SIAM J. NUMER. ANAL.
Vol. 58, No. 2, pp. 1217–1238

ERROR ANALYSIS FOR A FRACTIONAL-DERIVATIVE PARABOLIC PROBLEM ON QUASI-GRADED MESHES USING BARRIER FUNCTIONS*

NATALIA KOPTEVA† AND XIANGYUN MENG‡

Abstract. An initial-boundary value problem with a Caputo time derivative of fractional order $\alpha \in (0,1)$ is considered, solutions of which typically exhibit a singular behavior at an initial time. For this problem, we give a simple and general numerical-stability analysis using barrier functions, which yields sharp pointwise-in-time error bounds on quasi-graded temporal meshes with arbitrary degree of grading. L1-type and Alikhanov-type discretization in time are considered. In particular, those results imply that milder (compared to the optimal) grading yields optimal convergence rates in positive time. Semidiscretizations in time and full discretizations are addressed. The theoretical findings are illustrated by numerical experiments.

Key words. fractional-order parabolic equation, arbitrary degree of grading, pointwise-in-time error bounds, graded temporal mesh, L1 method, Alikhanov scheme

AMS subject classifications. 65M15, 65M60

DOI. 10.1137/19M1300686

1. Introduction. In this paper we give a simple and general numerical-stability analysis for an initial-boundary value problem with a Caputo time derivative of fractional order $\alpha \in (0,1)$.

- The subtle and sharp stability property (1.2), that we obtain, easily yields sharp pointwise-in-time error bounds for quasi-graded termporal meshes with arbitrary degree of grading. We are not aware of any such general results in the literature.
- In particular, our error bounds accurately predict that milder (compared to the optimal) grading yields optimal convergence rates in positive time. This finding is new and of practical importance.
- The simplicity of our approach is due to the usage of versatile barrier functions, which can be used in the analysis of any discrete fractional-derivative operator that satisfies the discrete maximum principle (or, more generally, is associated with an inverse-monotone matrix).
- Here this approach is employed in the error analysis of the L1 and Alikhanov L2-1_σ fractional-derivative operators, while in [11] it is used in the analysis of an L2-type discretization of order $3-\alpha$ in time. In [12] this methodology is generalized for semilinear fractional parabolic equations.

The Caputo fractional derivative in time, denoted here by D_t^α, is defined [5] by

$$D_t^\alpha u(\cdot,t) := \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha}\, \partial_s u(\cdot,s)\, ds \qquad \text{for } \; 0 < t \le T, \tag{1.1}$$

where $\Gamma(\cdot)$ is the Gamma function, and ∂_s denotes the partial derivative in s.

*Received by the editors November 20, 2019; accepted for publication (in revised form) February 27, 2020; published electronically April 16, 2020.
https://doi.org/10.1137/19M1300686
Funding: The work of the first author was supported by the Science Foundation Ireland grant SFI/12/IA/1683.
†Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland (natalia.kopteva@ul.ie).
‡Department of Mathematics, Beijing Jiaotong University, Beijing 100044, China (xymeng1@bjtu.edu.cn).

Our main stability result is that given an inverse-monotone fractional-derivative operator δ_t^α, associated with a temporal mesh $\{t_j\}_{j=0}^M$ on $[0,T]$ with $\tau := t_1$, and $\gamma \in \mathbb{R}$, under certain conditions on the mesh, the following is true for $\{V^j\}_{j=0}^M$:

$$\left.\begin{array}{c} |\delta_t^\alpha V^j| \lesssim (\tau/t_j)^{\gamma+1} \\ \forall j \geq 1, \quad V^0 = 0 \end{array}\right\} \;\Rightarrow\; |V^j| \lesssim \mathcal{V}^j := \tau t_j^{\alpha-1} \begin{cases} 1 & \text{if } \gamma > 0 \\ 1 + \ln(t_j/\tau) & \text{if } \gamma = 0 \\ (\tau/t_j)^\gamma & \text{if } \gamma < 0 \end{cases} \quad \forall j \geq 1. \tag{1.2}$$

This result is sharp in the sense that it is consistent with the analogous property for the continuous Caputo operator D_t^α; see Remark 1.1. The immediate usefulness of this property is due to the fact that truncation errors in time are typically bounded by negative powers of t_j.

It should be noted that while the explicit inverse of D_t^α is easily available, the proof of (1.2) for any discrete operator is quite nontrivial. As an alternative, discrete Grönwall inequalities were recently employed in the error analysis of L1- and Alikhanov-type schemes [13, 14, 15]. However, this approach involves intricate evaluations and, furthermore, yields less sharp error bounds (see Remarks 3.2 and 4.11 for a more detailed discussion). Our approach is entirely different and is substantially more concise as we obtain (1.2) using clever barrier functions, while the numerical results indicate that our error bounds are sharp in the pointwise-in-time sense.

The following fractional-order parabolic problem is considered:

$$\begin{aligned} &D_t^\alpha u + \mathcal{L}u = f(x,t) \quad \text{for } (x,t) \in \Omega \times (0,T], \\ &u(x,t) = 0 \quad \text{for } (x,t) \in \partial\Omega \times (0,T], \qquad u(x,0) = u_0(x) \quad \text{for } x \in \Omega. \end{aligned} \tag{1.3}$$

This problem is posed in a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ (where $d \in \{1,2,3\}$). The spatial operator $\mathcal{L}$ here is a linear second-order elliptic operator,

$$\mathcal{L}u := \sum_{k=1}^d \Big\{ -\partial_{x_k}(a_k(x,t)\,\partial_{x_k}u) + b_k(x,t)\,\partial_{x_k}u \Big\} + c(x,t)\,u, \tag{1.4}$$

with sufficiently smooth coefficients $\{a_k\}$, $\{b_k\}$ and c in $C(\bar\Omega)$, for which we assume that $a_k > 0$ in $\bar\Omega$, and also either $c \geq 0$ or $c - \frac{1}{2}\sum_{k=1}^d \partial_{x_k} b_k \geq 0$.

The first part of the paper is devoted to L1-type schemes for problem (1.3), which employ the discetization of $D_t^\alpha u$ defined, for $m = 1, \ldots, M$, by

$$\delta_t^\alpha U^m := \frac{1}{\Gamma(1-\alpha)} \sum_{j=1}^m \delta_t U^j \int_{t_{j-1}}^{t_j} (t_m - s)^{-\alpha}\, ds, \qquad \delta_t U^j := \frac{U^j - U^{j-1}}{t_j - t_{j-1}}, \tag{1.5}$$

when associated with the temporal mesh $0 = t_0 < t_1 < \cdots < t_M = T$ on $[0,T]$. The generality of our approach is demonstrated in the second part of the paper by extending the stability and error analysis to higher-order Alikhanov-type schemes [1].

Similarly to [2, 3, 4, 10, 13, 14, 16, 17, 20], our main interest will be in graded temporal meshes as they offer an efficient way of computing reliable numerical approximations of solutions singular at $t = 0$, which is typical for (1.3). In particular, [4, 10, 17, 20] give global-in-time error bounds on graded meshes for problems of type (1.3) for the L1 method [20, 10], the Alikhanov method [4], and a high-order

Petrov–Galerkin method in time [17]. There is also a lot of interest in the literature in optimal error bounds in positive time on uniform meshes; see, e.g., [6, 8, 9, 10].

- By contrast, here, as well as in the related paper [11], pointwise-in-time error bounds will be obtained, while an arbitrary degree of mesh grading (with uniform meshes included as a particular case) is allowed.
- More general temporal meshes, which may be viewed as obtained by adding new nodes in an arbitrary manner to a quasi-graded mesh, are also considered; see section 2.4 and Theorems 3.1 and 3.6.
- For both considered discretizations, when the optimal grading parameter $r = (p-\alpha)/\alpha$ is used, we recover the optimal global convergence rates of $p-\alpha$, where $p = 2$ for the L1 scheme and $p = 3$ for the Alikhanov scheme, as particular cases of our more general error bounds; see Remarks 3.3 and 4.6.
- Another straightforward particular case of our error bounds indicates that the optimal convergence rates $p - \alpha$ in positive time $t \gtrsim 1$ are attained using much milder grading with $r > p - \alpha$; see Remarks 3.2 and 4.5. The accuracy of these threshold values is demonstrated by the numerical results in sections 5.3–5.4.
- When dealing with the fractional parabolic case, for L1-type schemes we follow [10], while for Alikhanov-type schemes, our approach substantially differs from [1, 4] (as we aim at pointwise-in-time error bounds), so we build on some ideas from [11], which may be of independent interest.
- In the latter case, a much milder grading with $r = 2$ (compared to the optimal $r = 2/\alpha$) yields the optimal convergence order 2; see Remark 4.11.

Throughout the paper, it is assumed that there exists a unique solution of this problem such that $\|\partial_t^l u(\cdot,t)\|_{L_2(\Omega)} \lesssim 1 + t^{\alpha-l}$ for $l \le 3$. This is a realistic assumption, satisfied by typical solutions of (1.3) (see, e.g., [18], [20, section 2], [10, section 6]), in contrast to stronger assumptions of type $\|\partial^l u(\cdot,t)\|_{L_2(\Omega)} \lesssim 1$ frequently made in the literature (see, e.g., references in [7, Table 1.1]). Indeed, [19, Theorem 2.1] shows that if a solution u of (1.3) is less singular than we assume, then the initial condition u_0 is uniquely defined by the other data of the problem, which is clearly too restrictive. At the same time, our results can be easily applied to the case of u having no singularities or exhibiting a somewhat different singular behavior at $t = 0$.

Remark 1.1. The stability result (1.2) is sharp in the sense that it is consistent with the analogous property for the continuous Caputo operator D_t^α. Indeed, a calculation shows that if $v(0) = 0$ and $D_t^\alpha v(t) = F(t) := \min\{1, (\tau/t)^{\gamma+1}\}$ for $t > 0$, then the explicit representation $v(t) = J_t^\alpha F(t) = \{\Gamma(\alpha)\}^{-1}\int_0^t (t-s)^{\alpha-1} F(s)\, ds$ yields $v(t) \simeq \mathcal{V}(t)$ for $t \ge \tau$, where $\mathcal{V}(t)$ is a continuous version of $\mathcal{V}^j$ from (1.2).

Outline. Section 2 is devoted to the proof of the stability result (1.2) for the L1 discrete fractional-derivative operator. This result is then employed in section 3 to obtain pointwise-in-time error bounds for L1-type discretizations of the initial-value problem in section 3.1, as well as semidiscretizations and full discretizations of the initial-boundary-value problems in sections 3.2 and 3.3. The above error analysis is extended to the Alikhanov-type discretizations in section 4. Finally, our theoretical findings are illustrated by numerical experiments in section 5.

Notation. We write $a \simeq b$ when $a \lesssim b$ and $a \gtrsim b$, and $a \lesssim b$ when $a \le Cb$ with a generic constant C depending on Ω, T, u_0, and f, but not on the total numbers of degrees of freedom in space or time. Also, for $1 \le p \le \infty$, and $k \ge 0$, we shall use the standard norms in the spaces $L_p(\Omega)$ and the related Sobolev spaces $W_p^k(\Omega)$, while $H_0^1(\Omega)$ is the standard space of functions in $W_2^1(\Omega)$ vanishing on $\partial\Omega$.

2. Stability properties of the L1 discrete fractional-derivative operator.

2.1. Quasi-graded temporal meshes. Main stability result. Throughout the paper, we shall frequently assume that the temporal mesh is quasi-graded in the sense that, with some $r \ge 1$,

$$\tau := t_1 \simeq M^{-r}, \quad t_j \simeq \tau j^r, \quad \tau_j := t_j - t_{j-1} \simeq \tau^{1/r} t_j^{1-1/r} \quad \forall\, j = 1, \ldots, M. \tag{2.1}$$

For example, the standard graded temporal mesh $\{t_j = T(j/M)^r\}_{j=0}^M$ with some $r \ge 1$ (while $r = 1$ generates a uniform mesh) satisfies (2.1), in view of $\tau_j \simeq M^{-1}\, t_{j-1}^{1-1/r}$ and $t_j \le 2^r t_{j-1}$ for $j \ge 2$.

Furthermore, our results also apply to more general meshes that may be viewed as obtained by adding new nodes to any mesh of type (2.1); see section 2.4.

The key in our error analysis for L1-type discretizations is the following stability property.

Theorem 2.1 (stability). (i) *Let the temporal mesh satisfy* (2.1) *with* $1 \le r \le (2-\alpha)/\alpha$. *Given* $\gamma \in \mathbb{R}$ *and* $\{V^j\}_{j=0}^M$, *the stability property* (1.2) *holds true.*

(ii) *If* $\gamma \le \alpha - 1$, *then* (1.2) *holds true on an arbitrary temporal mesh* $\{t_j\}_{j=0}^M$.

(iii) *The above results remain valid if* $|\delta_t^\alpha V^j| \lesssim (\tau/t_j)^{\gamma+1}$ *in* (1.2) *is replaced by* $\delta_t^\alpha |V^j| \lesssim (\tau/t_j)^{\gamma+1}$.

Proof. (i) It suffices to prove part (i) only for $\gamma \ge \alpha - 1$ (as the result of part (ii) applies to the case $\gamma \le \alpha - 1$). The proof is presented in sections 2.2 and 2.3, where a few cases are considered separately.

