *Compact difference schemes for the modified anomalous fractional sub-diffusion equation and the fractional diffusion-wave equation*, J. Comput. Phys., 277 (2014), pp. 1–15.
[39] R. Wu, H. Ding, and C. Li, *Determination of coefficients of high-order schemes for Riemann-Liouville derivative*, The Scientific World Journal, 2014 (2014).
[40] S. B. Yuste and L. Acedo, *An explicit finite difference method and a new von Neumann-type stability analysis for fractional diffusion equations*, SIAM J. Numer. Anal., 42 (2005), pp. 1862–1874.
[41] M. Zayernouri, M. Ainsworth, and G. E. Karniadakis, *A unified Petrov-Galerkin spectral method for fractional pdes*, Comput. Methods Appl. Mech. Engrg., 283 (2015), pp. 1545–1569.
[42] F. Zeng, C. Li, F. Liu, and I. Turner, *The use of finite difference/element approaches for solving the time-fractional subdiffusion equation*, SIAM J. Sci. Comput., 35 (2013), pp. A2976–A3000.