(ii) This result is easily obtained from [10, Lemma 2.1(i)]. The latter implies that $|V^m| \lesssim \max_{j\le m}\{t_j^\alpha |\delta_t^\alpha V^j|\}$ on an arbitrary mesh. The assumptions on $\{V^j\}$ yield $t_j^\alpha |\delta_t^\alpha V^j| \lesssim t_j^\alpha (\tau/t_j)^{\gamma+1} = \tau^{\gamma+1} t_j^{\alpha-\gamma-1}$, which, combined with $\gamma \le \alpha - 1$, implies $t_j^\alpha |\delta_t^\alpha V^j| \lesssim \tau^{\gamma+1} t_m^{\alpha-\gamma-1}$ $\forall\, j \le m$. The desired assertion $|V^m| \lesssim \tau^{\gamma+1} t_m^{\alpha-\gamma-1}$ follows.

(iii) Imitate the proof of [10, Lemma 2.1(ii)]. To be more precise, let $W^0 = 0$ and $\delta_t^\alpha W^j = \max\{0, \delta_t^\alpha |V^j|\} \ge \delta_t^\alpha |V^j|$ $\forall\, j \ge 1$. Then $0 \le |V^j| \le W^j$ $\forall\, j \ge 1$ (as δ_t^α is associated with an M-matrix), while the results of parts (i) and (ii) apply to $\{W^j\}$. ∎

Remark 2.2. To a degree, the proof of Theorem 2.1 builds on the analysis in [10, Appendix A] for uniform grids, but now we address considerably more general meshes.

2.2. Proof of Theorem 2.1(i) for $\gamma \ge \alpha$. In this case, $(\tau/t_j)^{\gamma+1} \le (\tau/t_j)^{\alpha+1}$, so it suffices to consider only $\gamma = \alpha$. For the latter case, as the operator δ_t^α is associated with an M-matrix, it suffices to prove the following lemma.

Lemma 2.3. *Let the temporal mesh satisfy* (2.1) *with* $1 \le r \le (2-\alpha)/\alpha$. *Then there exists a discrete barrier function* $\{B^j\}_{j=0}^M$ *such that*

$$B^0 = 0, \qquad 0 \le B^j \lesssim t_j^{\alpha-1}, \qquad \delta_t^\alpha B^j \gtrsim \tau^\alpha t_j^{-\alpha-1} \qquad \forall j \ge 1. \tag{2.2}$$

Proof. Fix a sufficiently large number $2 \le p \lesssim 1$, and then set

$$\beta := 1 - \alpha, \qquad B(s) := \min\{(s/t_p) t_p^{-\beta}, s^{-\beta}\}, \qquad B^j := B(t_j). \tag{2.3}$$

Note that, when using the notation of type $\lesssim$, the dependence on p will be shown explicitly.

For $j \le p$, a straightforward calculation shows that $\delta_t^\alpha B^j = D_t^\alpha B(t_j) \simeq t_j^\beta t_p^{-\beta-1} \simeq p^{-r(2-\alpha)} t_j^\beta \tau^{-\beta-1}$, where we also used $t_p \simeq \tau p^r$ (in view of (2.1)). As $t_j \ge \tau$, we then get $\delta_t^\alpha B^j \gtrsim p^{-r(2-\alpha)} \tau^{-\beta} t_j^{\beta-1} \ge p^{-r(2-\alpha)} \tau^\gamma t_j^{-\gamma-1}$ $\forall \gamma \ge \alpha - 1$ including $\gamma = \alpha$.

Next, for $D_t^\alpha B(t)$ with $t > t_p$ one has

$$\Gamma(1-\alpha)\, D_t^\alpha B(t) = \underbrace{\int_0^{t_p} t_p^{-\beta-1}(t-s)^{-\alpha}\, ds}_{\geq t_p^{-\beta} t^{-\alpha}} - \underbrace{\beta \int_{t_p}^{t} s^{-\beta-1}(t-s)^{-\alpha}\, ds}_{=: t^{-1} I}\,.$$

Here, using $\hat{s} := s/t$ and $\hat{t}_p := t_p/t$, and noting that $\alpha + \beta = 1$, one gets

$$I = \beta \int_{\hat{t}_p}^{1} \hat{s}^{-\beta-1}(1-\hat{s})^{-\alpha}\, d\hat{s} = \hat{t}_p^{-\beta}(1-\hat{t}_p)^{\beta} \leq \hat{t}_p^{-\beta}(1-\beta\hat{t}_p).$$

Now, using $t^{-1}\hat{t}_p^{-\beta} = t_p^{-\beta} t^{-\alpha}$, one concludes that

$$\Gamma(1-\alpha)\, D_t^\alpha B(t) \geq t_p^{-\beta} t^{-\alpha}\,(\beta t_p/t) = \beta t_p^{\alpha} t^{-\alpha-1} \qquad \text{for } \ t > t_p. \tag{2.4}$$

So, to complete the proof, it remains to show that $|\delta_t^\alpha B^m - D_t^\alpha B(t_m)| \leq \frac{1}{2} D_t^\alpha B(t_m)$ for any $m > p$.

For the latter, note that $\Gamma(1-\alpha)[\delta_t^\alpha B^m - D_t^\alpha B(t_m)] = \sum_{j=1}^m \mu^j$, where, using the standard piecewise-linear interpolant B^I of B,

$$\mu^j := \int_{t_{j-1}}^{t_j} (B^I - B)'(s)\,(t_m - s)^{-\alpha} ds = \alpha \int_{t_{j-1}}^{t_j} (B - B^I)(s)\,(t_m - s)^{-\alpha-1} ds. \tag{2.5}$$

Clearly, $\mu^j = 0$ for $j \leq p$. For $p+1 \leq j \leq m-1$, one gets $|B - B^I| \lesssim \tau_j^2 |B''(t_{j-1})|$. For $j = m$, we shall use a similar but sharper bound $|B - B^I| \lesssim \tau_j (t_m - s)|B''(t_{j-1})|$. Combining these yields $|B - B^I| \leq \tau_j^2 \min\{1, (t_m - s)/\tau_m\} |B''(t_{j-1})|$ for $j > p$, where $|B''(t_{j-1})| \lesssim |B''(s)| \simeq s^{-\beta-2}$ (in view of $t_{j-1} \simeq t_j$). Also noting that, in view of (2.1), $\tau_j \simeq \tau^{1/r} t_j^{1-1/r} \simeq \tau^{1/r} s^{1-1/r}$, we arrive at

$$|\mu^j| \lesssim \tau^{2/r} \int_{t_{j-1}}^{t_j} s^{-\beta-2/r}\,(t_m - s)^{-\alpha-1}\, \min\{1, (t_m - s)/\tau_m\}\, ds \qquad \forall\, p > m.$$

This immediately yields the bound

$$|\delta_t^\alpha B^m - D_t^\alpha B(t_m)| \lesssim \tau^{2/r} \int_{t_p}^{t_m} s^{-\beta-2/r}\,(t_m - s)^{-\alpha-1}\, \min\{1, (t_m - s)/\tau_m\}\, ds. \tag{2.6}$$

For the latter, using the substitution $s = t_m \hat{s}$ and the notation $\hat{t}_j := t_j/t_m$, $\hat{\tau}_j := \tau_j/t_m$, one gets

$$|\delta_t^\alpha B^m - D_t^\alpha B(t_m)| \lesssim \tau^{2/r} t_m^{-2/r-1} \underbrace{\int_{\hat{t}_p}^{1} \hat{s}^{-\beta-2/r}(1-\hat{s})^{-\alpha-1} \min\{1, (1-\hat{s})/\hat{\tau}_m\}\, d\hat{s}}_{\lesssim \hat{t}_p^{\alpha-2/r} + \hat{\tau}_m^{-\alpha}}\,.$$

Here, when bounding the integral, it is convenient to separately consider the intervals $(\hat{t}_p, \max\{\frac{1}{2}, \hat{t}_p\})$, $(\max\{\frac{1}{2}, \hat{t}_p\}, 1-\hat{\tau}_m)$, and $(1-\hat{\tau}_m, 1)$, where $\hat{\tau}_m \leq \frac{1}{2}$ if p is sufficiently large (as, in view of (2.1), $\hat{\tau}_m \simeq 1/m \leq 1/p$). On these three intervals, the integrand is, respectively, $\lesssim s^{-\beta-2/r}$, $\lesssim (t_m - s)^{-\alpha-1}$ and $\lesssim (t_m - s)^{-\alpha}/\tau_m$, so the corresponding integrals are, respectively, $\lesssim \hat{t}_p^{\alpha-2/r}$, $\lesssim \hat{\tau}_m^{-\alpha}$, and $\lesssim \hat{\tau}_m^{-\alpha}$. Finally, note that

$\hat\tau_m = \tau_m/t_m \simeq (\tau/t_m)^{1/r}$, while, in view of $r \le (2-\alpha)/\alpha$, one has $(\tau/t_m)^{(2-\alpha)/r} \lesssim (\tau/t_m)^\alpha$. Now, a calculation shows that

$$\begin{aligned}
(2.7)\qquad |\delta_t^\alpha B^m - D_t^\alpha B(t_m)| &\lesssim \tau^{2/r} t_m^{-2/r-1}\Big[(t_p/t_m)^{\alpha-2/r} + (\tau/t_m)^{-\alpha/r}\Big]\\
&\lesssim (\tau/t_p)^{2/r} t_p^\alpha t_m^{-\alpha-1} + \underbrace{t_m^{-1}(\tau/t_m)^{(2-\alpha)/r}}_{\lesssim \tau^\alpha t_m^{-\alpha-1}}\\
&\lesssim \Big[(\tau/t_p)^{2/r} + (\tau/t_p)^\alpha\Big]\, t_p^\alpha t_m^{-\alpha-1}.
\end{aligned}$$

Combining this with (2.4) and choosing p sufficiently large yields the desired assertion $\delta_t^\alpha B^m \gtrsim t_p^\alpha t_m^{-\alpha-1}$ $\forall\, m > p$, and hence $\forall\, m \ge 1$. ∎

COROLLARY 2.4. *Lemma* 2.3 *remains valid if the temporal mesh is obtained by adding new nodes to any mesh of type* (2.1) *under the condition that the first mesh interval remains unchanged.*

Proof. Suppose the temporal mesh $\{t'_k\}$ is obtained by refining the mesh $\{t_j\}_{j=0}^M$ of type (2.1). For $t'_k \le t_p$, it is essential that $t'_1 = t_1 = \tau$, so the desired result is obtained exactly as in the proof of Lemma 2.3. For $t'_k > t_p$, the desired result is obtained by combining (2.4) with the bound $|D_t^\alpha(B^I - B)| \le \frac12 D_t^\alpha B$ at any $t'_k > t_p$, where B^I denotes the piecewise-linear interpolant on the new mesh $\{t'_k\}$. If $t'_k = t_m$ for some $m > p$, we again proceed exactly as in the proof of Lemma 2.3, as the same bounds on $B - B^I$ hold true (even though B^I is now the interpolant on a finer mesh). If $t'_k \in (t_m, t_{m+1})$ for $m \ge p$, then on (t_{m-1}, t'_k) one employs $|B - B^I| \lesssim \tau_m(t'_k - s)|B''(t_{m-1})| \lesssim \tau_m^2 \min\{1,\, (t'_k - s)/\tau_m\}|B''(t_{m-1})|$. Hence, one gets a version of (2.6) with t_m replaced by t'_k, which (in view of $t'_k \simeq t_m$) leads to the desired version of (2.7) at t'_k. ∎

2.3. Proof of Theorem 2.1(i) for $\gamma < \alpha$. We shall use the notation and some findings from the proof of Lemma 2.3. In particular, $\beta = 1-\alpha$, while $p \simeq 1$ was chosen sufficiently large in the proof of Lemma 2.3. When using the notation of type $\lesssim$, the dependence on γ and m, but not on p, will be shown explicitly.

For $m \ge 0$ and $\gamma < \alpha$, set

$$(2.8)\qquad p_m := 2^m p, \qquad B_m^j := \min\{t_j t_{p_m}^{-\beta-1}, t_j^{-\beta}\}, \qquad c_m := 2^{-m\gamma r} \;\Rightarrow\; c_m t_{p_m}^\gamma \simeq \tau^\gamma.$$

Here the final observation follows from (2.1) (which yields $t_{p_m} \simeq \tau p_m^r$).

Note that $B_0^j = B_j$, and, more generally, $B_m^j = B_j\big|_{p:=p_m}$, where B^j is from (2.3). Conveniently, in the proof of Lemma 2.3, the dependance on any sufficiently large p was shown explicitly. In particular, we recall that $\delta_t^\alpha B_m^j \ge 0$ for $j \ge 0$. Furthermore,

$$(2.9)\qquad \delta_t^\alpha B_0^j \gtrsim \tau^\gamma t_j^{-\gamma-1} \ \text{ for } 1 \le j \le p_0, \qquad c_m(\delta_t^\alpha B_m^j) \gtrsim \tau^\gamma t_j^{-\gamma-1} \ \text{ for } p_m < j \le p_{m+1}.$$

The first relation for $B_0^j = B^j$ can be found in the abovementioned proof for $\gamma \ge \alpha - 1$ (but is, in fact, valid for any fixed γ now that the dependence on p is inessential). The second relation in (2.9) follows from the bound of type (2.4) also obtained there: $\delta_t^\alpha B_m^j \gtrsim t_{p_m}^\alpha t_j^{-\alpha-1}$. The latter, indeed, implies $c_m(\delta_t^\alpha B_m^j) \gtrsim c_m t_{p_m}^\gamma t_j^{-\gamma-1} \simeq \tau^\gamma t_j^{-\gamma-1}$ for $p_m < j \le p_{m+1}$ (also using the final bound in (2.8)).

Now we are prepared to prove the following two lemmas, which are sufficient for establishing Theorem 2.1(i) for $\gamma \in (0, \alpha)$ and $\gamma \le 0$, respectively.

LEMMA 2.5. *Under the conditions of Theorem* 2.1(i), *suppose that* $\gamma \in (0, \alpha)$. *Then there exists a discrete barrier function* $\{\bar{B}^j\}_{j=0}^M$ *such that* $\bar{B}^0 = 0$, *while* $0 \le \bar{B}^j \lesssim t_j^{\alpha-1}$ *and* $\delta_t^\alpha \bar{B}^j \gtrsim \tau^\gamma t_j^{-\gamma-1}$ *for* $j \ge 1$.

Proof. Using (2.8), let $\bar{B}^j := \sum_{m=0}^\infty c_m B_m^j$. Then $\delta_t^\alpha \bar{B}^j \gtrsim \tau^\gamma t_j^{-\gamma-1}$ $\forall j \ge 1$ follows from (2.9), while $\sum_{m=0}^\infty c_m = C_\gamma := (1-2^{-\gamma r})^{-1}$, so $\bar{B}^j \le C_\gamma t_j^{-\beta} = C_\gamma t_j^{\alpha-1}$, which completes the proof. ∎

LEMMA 2.6. *Under the conditions of Theorem* 2.1(i), *suppose that* $\gamma \in [\alpha-1, 0]$. *If* $V^0 = 0$ *and* $|\delta_t^\alpha V^j| \lesssim \tau^\gamma t_j^{-\gamma-1}$ *for* $j = 1, \ldots, n \le M$, *then* $|V^n| \lesssim t_n^{\alpha-1}[1+\ln(t_n/\tau)]$ *if* $\gamma = 0$, *and* $|V^n| \lesssim t_n^{\alpha-1}(\tau/t_n)^\gamma$ *if* $\gamma \le 0$.

Proof. Using (2.8), let $\bar{B}^j := \sum_{m=0}^N c_m B_m^j$, where $N = 0$ if $n \le p$, and $N := \lceil \log_2(n/p) - 1 \rceil$ otherwise, so that $p_N < n \le p_{N+1}$. Note also that $N \lesssim \ln n \simeq \ln(t_n/\tau)$ (as $t_n/\tau \simeq n^r$ in view of (2.1)). Then $\delta_t^\alpha \bar{B}^j \gtrsim \tau^\gamma t_j^{-\gamma-1}$ for $1 \le j \le n$ follows from (2.9). Hence $|V^j| \lesssim \bar{B}^j$ $\forall j \le n$, in particular $|V^n| \lesssim \bar{B}^n$.

On the other hand, $\bar{B}^n \le t_n^{\alpha-1} \sum_{m=0}^N c_m$. When $\gamma = 0$, each $c_m = 1$, so $\sum_{m=0}^N c_m = 1 + N \simeq 1 + \ln(t_n/\tau)$, so, indeed, $\bar{B}^n \le t_n^{\alpha-1}[1+\ln(t_n/\tau)]$. When $\gamma \in [\alpha-1, 0)$, we get $\sum_{m=0}^N c_m = (c_{N+1}-1)/(c_1-1)$, where $c_{N+1} \simeq (\tau/t_{p_{N+1}})^\gamma \simeq (\tau/t_n)^\gamma$, while $C_\gamma := (c_1-1)^{-1} = (2^{|\gamma| r}-1)^{-1}$, so finally $|V^n| \lesssim \bar{B}^n \lesssim C_\gamma t_n^{\alpha-1}(\tau/t_n)^\gamma$. ∎

2.4. More general temporal meshes. Our main stability result, Theorem 2.1, remains valid for more general temporal meshes that may be viewed as obtained by adding new nodes to any mesh of type (2.1) under the condition that the first mesh interval remains unchanged. Indeed, an inspection of the proof of Corollary 2.4 reveals that not only Lemma 2.3 but also the results of Section 2.3 are valid for the above temporal mesh. Such more general meshes may be useful if the solution exhibits additional singularities away from the initial time.

LEMMA 2.7. *Theorem* 2.1 *remains valid if the temporal mesh satisfies the following weaker version of* (2.1)*:*

$$\tau := t_1 \simeq M^{-r}, \qquad \tau_j := t_j - t_{j-1} \lesssim \tau^{1/r} t_j^{1-1/r} \qquad \forall\, j = 1, \ldots, M. \tag{2.10}$$

Proof. It suffices to construct a submesh $\{t'_k\} \subset \{t_j\}$ that satisfies (2.1). Let $t'_1 := t_1$ and $t'_k := \min\{t_j : t_j \ge (Ck/M)^r\}$ (with an obvious modification near $t = T$). Here the constant C is chosen sufficiently large to ensure that $\{t_j\}$ is sufficiently dense within $\{t'_k\}$. To be more precise, whenever $t_j = t'_k$ one has, in view of (2.10), $\tau_j/t_j \lesssim (\tau/t'_k)^{1/r} \lesssim C^{-1}k^{-1}$. Hence t'_k is sufficiently close to $(Ck/M)^r$, which ensures that $\{t'_k\}$ satisfies all conditions in (2.1). ∎

3. Error analysis for L1-type discretizations.

3.1. Error estimation for a simplest example (without spatial derivatives). It is convenient to illustrate our approach to the estimation of the temporal-discretization error using a very simple example. Consider a fractional-derivative problem without spatial derivatives together with its discretization:

$$D_t^\alpha u(t) = f(t) \quad \text{for } t \in (0, T], \qquad u(0) = u_0, \tag{3.1a}$$

$$\delta_t^\alpha U^m = f(t_m) \quad \text{for } m = 1, \ldots, M, \qquad U^0 = u_0. \tag{3.1b}$$

Throughout this subsection, with slight abuse of notation, ∂_t will be used for $\frac{d}{dt}$, while $\delta_t u(t_j) := \tau_j^{-1}[u(t_j) - u(t_{j-1})]$.

The main result here is the following error estimate, to the proof of which we shall devote the remainder of the subsection.

THEOREM 3.1. *Let the temporal mesh either satisfy* (2.1) *with* $r \geq 1$, *or include a submesh of type* (2.1) *with the same first mesh interval. Suppose that* u *and* $\{U^m\}$ *satisfy* (3.1), *and* $|\partial_t^l u| \lesssim 1 + t^{\alpha-l}$ *for* $l = 1, 2$ *and* $t \in (0, T]$. *Then* $\forall\, m \geq 1$

$$|u(t_m) - U^m| \lesssim \mathcal{E}^m := \begin{cases} M^{-r}\, t_m^{\alpha-1} & \text{if } 1 \leq r < 2-\alpha, \\ M^{\alpha-2}\, t_m^{\alpha-1}[1+\ln(t_m/t_1)] & \text{if } r = 2-\alpha, \\ M^{\alpha-2}\, t_m^{\alpha-(2-\alpha)/r} & \text{if } r > 2-\alpha. \end{cases} \tag{3.2}$$

Remark 3.2 (convergence in positive time). Consider $t_m \gtrsim 1$. Then $\mathcal{E}^m \simeq M^{-r}$ for $r < 2-\alpha$ and $\mathcal{E}^m \simeq M^{\alpha-2}$ for $r > 2-\alpha$, i.e., in the latter case the optimal convergence rate is attained. For $r = 2-\alpha$ one gets an almost optimal convergence rate as now $\mathcal{E}^m \simeq M^{\alpha-2}\ln M$.

By contrast, [13, Theorem 3.1] (obtained by means of a discrete Grönwall inequality) gives a somewhat similar but less sharp error bound for graded meshes, as (in our notation) it involves the term $O(\tau^\alpha) = O(M^{-\alpha r})$, so, e.g., for $r = 2-\alpha$ the error bound [13, Equation 3.17] gives a considerably less sharp convergence rate of only $\alpha(2-\alpha)$. For $r = 1$, we have $\mathcal{E}^m \simeq M^{-1}$, so our error bound is consistent with [6, 8, 10] and is again sharper than [13, Equation 3.17].

Remark 3.3 (global convergence). Note that $\max_{m\geq1}\mathcal{E}^m \simeq \mathcal{E}^1 \simeq \tau_1^\alpha \simeq M^{-\alpha r}$ for $\alpha \leq (2-\alpha)/r$, while $\max_{m\geq1}\mathcal{E}^m \simeq \mathcal{E}^M \simeq M^{\alpha-2}$ otherwise. Consequently, Theorem 3.1 yields the global error bound $|u(t_m) - U^m| \lesssim M^{-\min\{\alpha r, 2-\alpha\}}$. This implies that the optimal grading parameter for global accuracy is $r = (2-\alpha)/\alpha$. Note that similar global error bounds were obtained in [13, 10, 20].

We first prove an auxiliary result.

LEMMA 3.4 (truncation error). *For a sufficiently smooth* u, *let* $r^m := \delta_t^\alpha u(t_m) - D_t^\alpha u(t_m)\ \forall\, m \geq 1$, *and*

$$\psi^1 := \sup_{s\in(0,t_1)}\big(s^{1-\alpha}|\delta_t u(t_1) - \partial_s u(s)|\big), \tag{3.3a}$$

$$\psi^j := t_j^{2-\alpha}\sup_{s\in(t_{j-1},t_j)}|\partial_s^2 u(s)| \qquad \forall\, j \geq 2. \tag{3.3b}$$

Then, under conditions (2.1) *on the temporal mesh,*

$$|r^m| \lesssim (\tau/t_m)^{\min\{\alpha+1,\,(2-\alpha)/r\}}\max_{j=1,\ldots,m}\{\psi^j\}. \tag{3.4}$$

Proof. To a large degree we shall follow the proofs of [10, Lemmas 2.3 and 2.3*], so some details will be skipped. First, recalling the definitions (1.1) and (1.5) of D_t^α and δ_t^α and using the auxiliary function $\chi := u - u^I$, we arrive at

$$\Gamma(1-\alpha)\, r^m = \sum_{j=1}^m\int_{t_{j-1}}^{t_j}(t_m-s)^{-\alpha}\underbrace{[\delta_t u(t_j) - \partial_s u(s)]}_{=-\chi'(s)}\, ds = \alpha\int_0^{t_m}(t_m-s)^{-\alpha-1}\chi(s)\, ds.$$

On $(0, t_1)$ note that $\chi(s) = -\int_s^{t_1}\chi'(\zeta)d\zeta$, so $|\chi(s)| \lesssim s^{\alpha-1}(t_1-s)\psi^1$ (see [10, Equation 2.7b] for details). Otherwise, $|\chi| \lesssim \tau_j^2 t_j^{\alpha-2}\psi^j$ on (t_{j-1}, t_j) for $1 < j < m$ and $|\chi| \lesssim \tau_m(t_m-s)t_m^{\alpha-2}\psi^m$ on on (t_{m-1}, t_m). Consequently, a calculation shows that

$$|r^m| \lesssim \mathring{\mathcal{J}}^m\,(\tau_1/t_m)^{\alpha+1}\,\psi^1 + \mathcal{J}^m\max_{j=2,\ldots,m}\big\{\nu_{m,j}(\tau_j/t_j)^{2-\alpha}(t_j/t_m)^{\alpha+1}\,\psi^j\big\}. \tag{3.5}$$

Note that in various places here we also used $t_{j-1} \simeq t_j \simeq s$ for $s \in (t_{j-1}, t_j)$, $j > 1$. The notation in (3.5) is as follows:

$$\mathring{\mathcal{J}}^m := (t_m/\tau_1)^{\alpha+1} \int_0^{t_1} s^{\alpha-1}(t_1 - s)\,(t_m - s)^{-\alpha-1} ds \lesssim 1,$$

$$\mathcal{J}^m := \tau_m^\alpha\, t_m^{\alpha/r+1} \int_{t_1}^{t_m} s^{-\alpha/r-1}\,(t_m - s)^{-\alpha-1}\, \min\{1, (t_m - s)/\tau_m\}\, ds \lesssim 1,$$

$$\nu_{m,j} := (\tau_j/\tau_m)^\alpha\,(t_j/t_m)^{-\alpha(1-1/r)} \simeq 1.$$

Here the bound on $\nu_{m,j}$ follows from $\tau_j/\tau_m \simeq (t_j/t_m)^{1-1/r}$ (in view of (2.1)).

For the estimation of quantities of type $\mathring{\mathcal{J}}^m$ and $\mathcal{J}^m$, we refer the reader to [10]. In particular, for $\mathring{\mathcal{J}}^m$, we first use the observation that $(t_1 - s)/(t_m - s) \le t_1/t_m$ for $s \in (0, t_1)$. Then for $\mathring{\mathcal{J}}^m$ and $\mathcal{J}^m$, it is helpful to, respectively, use the substitutions $\hat{s} = s/t_1$ and $\hat{s} = s/t_m$, while for $\mathcal{J}^m$ we also employ $(t_1/t_m)^{-\alpha/r} \simeq (\tau_m/t_m)^{-\alpha}$ (also in view of (2.1)).

Combining the above observations with (3.5) yields

$$|r^m| \lesssim \max_{j=1,\dots,m} \bigl\{ \underbrace{(\tau_j/t_j)^{2-\alpha}}_{\simeq(\tau/t_j)^{(2-\alpha)/r}}\ (t_j/t_m)^{\alpha+1}\,\psi^j \bigr\},$$

where we also used $\tau_j/t_j \simeq (\tau/t_j)^{1/r}$ (in view of (2.1)). The desired bound (3.4) follows as $\tau \le t_j \le t_m$. ∎

Corollary 3.5 (more general meshes). *Lemma 3.4 remains valid if the temporal mesh is obtained by adding new nodes to any mesh of type (2.1) under the condition that the first mesh interval remains unchanged.*

Proof. Suppose the temporal mesh $\{t'_k\}$ is obtained by refining the mesh $\{t_j\}_{j=0}^M$ of type (2.1). We again employ $\chi = u - u^I$, only now u^I denotes the piecewise-linear interpolant on the finer mesh $\{t'_k\}$. As $t'_1 = t_1$, the estimation of integrals over $(0, t_1)$ remains unchanged. If $t'_k = t_m$ for some $m > 1$, then we proceed exactly as in the proof of Lemma 3.4, as the same bound $|\chi| \lesssim \tau_j^2 \min\{1, (t_m - s)/\tau_m\}\, t_j^{\alpha-2}\psi^j$ holds true on $(t_{j-1}, t_j)\ \forall\, j > 1$ (even though u^I is now the interpolant on a finer mesh). We also use $\max_{j\le m} \psi^j \simeq \Psi'_k := \max_{l\le k}\{\psi'^l\}$. If $t'_k \in (t_{m-1}, t_m)$ for some $m > 1$, then one has a similar bound $|\chi| \lesssim \tau_j^2 \min\{1, (t'_k - s)/\tau_m\}\, t_j^{\alpha-2}\,\Psi'_k$ on $(t_{j-1}, \min\{t_j, t'_k\})$ $\forall\, j > 1$. So for the truncation error at t'_k we get a version of (3.5), in which (including its ingredients) t_m is replaced by $t'_k \simeq t_m$ and ψ^j is replaced by Ψ'_k. This again leads to the desired version of (3.4) for the mesh $\{t'_k\}$. ∎

Proof of Theorem 3.1. Consider the error $e^m := u(t_m) - U^m$, for which (3.1) implies $e^0 = 0$ and $\delta_t^\alpha e^m = r^m\ \forall\, m \ge 1$, where the truncation error r^m is from Lemma 3.4 and hence satisfies (3.4). Furthermore, under the conditions (2.1) on the temporal mesh (or its submesh), one has $\psi^1 \lesssim 1$ (in view of $|\delta_t u(t_1)| \le \tau_1^{-1}\int_0^{t_1} |\partial_s u|\, ds \lesssim \tau_1^{\alpha-1}$) and $\psi^j \lesssim 1$ for $j \ge 2$ (in view of $s \simeq t_j$ for $s \in (t_{j-1}, t_j)$ for this case). Consequently, in view of Lemma 3.4 and Corollary 3.5, we arrive at

$$|r^m| \lesssim (\tau/t_m)^{\gamma+1} \quad \forall\, m \ge 1, \qquad \text{where}\ \ \gamma + 1 := \min\{\alpha + 1, (2 - \alpha)/r\}.$$

Next consider three cases.

Case $1 \le r < 2 - \alpha$. Then both $(2 - \alpha)/r > 1$ and $\alpha + 1 > 1$, so $\gamma > 0$. An application of Theorem 2.1(i) for this case yields $|e^m| \lesssim \tau\, t_m^{\alpha-1}$, where $\tau \simeq M^{-r}$.

Case $r = 2 - \alpha$. Then $(2 - \alpha)/r = 1$, while $\alpha + 1 > 1$, so $\gamma = 0$. An application of Theorem 2.1(i) yields $|e^m| \lesssim \tau\, t_m^{\alpha-1}[1 + \ln(t_m/t_1)]$, where $\tau \simeq M^{-r} = M^{\alpha-2}$.

Case $r > 2-\alpha$. Then $(2-\alpha)/r < 1$, while $\alpha+1 > 1$, so $\gamma+1 = (2-\alpha)/r < 1$. An application of Theorem 2.1(where part (i) of this theorem is used if $r \le (2-\alpha)/\alpha$ and part (ii) is used otherwise) yields $|e^m| \lesssim \tau\, t_m^{\alpha-1}(\tau/t_m)^{(2-\alpha)/r-1} \simeq \tau^{(2-\alpha)/r} t_m^{\alpha-(2-\alpha)/r}$, where $\tau^{(2-\alpha)/r} \simeq M^{\alpha-2}$. ∎

3.2. Error analysis for the L1 semidiscretization in time. Consider the semidiscretization of our problem (1.3) in time using the L1 method:

$$\delta_t^\alpha U^m + \mathcal{L}U^m = f(\cdot, t_m) \text{ in } \Omega, \quad U^m = 0 \text{ on } \partial\Omega \quad \forall\, m = 1, \ldots, M; \quad U^0 = u_0. \tag{3.6}$$

THEOREM 3.6. *Let the temporal mesh either satisfy* (2.1) *with $r \ge 1$, or include a submesh of type* (2.1) *with the same first mesh interval. Given $p \in \{2, \infty\}$, suppose that u is from* (1.3), (1.4), *with $c - p^{-1}\sum_{k=1}^d \partial_{x_k} b_k \ge 0$, and $\|\partial_t^l u(\cdot, t)\|_{L_p(\Omega)} \lesssim 1+t^{\alpha-l}$ for $l = 1, 2$ and $t \in (0, T]$. Then for $\{U^m\}$ from* (3.6), *one has*

$$\|u(\cdot, t_m) - U^m\|_{L_p(\Omega)} \lesssim \mathcal{E}^m \qquad \forall\, m = 1, \ldots, M, \tag{3.7}$$

where $\mathcal{E}^m$ is from (3.2).

Proof. For the error $e^m := u(\cdot, t_m) - U^m$, using (1.3) and (3.6), and imitating the proof of [10, Theorem 3.1], one gets a version of [10, Equation 3.4]:

$$\delta_t^\alpha \|e^m\|_{L_p(\Omega)} \le \|r^m\|_{L_p(\Omega)} \qquad \forall\, m = 1, \ldots, M. \tag{3.8}$$

Here the truncation error $r^m := \delta_t^\alpha u(\cdot, t_m) - D_t^\alpha u(\cdot, t_m)$ is estimated in Lemma 3.4 and hence satisfies (3.4). The desired error bound is obtained by closely imitating the proof of Theorem 3.1. Importantly, parts (i) and (ii) of Theorem 2.1 remain applicable to (3.8) in view of Theorem 2.1(iii). ∎

3.3. Error analysis for full L1-type discretizations. Similarly to section 3.2, one can easily combine the analysis of section 3.1 with [10, sections 4–5] to obtain error bounds of type (3.2) for full discretizations of problem (1.3) with $\mathcal{L} = \mathcal{L}(t)$, whether finite differences or finite elements are employed as spatial discreziations. We shall give a flavor of such results.

3.3.1. Finite difference discretizations. Consider our problem (1.3)–(1.4) in the spatial domain $\Omega = (0, 1)^d \subset \mathbb{R}^d$. Suppose that the standard finite difference operator $\mathcal{L}_h$ from [10, section 4] is employed as a spatial discretization on a uniform tensor-product mesh Ω_h of size h. We shall assume that h is sufficiently small so that $\mathcal{L}_h$ satisfies the discrete maximum principle. Then, under the conditions of Theorem 3.6 with $p = \infty$, and additionally assuming that $\|\partial_{x_k}^l u(\cdot, t)\|_{L_\infty(\Omega)} \lesssim 1$ for $l = 3, 4$, $k = 1, \ldots, d$, and $t \in (0, T]$, one easily gets the following version of [10, Theorem 4.1]:

$$\|u(\cdot, t_m) - U^m\|_{\infty\,;\Omega_h} \lesssim \mathcal{E}^m + t_m^\alpha\, h^2 \qquad \forall\, m = 1, \ldots, M, \tag{3.9}$$

where $\|\cdot\|_{\infty\,;\Omega_h} := \max_{\Omega_h} |\cdot|$ denotes the spatial nodal maximum norm, while $\mathcal{E}^m$ is from (3.2).

3.3.2. Finite element discretizations. Discretize (1.3)–(1.4), posed in a general bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, by applying a standard Galerkin finite element spatial approximation to the temporal semidiscretization (3.6). A Lagrange finite element space $S_h \subset H_0^1(\Omega) \cap C(\bar{\Omega})$ of fixed degree $\ell \geq 1$, relative to a quasi-uniform simplicial triangulation of Ω, is employed, as in [10, section 5]. Then, under the conditions of Theorem 3.6 with $p = 2$, one easily gets the following version of [10, Theorem 5.1]:

$$\begin{aligned}(3.10)\qquad \|u(\cdot, t_m) - u_h^m\|_{L_2(\Omega)} &\lesssim \|u_0 - u_h^0\|_{L_2(\Omega)} + \mathcal{E}^m + \max_{t\in\{0,t_m\}} \|\rho(\cdot, t)\|_{L_2(\Omega)} \\ &\quad + \int_0^{t_m} \|\partial_t \rho(\cdot, t)\|_{L_2(\Omega)}\, dt \qquad \forall\, m = 1, \ldots, M.\end{aligned}$$

Here $u_h^m \in S_h$ is the finite element solution at time t_m, $\mathcal{E}^m$ is from (3.2), and $\rho(\cdot, t) := \mathcal{R}_h u(t) - u(\cdot, t)$ is the error of the standard Ritz projection $\mathcal{R}_h u(t) \in S_h$ of $u(\cdot, t)$. Under additional realistic assumptions on u, the final two terms in the above error estimate can be bounded by $O(h^{\ell+1})$, where h is the triangulation diameter [10, section 5]. One can also estimate the error in the norm $L_\infty(\Omega)$ imitating [10, section 5.2].

4. Generalization for the Alikhanov discrete fractional-derivative operator. In this section we shall show that the above error analysis is not restricted to L1 discretizations, but may be extended, without major modifications, to other discretizations. Here we shall focus on a higher-order discrete fractional-derivative operator proposed by Alikhanov [1], while in [11] a similar analysis is generalized for another higher-order scheme.

A stability property of type (1.2) will be established in section 4.2. Next, in section 4.3, the truncation error will be estimated and the error for the simplest problem without spatial derivatives will be bounded by a quantity similar to $\mathcal{E}^m$ in (3.2). A stability property of type (1.2) for Alikhanov-type semidiscretizations will be obtained in section 4.4, which will allow to extend our error analysis to this case. Finally, error bounds for full discretizations will be briefly discussed in section 4.5.

4.1. Alikhanov discrete fractional-derivative operator. Discrete maximum principle. The discrete fractional-derivative operator proposed by Alikhanov is associated with the point

$$(4.1a)\qquad t_m^* := t_{m-\alpha/2} = t_m - \tfrac{1}{2}\alpha\tau_m.$$

In the definition of this operator, as well as in its analysis, we shall employ three standard *Lagrange interpolation operators* with the following interpolation points:

$$\Pi_{1,j}\ :\ \{t_{j-1}, t_j\}, \qquad \Pi_{2,j}\ :\ \{t_{j-1}, t_j, t_{j+1}\}, \qquad \Pi_{2,j}^*\ :\ \{t_{j-1}, t_j^*, t_j\}.$$

Now, applying $\Pi_{2,j}$ to the computed solution values $\{U^j\}$ on (t_{j-1}, t_j) for $j < m$ and $\Pi_{1,m}$ on on (t_{m-1}, t_m^*), we define an alternative discretization for the fractional operator D_t^α $\forall\, m = 1, \ldots, M$:

$$(4.1b)\qquad \delta_t^{\alpha,*} U^m := D_t^\alpha(\Pi^m U)(t_m^*), \qquad \Pi^m := \begin{cases} \Pi_{2,j} & \text{on } (t_{j-1}, t_j) \quad \forall j < m, \\ \Pi_{1,m} & \text{on } (t_{m-1}, t_m^*). \end{cases}$$

Note that the interpolation operator $\Pi_{2,j}^*$ is not used in the definition of $\delta_t^{\alpha,*}$ but will be useful in the estimation of the truncation error. In particular, for the final interval (t_{m-1}, t_m^*) it will occasionally be convenient to employ the representation $\Pi_{1,m} = \Pi_{2,m}^* + (\Pi_{1,m} - \Pi_{2,m}^*)$, as the choice (4.1a) ensures for any sufficiently smooth function v that

$$(4.2)\qquad \int_{t_{m-1}}^{t_m^*} (\Pi_{1,m} v - \Pi_{2,m}^* v)'(s)\, (t_m^* - s)^{-\alpha} ds = 0.$$

Indeed, here $\Pi_{1,m}v - \Pi^*_{2,m}v = C(s - t_{m-1})(t_m - s)$, with some constant C, so one has $(\Pi_{1,m}v - \Pi^*_{2,m}v)' = 2C(s - t_{m-1/2})$ and, consequently, (4.1a) yields (4.2).

Remark 4.1 (discrete maximum principle). Sufficient conditions for the operator $\delta_t^{\alpha,*}$ to be associated with an M-matrix, and, hence, satisfy the discrete maximum principle, are given by [4, Lemma 4] (see also [4, Remark 3]) and [14, Assumption M1]. In particular, throughout this section we shall assume that either $0.4656 \le \rho_j \le \rho_{j-1}$ $\forall\, j \ge 2$ [4] or $\rho_j \ge 4/7$ $\forall\, j \ge 2$ [14], where $\rho_j := \tau_{j+1}/\tau_j$. It is sufficient for the discrete maximum principle, and it is satisfied, for example, by the standard graded mesh $\{t_j = T(j/M)^r\}_{j=0}^M$ with any $r \ge 1$.

4.2. Stability theorem for the Alikhanov scheme. To generalize the above error analysis to the Alikhanov scheme, we need to extend the stability result given by Theorem 2.1 to the operator $\delta_t^{\alpha,*}$.

THEOREM 4.2 (stability). *Let the temporal mesh* $\{t_j\}_{j=0}^M$ *satisfy the condition from Remark* 4.1.

(i) *Let the temporal mesh additionally satisfy* (2.1) *with* $1 \le r \le (3-\alpha)/\alpha$. *Given* $\gamma \in \mathbb{R}$ *and* $\{V^j\}_{j=0}^M$, *the stability property* (1.2) *holds true with* δ_t^α *replaced by* $\delta_t^{\alpha,*}$.

(ii) *If* $\gamma \le \alpha - 1$, *then, without further restrictions on the mesh,* (1.2) *holds true with* δ_t^α *replaced by* $\delta_t^{\alpha,*}$.

(iii) *The above results remain valid if* $|\delta_t^{\alpha,*}V^j| \lesssim (\tau/t_j)^{\gamma+1}$ *in* (1.2) *is replaced by* $\delta_t^{\alpha,*}|V^j| \lesssim (\tau/t_j)^{\gamma+1}$.

Proof. (i) This part is obtained similarly to the proof of Theorem 2.1(i), only with a few changes in obtaining a version of Lemma 2.3 for $\delta_t^{\alpha,*}$; see Lemma 4.3 below.

(ii) This part is obtained exactly as in the proof of Theorem 2.1, only instead of [10, Lemma 2.1(i)] we now employ a similar [4, Lemma 5] for $\delta_t^{\alpha,*}$.

(iii) This part is obtained exactly as in the proof of Theorem 2.1. ∎

LEMMA 4.3 (Lemma 2.3 for Alikhanov scheme). *Under the conditions of Theorem* 4.2(i) *on the temporal mesh, the discrete barrier function* $\{B^j\}_{j=0}^M$ *from* (2.3) *satisfies* (2.2) *with* δ_t^α *replaced by* $\delta_t^{\alpha,*}$.

Proof. As $t_j^* \simeq t_j$ (in view of (4.1a)), it suffices to prove that $\delta_t^{\alpha,*}B^j \ge \tau^\alpha(t_j^*)^{-\alpha-1}$ $\forall j \ge 1$. For the latter, we closely imitate the proof of Lemma 2.3. In particular, for $j \le p$ one gets $\delta_t^{\alpha,*}B^j = D_t^\alpha B(t_j^*)$. When estimating $\delta_t^{\alpha,*}B^m$ for $m > p$, a few modifications are required that we now describe.

For $D_t^\alpha B(t_m^*)$ we have (2.4), while, in view of (4.2), $\delta_t^{\alpha,*}B^m = D_t^\alpha(I_2B)(t_m^*)$, where $I_2B := \Pi_{2,j}B$ on (t_{j-1}, t_j) for $j < m$ and $I_2B := \Pi^*_{2,j}B$ on (t_{m-1}, t_m^*) (with interpolation points $\{t_{m-1}, t_m^*, t_m\}$), i.e., I_2B is a piecewise quadratic interpolant. Now $\Gamma(1-\alpha)[\delta_t^{\alpha,*}B^m - D_t^\alpha B(t_m^*)] = \sum_{j=p}^m \mu^j$, where (compare with (2.5))

$$\mu^j := \alpha \int_{t_{j-1}}^{\min\{t_j, t_m^*\}} (B - I_2B)(s)\,(t_m^* - s)^{-\alpha-1} ds.$$

The estimation of μ^j for $j > p$ is similar to the case of the L1 scheme, only now we use a sharper bound $|B - I_2B| \lesssim \tau_j^3 \min\{1, (t_m^* - s)/(t_m^* - t_{m-1})\}|B'''(t_{j-1})|$, where $|B'''(t_{j-1})| \lesssim s^{-\beta-3}$. So now we get the following version of (2.6), in which the factors that differ from the proof of Lemma 2.3 are framed:

$$\sum_{p+1}^m |\mu^j| \lesssim \tau^{\boxed{3}/r} \int_{t_p}^{t_m^*} s^{-\beta-\boxed{3}/r}\,(t_m - s)^{-\alpha-1}\,\min\{1, (t_m^* - s)/(t_m^* - t_{m-1})\}\,ds.$$

This leads to the following version of (2.7):

$$\sum_{p+1}^{m}|\mu^j| \lesssim \tau^{\boxed{3}/r}(t_m^*)^{-\boxed{3}/r-1}\Big[(t_p/t_m^*)^{\alpha-\boxed{3}/r}+(\tau/t_m^*)^{-\alpha/r}\Big]$$
$$\lesssim (\tau/t_p)^{\boxed{3}/r}\,t_p^{\alpha}(t_m^*)^{-\alpha-1}+\underbrace{(t_m^*)^{-1}(\tau/t_m^*)^{(\boxed{3}-\alpha)/r}}_{\lesssim\tau^{\alpha}(t_m^*)^{-\alpha-1}}$$
$$\lesssim \Big[(\tau/t_p)^{\boxed{3}/r}+(\tau/t_p)^{\alpha}\Big]\,t_p^{\alpha}(t_m^*)^{-\alpha-1}, \tag{4.3}$$

where in the second line we employed $(\tau/t_m^*)^{(3-\alpha)/r} \lesssim (\tau/t_m^*)^{\alpha}$ (in view of $r \le (3-\alpha)/\alpha$).

It remains to get a similar bound on $|\mu^p|$ (where $p < m$). As B' abruptly changes at t_p, we now employ $|B - I_2B| = |B - \Pi_{2,j}B| \lesssim \max_{[t_{p-1},t_{p+1}]}|B - B(t_p)| \lesssim \tau_p t_p^{-\beta-1}$. (Note that when using the latter bound, we rely on the property $\tau_j \simeq \tau_{j+1}$ for the stability of the interpolating operator $\Pi_{2,j}$ in the sense that $\max_{[t_{j-1},t_j]}|\Pi_{2,j}v| \lesssim \max_{[t_{j-1},t_{j+1}]}|v|$ for any continuous v.) Now a calculation shows that

$$|\mu^p| \lesssim \tau_p t_p^{-\beta-1}\int_{t_{p-1}}^{t_p}(t_m^*-s)^{-\alpha-1}ds \lesssim \tau_p t_p^{-\beta-1}(t_{p+1}^*/t_m^*)^{\alpha+1}\underbrace{\int_{t_{p-1}}^{t_p}(t_{p+1}^*-s)^{-\alpha-1}ds}_{\lesssim\tau_p^{-\alpha}},$$

where, in the second relation, we employed the observation $(t_{p+1}^*-s)/(t_m^*-s) \le t_{p+1}^*/t_m^*$ $\forall s \in (0, t_{p+1}^*)$ (in view of $t_{p+1}^* \le t_m^*$). Next, $|\mu^p| \lesssim (\tau_p/t_p)^{\beta}\, t_p^{\alpha} t_m^{-\alpha-1}$, and, in view of $\tau_p/t_p = (\tau/t_p)^{1/r}$ (by (2.1)), one gets $|\mu^p| \lesssim (\tau/t_p)^{\beta/r}\, t_p^{\alpha} t_m^{-\alpha-1}$.

Finally, combining the latter bound with (4.3), we conclude that $|\delta_t^{\alpha,*}B^m - D_t^{\alpha}B(t_m^*)| \lesssim \sum_{j=p}^{m}|\mu^j|$ will be dominated by $\frac{1}{2}D_t^{\alpha}B(t_m^*)$ from (2.4) if p is chosen sufficiently large. ∎

4.3. Error analysis of the Alikhanov scheme for a simplest example (without spatial derivatives). Consider a fractional-derivative problem without spatial derivatives together with its discretization using $\delta_t^{\alpha,*}$ from (4.1):

$$D_t^{\alpha}u(t) = f(t) \quad \text{for } t \in (0,T], \qquad u(0) = u_0, \tag{4.4a}$$
$$\delta_t^{\alpha,*}U^m = f(t_m^*) \quad \text{for } m = 1,\dots,M, \qquad U^0 = u_0. \tag{4.4b}$$

Then for the error we have a version of Theorem 3.1.

THEOREM 4.4. *Let the temporal mesh satisfy the condition from Remark* 4.1 *and* (2.1) *with $r \ge 1$. Suppose that u and $\{U^m\}$ satisfy* (3.1), *and $|\partial_t^l u| \lesssim 1 + t^{\alpha-l}$ for $l = 1, 3$ and $t \in (0,T]$. Then $\forall\, m \ge 1$*

$$|u(t_m) - U^m| \lesssim \mathcal{E}^{m,*} := \begin{cases} M^{-r}\,t_m^{\alpha-1} & \text{if } 1 \le r < 3-\alpha,\\ M^{\alpha-3}\,t_m^{\alpha-1}[1+\ln(t_m/t_1)] & \text{if } r = 3-\alpha,\\ M^{\alpha-3}\,t_m^{\alpha-(3-\alpha)/r} & \text{if } r > 3-\alpha. \end{cases} \tag{4.5}$$

Remark 4.5 (convergence in positive time). Consider $t_m \gtrsim 1$. Then $\mathcal{E}^{m,*} \simeq M^{-r}$ for $r < 3-\alpha$ and $\mathcal{E}^{m,*} \simeq M^{\alpha-3}$ for $r > 3-\alpha$, i.e., in the latter case the optimal convergence rate is attained. For $r = 3-\alpha$ one gets an almost optimal convergence rate as now $\mathcal{E}^{m,*} \simeq M^{\alpha-3}\ln M$.

Remark 4.6 (global convergence). Note that $\max_{m\geq 1}\mathcal{E}^{m,*}\lesssim\mathcal{E}^{1,*}\simeq\tau_1^\alpha\simeq M^{-\alpha r}$ for $\alpha\leq(3-\alpha)/r$, while $\max_{m\geq 1}\mathcal{E}^{m,*}\simeq\mathcal{E}^{M,*}\simeq M^{\alpha-3}$ otherwise. Consequently, Theorem 4.4 yields the global error bound $|u(t_m)-U^m|\lesssim M^{-\min\{\alpha r,3-\alpha\}}$. This implies that the optimal grading parameter for global accuracy is $r=(3-\alpha)/\alpha$. Note that a similar global error bound was obtained in [4].

The proof is, to a large degree, similar to the arguments in section 3.1, with slight modifications in the truncation error estimation.

LEMMA 4.7 (truncation error). *For a sufficiently smooth* u, *let* $r^m:=\delta_t^{\alpha,*}u(t_m)-D_t^\alpha u(t_m^*)\ \forall\, m\geq 1$, *and*

$$\psi^1:=\sup_{s\in(0,t_2)}\big(s^{1-\alpha}|\partial_s u(s)|\big)+t_2^{-\alpha}\mathrm{osc}\big(u,[0,t_2]\big), \tag{4.6a}$$

$$\psi^j:=t_j^{3-\alpha}\sup_{s\in(t_{j-1},t_{j+1})}|\partial_s^3 u(s)|\quad\forall\, 2\leq j\leq M-1,\qquad \psi^M:=\psi^{M-1}. \tag{4.6b}$$

Then, under conditions (2.1) *on the temporal mesh, one has*

$$|r^m|\lesssim(\tau_1/t_m)^{\min\{\alpha+1,\,(3-\alpha)/r\}}\max_{j=1,\ldots,m}\{\psi^j\}\qquad\forall\, m\geq 1. \tag{4.7}$$

Proof. We imitate the proof of Lemma 3.4, and also use the notation I_2 and some observations from the proof of Lemma 4.3. Recall that, in view of (4.2), $\delta_t^{\alpha,*}u(t_m)=D_t^\alpha(I_2u)(t_m^*)$ where $I_2=\Pi_{2,j}$ on (t_{j-1},t_j) for $j<m$ and $I_2:=\Pi_{2,j}^*$ on (t_{m-1},t_m^*).

Next, recalling the definition (1.1) of D_t^α and using the auxiliary function $\chi:=u-I_2u$, which satisfies $\chi(t_m^*)=0$, we arrive at

$$\Gamma(1-\alpha)\,r^m=\int_0^{t_m^*}(t_m^*-s)^{-\alpha}\underbrace{\partial_s[I_2u(s)-u(s)]}_{=-\chi'(s)}ds=\alpha\int_0^{t_m^*}(t_m^*-s)^{-\alpha-1}\chi(s)\,ds.$$

Let $t_1^{**}:=\min\{t_1,t_m^*\}$ and consider the intervals $(0,t_1^{**})$ and (t_1^{**},t_m^*) separately. On $(0,t_1^{**})$ note that $\chi(t_1^{**})=0$ implies $\chi(s)=-\int_s^{t_1^{**}}\chi'(\zeta)d\zeta$, where $|\chi'|\leq|\partial_s u|+|\partial_s(I_2u)|$, while $|\partial_s(I_2u)|\lesssim t_2^{-1}\mathrm{osc}(u,[0,t_2])\leq s^{\alpha-1}t_2^{-\alpha}\mathrm{osc}(u,[0,t_2])$ (in view of $s\leq\tau_1\simeq\tau_2$), so $|\chi(s)|\lesssim s^{\alpha-1}(t_1^{**}-s)\psi^1$. Note also that $|\chi|\lesssim\tau_j^3t_j^{\alpha-3}\psi^j$ on (t_{j-1},t_j) for $1<j<m$ and $|\chi|\lesssim\tau_m^2(t_m^*-s)t_m^{\alpha-3}\psi^m$ on (t_{m-1},t_m^*) if $m>1$. Consequently, a calculation shows that we get a version of (3.5):

$$|r^m|\lesssim\mathring{\mathcal{J}}^m\,(\tau_1/t_m)^{\alpha+1}\,\psi^1+\mathcal{J}^m\max_{j=2,\ldots,m}\{\nu_{m,j}(\tau_j/t_j)^{\boxed{3}-\alpha}(t_j/t_m)^{\alpha+1}\,\psi^j\}, \tag{4.8}$$

where for convenience, the factors that differ from the proof of Lemma 3.4 are framed. Note that in various places we also use $t_j^*\simeq t_j\simeq t_{j+1}$ for $j\geq 1$ and $s\simeq t_j$ on (t_{j-1},t_j). The notation in (4.8) is as follows:

$$\mathring{\mathcal{J}}^m:=(t_m/\tau_1)^{\alpha+1}\int_0^{t_1^{**}}s^{\alpha-1}(t_1^{**}-s)\,(t_m^*-s)^{-\alpha-1}ds\lesssim 1,$$

$$\mathcal{J}^m:=\tau_m^\alpha\,t_m^{\alpha/r+1}\int_{t_1^{**}}^{t_m^*}s^{-\alpha/r-1}\,(t_m^*-s)^{-\alpha-1}\,\min\{1,(t_m^*-s)/\tau_m\}\,ds\lesssim 1,$$

$$\nu_{m,j}:=(\tau_j/\tau_m)^\alpha\,(t_j/t_m)^{-\alpha(1-1/r)}\simeq 1.$$

Here the bound on $\nu_{m,j}$ follows from $\tau_j/\tau_m\simeq(t_j/t_m)^{1-1/r}$ (in view of (2.1)). For the estimation of quantities of type $\mathring{\mathcal{J}}^m$ and $\mathcal{J}^m$, we refer the reader to [10]. In particular, for $\mathring{\mathcal{J}}^m$, we first use the observation that $(t_1^{**}-s)/(t_m^*-s) \le t_1^{**}/t_m^* \simeq t_1/t_m$ for $s\in(0,t_1^{**})$. Then for $\mathring{\mathcal{J}}^m$ and $\mathcal{J}^m$, it is helpful to, respectively, use the substitutions $\hat{s}=s/t_1^{**}$ and $\hat{s}=s/t_m^*$, while for $\mathcal{J}^m$ we also employ $(t_1^{**}/t_m^*)^{-\alpha/r}\simeq(t_1/t_m)^{-\alpha/r}\simeq(\tau_m/t_m)^{-\alpha}$ (also in view of (2.1)).

Combining the above observations with (4.8) yields

$$|r^m| \lesssim \max_{j=1,\dots,m}\bigl\{\underbrace{(\tau_j/t_j)^{\boxed{3}-\alpha}}_{\simeq(\tau/t_j)^{(3-\alpha)/r}}(t_j/t_m)^{\alpha+1}\,\psi^j\bigr\},$$

where we also used $\tau_j/t_j\simeq(\tau/t_j)^{1/r}$ (in view of (2.1)). The desired bound (4.7) follows as $\tau\le t_j\le t_m$. ∎

Proof of Theorem 4.4. Consider the error $e^m := u(t_m)-U^m$, for which (4.4) implies $e^0=0$ and $\delta_t^{\alpha,*}e^m = r^m\ \forall m\ge 1$, where the truncation error r^m is from Lemma 4.7 and hence satisfies (4.7). Furthermore, under the conditions (2.1) on the temporal mesh, one has $\psi^1\lesssim 1$ (in view of $\operatorname{osc}\bigl(u,[0,t_2]\bigr)\le\int_0^{t_2}|\partial_s u|\,ds\lesssim t_2^\alpha$) and $\psi^j\lesssim 1$ for $j\ge 2$ (in view of $s\simeq t_j$ for $s\in(t_{j-1},t_j)$ for this case). Consequently, we arrive at

$$|r^m|\lesssim(\tau/t_m)^{\gamma+1}\quad\forall m\ge 1,\qquad\text{where }\ \gamma+1:=\min\{\alpha+1,(3-\alpha)/r\}.$$

The remainder of the proof employs Theorem 4.2 and closely follows the proof of Theorem 3.1. In particular, the three cases $1\le r<3-\alpha$, $r=3-\alpha$, and $r>3-\alpha$ are considered separately, while $\tau\simeq M^{-r}$ now implies $\tau^{(3-\alpha)/r}\simeq M^{\alpha-3}$. ∎

4.4. Error analysis for the Alikhanov-type semidiscretization in time. Consider the semidiscretization of our problem (1.3) in time using $\delta_t^{\alpha,*}$ from (4.1):

(4.9a)
$$\delta_t^{\alpha,*}U^m+\mathcal{L}U^{m,*}=f(\cdot,t_m^*)\ \text{ in }\Omega,\quad U^m=0\ \text{ on }\partial\Omega\quad\forall m=1,\dots,M;\quad U^0=u_0.$$

where, in view of (4.1a), we use a second-order discretization for $\mathcal{L}u(\cdot,t_m^*)$ with

$$U^{m,*}:=\tfrac12\alpha\,U^{m-1}+\bigl(1-\tfrac12\alpha\bigr)U^m. \tag{4.9b}$$

To simplify the presentation, here we shall consider only standard graded temporal meshes, which clearly satisfy both the condition from Remark 4.1 and (2.1). We shall also make some simplifying assumptions of $\mathcal{L}$.

LEMMA 4.8 (stability for fractional parabolic case). *Given $\gamma\in\mathbb{R}$, let $\{t_j=T(j/M)^r\}_{j=0}^M$ for some $1\le r\le(3-\alpha)/\alpha$ if $\gamma>\alpha-1$ or for some $r\ge1$ if $\gamma\le\alpha-1$. Also, let $\mathcal{L}$ of* (1.4) *be independent of t with $b_k=0\ \forall k$. Then for $\{U^j\}_{j=0}^M$ from* (4.9) *one has*

$$\left.\begin{array}{c}\|f(\cdot,t_j^*)\|_{L_2(\Omega)}\lesssim(\tau/t_j)^{\gamma+1}\\ \forall j\ge1,\quad U^0=0\ \text{ in }\bar\Omega\end{array}\right\}\ \Rightarrow\ \|U^j\|_{L_2(\Omega)}\lesssim\mathcal{V}^j(\tau;\gamma)\quad\forall j\ge1, \tag{4.10}$$

where $\mathcal{V}^j=\mathcal{V}^j(\tau;\gamma)$ is defined in (1.2).

Proof. (i) Throughout the proof, we shall use the notation

$$\delta_t^{\alpha,*}U^m=\sum_{j=0}^m\kappa_{m,j}U^j,\qquad f^j:=f(\cdot,t_j^*),\qquad\rho_j:=\tau_{j+1}/\tau_j,$$

where, in view of Remark 4.1, $\kappa_{m,m} > 0$, while $\kappa_{m,j} \le 0 \;\forall\, j < m$. An inspection of some arguments in [4] shows (see Remark 4.9 below for further details) that there exists a constant $c_0 = c_0(\alpha, r) \in (0,1)$ such that $\frac{1}{2}\alpha\kappa_{m,m} < c_0(1-\frac{1}{2}\alpha)|\kappa_{m,m-1}|$ $\forall\, m \ge 2$. Next, we claim that there is a sufficiently large $1 \le K \lesssim 1$ (where $K = K(\alpha, r)$ is independent of M) such that

$$\tfrac{1}{2}\alpha \le \kappa_{m,m}^{-1/2}\,\kappa_{m-1,m-1}^{-1/2}\,\big(1-\tfrac{1}{2}\alpha\big)|\kappa_{m,m-1}| \qquad \forall m \ge K+1. \tag{4.11}$$

Indeed, it suffices to check that $c_0^2 \le \kappa_{m,m}/\kappa_{m-1,m-1}$, while, in view of $\rho_j \le \rho_{j-1}$ $\forall\, j \ge 2$ a calculation shows that $\tau_{m-1}^{\alpha}\kappa_{m-1,m-1} \le \tau_m^{\alpha}\kappa_{m,m}$, hence it suffices to check that $c_0^{2/\alpha} \le \tau_{m-1}/\tau_m$, which can be ensured by choosing $K = K(\alpha, r)$ sufficiently large (see the proof of [11, Corollary 3.3] for further details).

We shall consider the cases $K = 1$ and $K > 1$ separately in parts (ii) and (iii).

(ii) Suppose $K = 1$ in (4.11). Then

$$\delta_t^{\alpha,*} w^m \le \|f^m\|_{L_2(\Omega)}, \quad w^m := \sqrt{\|U^m\|^2_{L_2(\Omega)} + \kappa_{m,m}^{-1}\big(1-\tfrac{1}{2}\alpha\big)\langle \mathcal{L}U^m, U^m\rangle}\,. \tag{4.12}$$

Indeed, in view of (4.11), taking the inner product of the equation from (4.9a) with U^m, one gets

$$\kappa_{m,m}(w^m)^2 \le |\kappa_{m,m-1}| w^m w^{m-1} + \sum_{j=1}^{m-2} |\kappa_{m,j}| \underbrace{\langle U^m, U^j\rangle}_{\le w^m w^j} + \underbrace{\langle U^m, f^m\rangle}_{\le w^m \|f^m\|_{L_2(\Omega)}}.$$

Dividing by w^m, we arrive at (4.12). Now an application of Theorem 4.2(iii) yields $w^j \lesssim \mathcal{V}^j$, and hence the desired result.

(iii) Suppose that $1 < K \lesssim 1$. We imitate part (ii) in the proof of [11, Theorem 3.2]. First, for $m \le K$, using $\tau_m \simeq \tau_1$ and (4.9), one gets $w^m \lesssim w^{m-1} + \sum_{j=0}^{m-2} \|U^j\|_{L_2(\Omega)} + \tau_1^{\alpha}\|f^m\|_{L_2(\Omega)}$. Here $\|f^m\|_{L_2(\Omega)} \lesssim 1$, so $\|U^m\|_{L_2(\Omega)} \lesssim \tau_1^{\alpha} \simeq \mathcal{V}^m$ $\forall\, m \le K$.

It remains to estimate the values of $\{\mathring{U}^j\}_{j=0}^M := \{0, \dots, 0, U^{K+1}, \dots, U^M\}$ (i.e., $\mathring{U}^j$ is set to 0 for $j \le K$ and to U^j otherwise). Note that $\delta_t^{\alpha,*}\mathring{U}^m = 0$ for $m \le K$ and $|\delta_t^{\alpha,*}\mathring{U}^{K+1}| \lesssim 1$. Consider $m \ge K+2$. Then, by (4.1b), one has $\delta_t^{\alpha,*}\mathring{U}^m = \delta_t^{\alpha,*}U^m - D_t^{\alpha}\Pi^m[U - \mathring{U}](t_m^*)$. As $\Pi^m[U - \mathring{U}]$ has support on $(0, t_{K+1})$, vanishes at 0 and $t_{K+1} \le t_m^*$, while its absolute value $\lesssim \tau_1^{\alpha}$, so, recalling (1.1) and applying an integration by parts yields $|D_t^{\alpha}\Pi^m[U - \mathring{U}](t_m)| \lesssim \tau_1^{\alpha}\int_0^{t_{K+1}} (t_m^* - s)^{-\alpha-1} ds \lesssim (\tau_1/t_m)^{\alpha+1}$ (where we also used $t_{K+1} \simeq \tau_1$ and $t_m^* - s \gtrsim t_m^* - t_{K+1} \simeq t_m$). Consequently, for $m \ge K+2$ one concludes that $|\delta_t^{\alpha,*}\mathring{U}^m|$ is $\lesssim (\tau_1/t_m)^{\gamma+1}$ if $\gamma \le \alpha$ and $\lesssim (\tau_1/t_m)^{\alpha+1}$ otherwise.

Finally, we restrict the problem for $\{\mathring{U}^j\}_{j=K-1}^M$ to the mesh $\{t_j\}_{j=K-1}^M$ and note that for the Alikhanov-type operator $\mathring{\delta}_t^{\alpha,*}$ associated with the latter mesh one gets $\mathring{\delta}_t^{\alpha,*}\mathring{U}^K = 0$ and $\mathring{\delta}_t^{\alpha,*}\mathring{U}^m = \delta_t^{\alpha,*}\mathring{U}^m$ for $m \ge K+1$. Now, in view of (4.11), an application of the result of part (ii) yields $\|\mathring{U}^j\|_{L_2(\Omega)} \lesssim \mathcal{V}^j$, which leads to the desired bound on $\|\mathring{U}^j\|_{L_2(\Omega)}$. ∎

Remark 4.9. Comparing our notation $\{\kappa_{m,j}\}$ with [4, Equation 7], the relation [4, Equation 15] can be rewritten as $\sigma|\kappa_{m,m-1}| - (1-\sigma)\kappa_{m,m} > 0$, where $\sigma = 1 - \frac{1}{2}\alpha$, while a sufficient condition for the latter is given by [4, Equation 17] and is satisfied by our mesh. Furthermore, an inspection of the proof of [4, Lemma 4] shows such that in the second relation in [4, Equation 41], one can include a constant factor $c_1(\alpha, \sigma\bar{\rho}) \in (0,1)$ in the right-hand side, where $\bar{\rho} := \max \rho_j = \rho_1$ on our mesh.

(The latter observation can be shown by inspecting the proof of [4, Lemma 2] and replacing the the piecewise-constant approximation of $(t_{k+\sigma}-\eta)^{-\alpha}$ by a piecewise-linear one.) Then (under the same sufficient condition) one obtains a stronger version of [4, Equation 15]: $\sigma^*|\kappa_{m,m-1}|-(1-\sigma^*)\kappa_{m,m}>0$ with $(2\sigma^*-1)/\sigma^*:=c_1(2-\sigma)/\sigma$, i.e., $\sigma^*=\sigma^*(\alpha,\sigma\bar\rho)<\sigma$. Consequently, $(1-\sigma)\kappa_{m,m}<c_0\sigma|\kappa_{m,m-1}|$, where $c_0:=\frac{\sigma^*}{1-\sigma^*}\frac{1-\sigma}{\sigma}\in(0,1)$. Recalling that $\sigma=1-\frac12\alpha$, we conclude that $\forall\, m\ge 2$ one has $\frac12\alpha\kappa_{m,m}<c_0(1-\frac12\alpha)|\kappa_{m,m-1}|$.

THEOREM 4.10. *Let* $\{t_j=T(j/M)^r\}_{j=0}^M$ *for some* $r\ge 1$. *Suppose that* u *is from* (1.3), (1.4), *where* $\mathcal{L}$ *of* (1.4) *is independent of* t *with* $b_k=0\ \forall\, k$. *Also, let* $\|\partial_t^l u(\cdot,t)\|_{L_2(\Omega)}\lesssim 1+t^{\alpha-l}$ *for* $l=1,3$ *and* $\|\partial_t^2\mathcal{L}u(\cdot,t)\|_{L_2(\Omega)}\lesssim 1+t^{\alpha-2}\ \forall\, t\in(0,T]$. *Then for* $\{U^m\}$ *from* (4.9), *one has*

$$\|u(\cdot,t_m)-U^m\|_{L_2(\Omega)}\lesssim \mathcal{E}^{m,**}:=\mathcal{E}^{m,*}+M^{-2}\begin{cases} t_m^{2\alpha-2/r} & \text{if } 2/r<\alpha+1,\\ 0, & \text{otherwise}\end{cases}\qquad \forall\, m\ge 1,$$

where $\mathcal{E}^{m,*}$ *is from* (4.5).

Proof. For the error $e^m:=u(\cdot,t_m)-U^m$, using (1.3) and (4.9), one immediately gets $e^0=0$ and $\forall\, m\ge 1$

$$\delta_t^{\alpha,*}e^m+\mathcal{L}e^{m,*}=r^m+R^m,\qquad \text{where}\quad R^m:=\mathcal{L}u^{m,*}-\mathcal{L}u(\cdot,t_m^*) \tag{4.13}$$

with the notation $u^{m,*}:=\frac12\alpha\, u(\cdot,t_{m-1})+\big(1-\frac12\alpha\big)u(\cdot,t_m)$, and the truncation error r^m from Lemma 4.7 that satisfies (4.7). So under our assumptions of u one has $\|R^m\|_{L_2(\Omega)}\lesssim \tau_m^2t_m^{\alpha-2}\simeq\tau^{2/r}t_m^{\alpha-2/r}$ in view of (2.1), and also $\|r^m\|_{L_2(\Omega)}\lesssim(\tau/t_m)^{\gamma+1}$, where $\gamma+1:=\min\{\alpha+1,(3-\alpha)/r\}$ (see the proof of Theorem 4.4).

If $2/r\ge\alpha+1\ge\gamma+1$, then $\|R^m\|_{L_2(\Omega)}\lesssim(\tau/t_m)^{2/r}\le(\tau/t_m)^{\gamma+1}$, so Lemma 4.8 yields $\|e^m\|_{L_2(\Omega)}\lesssim\mathcal{V}^m(\tau;\gamma)$. Otherwise, Lemma 4.8 yields $\|e^m\|_{L_2(\Omega)}\lesssim\mathcal{V}^m(\tau;\gamma)+\tau^\alpha\mathcal{V}^m(\tau;\gamma')$, where $\gamma':=2/r-\alpha-1<0$ implies $\tau^\alpha\mathcal{V}^m(\tau;\gamma')=\tau^{2/r}t_m^{2\alpha-2/r}$, where $\tau^{2/r}\simeq M^{-2}$. Finally, imitating the proof of Theorem 4.4, one gets $\mathcal{V}^m(\tau;\gamma)\lesssim\mathcal{E}^{m,*}$, so combining our findings we arrive at the desired error bound. ∎

Remark 4.11 (convergence in positive time). Consider $t_m\gtrsim 1$. Then, in view of Remark 4.5, $\mathcal{E}^{m,**}\simeq M^{-r}$ for $r\le 2$. Otherwise, $2/r<1<\alpha+1$ so $\mathcal{E}^{m,**}\simeq M^{-2}$. In summary, $\mathcal{E}^{m,**}\simeq M^{-\min\{r,2\}}$, and $r=2$ yields the optimal convergence rate 2.

By contrast, [14, Theorem 3.9] (obtained by means of a discrete Grönwall inequality [15]) gives a somewhat similar but less sharp error bound for graded meshes, as (in our notation) it involves the term $O(\tau^\alpha)=O(M^{-\alpha r})$, so, e.g., for $r=2$ it gives a considerably less sharp convergence rate of only 2α. For $r=1$, we have $\mathcal{E}^{m,**}\simeq M^{-1}$, so our error bound is again sharper than those in [14, Theorem 3.9].

Remark 4.12 (global convergence). In view of Remark 4.6, $\max_{m\ge1}\mathcal{E}^{m,*}\simeq M^{-\min\{\alpha r,3-\alpha\}}$. If $\alpha r\ge 1$, then $2/r\le 2\alpha<\alpha+1$ so $\max_{m\ge1}\mathcal{E}^{m,**}\simeq\max_{m\ge1}\mathcal{E}^{m,*}+M^{-2}\simeq M^{-\min\{\alpha r,2\}}$. Otherwise, $\alpha r<1$ implies $\max_{m\ge1}\mathcal{E}^{m,*}\simeq M^{-\alpha r}$, while $\max_{m\ge1}M^{-2}t_m^{2\alpha-2/r}$ is attained at $m=1$ and is $\simeq\tau^{2\alpha}\simeq M^{-2\alpha r}$, so $\max_{m\ge1}\mathcal{E}^{m,**}\simeq M^{-\alpha r}$. Consequently, Theorem 4.10 yields the global error bound $|u(t_m)-U^m|\lesssim M^{-\min\{\alpha r,2\}}$. This implies that the optimal grading parameter for global accuracy is $r=2/\alpha$. Note that a similar global error bound was obtained in [4].

4.5. Alikhanov-type finite element discretizations. Discretize (1.3)–(1.4), posed in a general bounded Lipschitz domain $\Omega\subset\mathbb{R}^d$, by applying a standard

Galerkin finite element spatial approximation, described in section 3.3.2, to the temporal semidiscretization (4.9). Then for the full discretization one can easily generalize the stability result given by Lemma 4.8. Furthermore, for $e_h^m := \mathcal{R}_h u(t_m) - u_h^m \in S_h$, where $u_h^m \in S_h$ is the finite element solution at time t_m, and $\mathcal{R}_h u(t) \in S_h$ is the Ritz projection of $u(\cdot, t)$, a standard calculation (see, e.g., [10, Theorem 5.1]) yields

$$\langle \delta_t^{\alpha,*} e_h^m, v_h \rangle + \mathcal{A}(e_h^{m,*}, v_h) = \langle \delta_t^{\alpha,*} \rho + r^m + R^m, v_h \rangle \quad \forall v_h \in S_h.$$

Here $\mathcal{A}(\cdot,\cdot)$ is the standard bilinear form associated with $\mathcal{L}$, and $\rho(\cdot, t) := \mathcal{R}_h u(t) - u(\cdot, t)$. So, under the conditions of Theorem 4.10 and assuming that $u_h^0 = \mathcal{R}_h u(0)$, one gets the following version of [11, Theorem 5.5]:

$$\begin{aligned} \|u(\cdot, t_m) - u_h^m\|_{L_2(\Omega)} \lesssim \mathcal{E}^{m,**} + \|\rho(\cdot, t_m)\|_{L_2(\Omega)} \\ + t_m^\alpha \sup_{t \in (0, t_m)} \left\{ t^{1-\alpha} \|\partial_t \rho(\cdot, t)\|_{L_2(\Omega)} \right\} \quad \forall\, m \geq 1, \end{aligned}$$

where $\mathcal{E}^{m,**}$ is from Theorem 4.10. Under additional realistic assumptions on u, the final two terms in the above error estimate can be bounded by $O(h^{\ell+1})$, where h is the triangulation diameter [10, section 5].

5. Numerical results.

5.1. Fractional parabolic test with finite elements. Our first fractional-order parabolic test problem is (1.3) with $\mathcal{L} = -(\partial_{x_1}^2 + \partial_{x_2}^2)$, posed in the domain $\Omega \times [0, 1]$ from [10, section 7] with $\partial\Omega$ parameterized by $x_1(l) := \frac{2}{3} R \cos\theta$ and $x_2(l) := R \sin\theta$, where $R(l) := 0.8 + \cos^2 l$ and $\theta(l) := l + e^{(l-5)/2} \sin(l/2) \sin l$ for $l \in [0, 2\pi]$. We choose f, as well as the initial and nonhomogeneous boundary conditions, so that the unique exact solution $u = t^\alpha [1 + \ln(x - y/3 + 7)]$. This problem is discretized in space (with an obvious modification for the case of nonhomogeneous boundary conditions) using lumped-mass linear finite elements on quasiuniform Delaunay triangulations of Ω (with "DOF" denoting the number of degrees of freedom in space). The errors will be computed in the approximate $L_2(\Omega)$ norm as $\|u_h - u^I\|_{L_2(\Omega)}$, where $u^I \in S_h$ is the piecewise-linear interpolant in Ω. All numerical experiments will use the graded temporal mesh $\{t_j = T(j/M)^r\}_{j=0}^M$.

For the L1 method, we have the error bounds (3.7) and (3.10). These error bounds are consistent with the numerical rates of convergence given in [6] for errors in positive time and $r = 1$, as well as those in [20, 10] for errors in the maximum norm in time and various r. Additionally, consider the case $r > 2 - \alpha$, for which our error bounds predict the optimal convergence rate of $2 - \alpha$ with respect to time at $t \gtrsim 1$ (see Remark 3.2). This agrees with the numerical convergence rates given in Table 5.1 for the L1 method with $r = (2 - \alpha)/0.9$.

TABLE 5.1

Fractional-order parabolic test problem from section 5.1: $L_2(\Omega)$ errors at $t = 1$ (odd rows) and computational rates q in M^{-q} (even rows) for the L1 method with $r = (2-\alpha)/.9$ and the Alikhanov method with $r = 2$, spatial DOF $= 398410$.

	L1 method, $r = \frac{2-\alpha}{.9}$				Alikhanov method, $r = 2$			
	$M = 2^6$	$M = 2^7$	$M = 2^8$	$M = 2^9$	$M = 2^6$	$M = 2^7$	$M = 2^8$	$M = 2^9$
$\alpha = 0.3$	8.35e-5	2.58e-5	7.98e-6	2.48e-6	5.79e-6	1.25e-6	2.83e-7	7.04e-8
	1.69	1.69	1.69		2.21	2.15	2.01	
$\alpha = 0.5$	2.66e-4	9.47e-5	3.38e-5	1.20e-5	7.10e-6	1.58e-6	3.67e-7	9.15e-8
	1.49	1.49	1.49		2.16	2.11	2.00	
$\alpha = 0.7$	5.61e-4	2.30e-4	9.44e-5	3.88e-5	7.38e-6	1.72e-6	4.09e-7	1.03e-7
	1.29	1.28	1.28		2.10	2.07	1.99	

TABLE 5.2
Fractional-order parabolic test problem from section 5.2: maximum nodal errors at $t = 1$ (odd rows) and computational rates q in M^{-q} or N^{-q} (even rows) for the L1 method with $r = (2-\alpha)/.9$.

	errors and convergence rates in time $N = M$				errors and convergence rates in space $M = N^2$			
	$M = 2^5$	$M = 2^6$	$M = 2^7$	$M = 2^8$	$N = 2^3$	$N = 2^4$	$N = 2^5$	$N = 2^6$
$\alpha = 0.3$	6.99e-4	2.30e-4	7.45e-5	2.39e-5	2.81e-3	7.36e-4	1.87e-4	4.86e-5
	1.60	1.63	1.64		1.93	1.97	1.95	
$\alpha = 0.5$	1.54e-3	5.75e-4	2.10e-4	7.59e-5	2.87e-3	7.34e-4	1.84e-4	4.86e-5
	1.43	1.45	1.47		1.97	1.99	1.92	
$\alpha = 0.7$	3.05e-3	1.28e-3	5.29e-4	2.17e-4	3.15e-3	7.84e-4	1.91e-4	4.86e-5
	1.25	1.27	1.28		2.01	2.04	1.97	

For the Alikhanov method, we have the error bounds of Theorem 4.10 and section 4.5. Note that they are consistent with the numerical rates of convergence given in [4] for errors in the maximum norm in time and various r. Additionally, here we numerically investigate the case $r = 2$, for which our error bounds predict the optimal convergence rate 2 with respect to time at $t \gtrsim 1$ (see Remark 4.11). This clearly agrees with the numerical convergence rates given in Table 5.1 for the the Alikhanov method.

5.2. Fractional parabolic test with finite differences. To test the error bound (3.9) given in section 3.3.1 for finite difference discretizations in space combined with the L1 scheme in time, we shall employ another test problem. Consider (1.3) with $\mathcal{L} = -(\partial^2_{x_1} + \partial^2_{x_2}) + (1 + x_1 + x_2 + t)$, the initial condition $u_0 = \sin x_1 \sin x_2$, and $f = x_1(\pi - x_1)x_2(\pi - x_2)(1 + t^4) + t^2$, posed in the domain $\Omega \times [0, 1]$ with the square spatial domain $\Omega = (0, \pi)^2$ (this test is a modification of [20, Example 6.2]). The spatial mesh was a uniform tensor product mesh of size $h = \pi/N$ (i.e., with N equal mesh intervals in each coordinate direction). As the exact solution is unknown, the errors were computed using the two-mesh principle.

We focus on the most interesting case of the graded temporal mesh with $r > 2-\alpha$, for which our error bound (3.9) predicts the optimal convergence rate of $2 - \alpha$ with respect to time at $t \gtrsim 1$ (in view of Remark 3.2). This clearly agrees with the numerical convergence rates given in Table 5.2 for the grading parameter $r = (2 - \alpha)/0.9$.

5.3. L1 method: pointwise sharpness of the error estimate for the initial-value problem. Here, to demonstrate the sharpness of the error estimate (3.2) given by Theorem 3.1 for the L1 method, we consider the simplest initial-value fractional-derivative test problem (3.1) with the simplest typical exact solution $u(t) := t^\alpha$. Table 5.3 shows the errors and the corresponding convergence rates at $t = 1$, which agree with (3.2), in view of Remark 3.2. In particular, the latter implies that the errors are $\lesssim M^{-\min\{r,2-\alpha\}}$ for $r \neq 2 - \alpha$ and $\lesssim M^{-(2-\alpha)} \ln M$ for $r = 2 - \alpha$. The maximum errors and corresponding convergence rates for various α and r are given in [20, 10], and they confirm the conclusions of Remark 3.3, which predicts from the pointwise bound (3.2) that the global errors are $\lesssim M^{-\min\{\alpha r,2-\alpha\}}$.

Furthermore, in Figure 5.1, the pointwise errors for various r are compared with the pointwise theoretical error bound (3.2), and again, with the exception of a few initial mesh nodes, we observe remarkably good agreement. Note that Figure 5.1 only

Table 5.3
L1 method applied to the initial-value test problem: errors at $t = 1$ (odd rows) and computational rates q in M^{-q} (even rows) for $r = 1$, $r = 2 - \alpha$ and $r = (2 - \alpha)/.95$.

		$M = 2^7$	$M = 2^9$	$M = 2^{11}$	$M = 2^{13}$	$M = 2^{15}$	$M = 2^{17}$
$r = 1$	$\alpha = 0.3$	1.182e-3	2.939e-4	7.333e-5	1.832e-5	4.578e-6	1.144e-6
		1.004	1.001	1.001	1.000	1.000	
	$\alpha = 0.5$	1.953e-3	4.883e-4	1.221e-4	3.052e-5	7.629e-6	1.907e-6
		1.000	1.000	1.000	1.000	1.000	
	$\alpha = 0.7$	2.489e-3	6.433e-4	1.642e-4	4.163e-5	1.050e-5	2.640e-6
		0.976	0.985	0.990	0.994	0.996	
$r = 2 - \alpha$	$\alpha = 0.3$	1.201e-4	1.310e-5	1.401e-6	1.477e-7	1.540e-8	1.592e-9
		1.598	1.612	1.623	1.631	1.637	
	$\alpha = 0.5$	5.039e-4	7.407e-5	1.063e-5	1.500e-6	2.089e-7	2.878e-8
		1.383	1.400	1.413	1.422	1.430	
	$\alpha = 0.7$	1.267e-3	2.495e-4	4.782e-5	8.986e-6	1.663e-6	3.042e-7
		1.172	1.192	1.206	1.217	1.225	
$r = \frac{2-\alpha}{.95}$	$\alpha = 0.3$	1.035e-4	1.074e-5	1.094e-6	1.098e-7	1.092e-8	1.076e-9
		1.634	1.648	1.658	1.665	1.671	
	$\alpha = 0.5$	4.469e-4	6.276e-5	8.609e-6	1.161e-6	1.546e-7	2.039e-8
		1.416	1.433	1.445	1.454	1.461	
	$\alpha = 0.7$	1.143e-3	2.164e-4	3.984e-5	7.192e-6	1.279e-6	2.250e-7
		1.201	1.221	1.235	1.245	1.254	

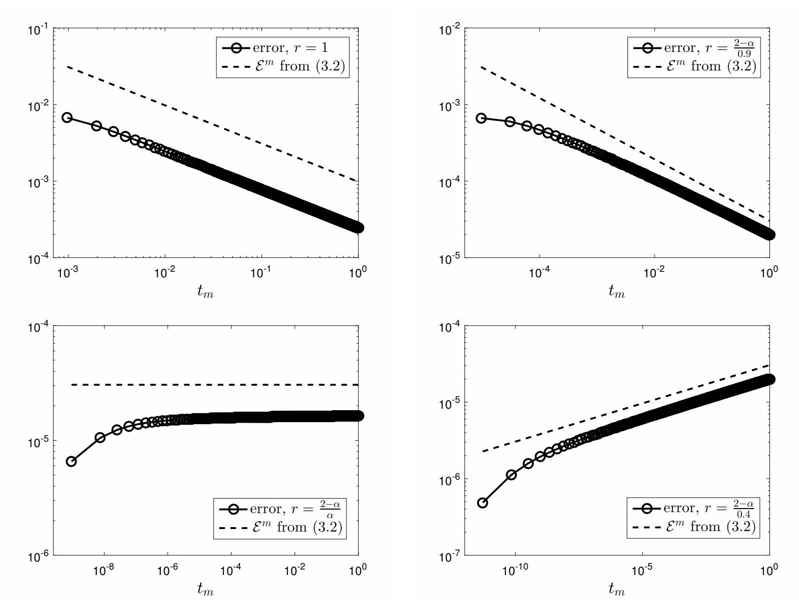

Fig. 5.1. *L1 method applied to the initial-value test problem: pointwise errors for $\alpha = 0.5$ and $M = 1024$, cases $r = 1$, $r = (2 - \alpha)/0.9$, $r = (2 - \alpha)/\alpha$ and $r = (2 - \alpha)/0.4$.*

addresses the case $\alpha = 0.5$, but for other values of α we observed similar consistency of (3.2) with the actual pointwise errors.

5.4. Alikhanov method: pointwise sharpness of the error estimate for the initial-value problem. Next, we turn to the Alikhanov method and, to demonstrate the sharpness of the error estimate (4.5) given by Theorem 4.4, consider the simplest initial-value fractional-derivative test problem (4.4) with the same simplest typical exact solution $u(t) := t^\alpha$. Table 5.4 shows the errors and the corresponding convergence rates at $t = 1$, which agree with (4.5), in view of Remark 4.5. In particular, the latter implies that the errors are $\lesssim M^{-\min\{r,3-\alpha\}}$ for $r \neq 3 - \alpha$. The maximum errors and corresponding convergence rates given in Table 5.5 clearly confirm the conclusions of Remark 4.6, which predicts from the pointwise bound (4.5) that the global errors are $\lesssim M^{-\min\{\alpha r,3-\alpha\}}$.

TABLE 5.4
Alikhanov method applied to the initial-value test problem: errors at $t = 1$ (odd rows) and computational rates q in M^{-q} (even rows) for $r = 1$, $r = 2$ and $r = (3 - \alpha)/.95$.

		$M = 2^6$	$M = 2^8$	$M = 2^{10}$	$M = 2^{12}$	$M = 2^{14}$	$M = 2^{16}$
$r = 1$	$\alpha = 0.3$	1.325e-3	3.306e-4	8.260e-5	2.065e-5	5.162e-6	1.290e-6
		1.002	1.000	1.000	1.000	1.000	
	$\alpha = 0.5$	1.530e-3	3.819e-4	9.543e-5	2.386e-5	5.964e-6	1.491e-6
		1.001	1.000	1.000	1.000	1.000	
	$\alpha = 0.7$	1.236e-3	3.087e-4	7.715e-5	1.929e-5	4.821e-6	1.205e-6
		1.001	1.000	1.000	1.000	1.000	
$r = 2$	$\alpha = 0.3$	3.891e-5	2.446e-6	1.530e-7	9.560e-9	5.975e-10	3.734e-11
		1.996	1.999	2.000	2.000	2.000	
	$\alpha = 0.5$	6.079e-5	3.940e-6	2.502e-7	1.576e-8	9.885e-10	6.190e-11
		1.974	1.988	1.995	1.997	1.999	
	$\alpha = 0.7$	6.450e-5	4.436e-6	2.936e-7	1.902e-8	1.216e-09	7.720e-11
		1.931	1.959	1.974	1.984	1.989	
$r = \frac{3-\alpha}{.95}$	$\alpha = 0.3$	1.085e-5	3.241e-7	8.953e-9	2.363e-10	6.058e-12	1.509e-13
		2.532	2.589	2.622	2.643	2.664	
	$\alpha = 0.5$	2.710e-5	1.057e-6	3.839e-8	1.337e-9	4.529e-11	1.517e-12
		2.340	2.392	2.422	2.442	2.450	
	$\alpha = 0.7$	3.962e-5	2.017e-6	9.638e-8	4.431e-9	1.986e-10	8.791e-12
		2.148	2.194	2.221	2.240	2.249	

TABLE 5.5
Alikhanov method applied to the initial-value test problem: maximum nodal errors (odd rows) and computational rates q in M^{-q} (even rows) for $r = 1$, $r = 2/\alpha$ and $r = (3 - \alpha)/\alpha$.

		$M = 2^6$	$M = 2^8$	$M = 2^{10}$	$M = 2^{12}$	$M = 2^{14}$	$M = 2^{16}$
$r = 1$	$\alpha = 0.3$	2.477e-2	1.634e-2	1.078e-2	7.115e-3	4.694e-3	3.097e-3
		0.300	0.300	0.300	0.300	0.300	
	$\alpha = 0.5$	1.164e-2	5.819e-3	2.909e-3	1.455e-3	7.273e-4	3.637e-4
		0.500	0.500	0.500	0.500	0.500	
	$\alpha = 0.7$	3.919e-3	1.485e-3	5.627e-4	2.132e-4	8.079e-5	3.061e-5
		0.700	0.700	0.700	0.700	0.700	
$r = \frac{2}{\alpha}$	$\alpha = 0.3$	5.865e-5	3.665e-6	2.291e-7	1.432e-8	8.949e-10	5.593e-11
		2.000	2.000	2.000	2.000	2.000	
	$\alpha = 0.5$	5.250e-5	3.281e-6	2.051e-7	1.282e-8	8.011e-10	5.007e-11
		2.000	2.000	2.000	2.000	2.000	
	$\alpha = 0.7$	4.232e-5	2.645e-6	1.653e-7	1.033e-8	6.458e-10	4.036e-11
		2.000	2.000	2.000	2.000	2.000	
$r = \frac{3-\alpha}{\alpha}$	$\alpha = 0.3$	5.505e-5	1.659e-6	4.472e-8	1.142e-9	2.833e-11	6.923e-13
		2.526	2.607	2.646	2.667	2.677	
	$\alpha = 0.5$	3.976e-5	1.379e-6	4.508e-8	1.439e-9	4.542e-11	1.425e-12
		2.425	2.467	2.485	2.493	2.497	
	$\alpha = 0.7$	3.425e-5	1.498e-6	6.307e-8	2.619e-9	1.083e-10	4.469e-12
		2.257	2.285	2.295	2.298	2.299	

Note that, similarly to Figure 5.1, we observed the pointwise behavior of the errors consistent with (4.5); see also [11, Figure 6.2] for similar graphs of pointwise errors of an L2-type method.

REFERENCES

[1] A. A. Alikhanov, *A new difference scheme for the time fractional diffusion equation*, J. Comput. Phys., 280 (2015), pp. 424–438.

[2] H. Brunner, *The numerical solution of weak singular Volterra integral equations by collocation on graded meshes*, Math. Comp., 45 (1985), pp. 417–437.

[3] H. Brunner, *Collocation methods for Volterra integral and related functional differential equations*, Cambridge University Press, Cambridge, UK, 2004.

[4] H. Chen and M. Stynes, *Error analysis of a second-order method on fitted meshes for a time-fractional diffusion problem*, J. Sci. Comput., 79 (2019), pp. 624–647.

[5] K. Diethelm, *The analysis of fractional differential equations*, Lecture Notes in Math., Springer-Verlag, Berlin, 2010.

[6] J. L. Gracia, E. O'Riordan and M. Stynes, *Convergence in positive time for a finite difference method applied to a fractional convection-diffusion problem*, Comput. Methods Appl. Math., 18 (2018), pp. 33–42

[7] B. Jin, R. Lazarov, and Z. Zhou, *Two fully discrete schemes for fractional diffusion and diffusion-wave equations with nonsmooth data*, SIAM J. Sci. Comput., 38 (2016), pp. A146–A170.

[8] B. Jin, R. Lazarov, and Z. Zhou, *An analysis of the $L1$ scheme for the subdiffusion equation with nonsmooth data*, IMA J. Numer. Anal., 36 (2016), pp. 197–221.

[9] B. Jin, R. Lazarov and Z. Zhou, *Numerical methods for time-fractional evolution equations with nonsmooth data: a concise overview*, Comput. Methods Appl. Mech. Engrg., 346 (2019), pp. 332–358.

[10] N. Kopteva, *Error analysis of the $L1$ method on graded and uniform meshes for a fractional-derivative problem in two and three dimensions*, Math. Comp., 88 (2019), pp. 2135–2155.

[11] N. Kopteva, *Error analysis of an L2-type method on graded meshes for a fractional-order parabolic problem*, Math. Comp., to appear.

[12] N. Kopteva, *Error analysis for time-fractional semilinear parabolic equations using upper and lower solutions*, preprint, arXiv:2001.04452, 2020, https://arxiv.org/abs/2001.04452.

[13] H.-L. Liao, D. Li, and J. Zhang, *Sharp error estimate of the nonuniform $L1$ formula for linear reaction-subdiffusion equations*, SIAM J. Numer. Anal., 56 (2018), pp. 1112–1133.

[14] H.-L. Liao, W. McLean, and J. Zhang, *A second-order scheme with nonuniform time steps for a linear reaction-subdiffusion problem*, preprint, arXiv:1803.09873v4, 2018, https://arxiv.org/abs/1803.09873v4.

[15] H.-L. Liao, W. McLean, and J. Zhang, *A discrete Grönwall inequality with application to numerical schemes for fractional reaction-subdiffusion problems*, SIAM J. Numer. Anal., 57 (2019), pp. 218–237.

[16] W. McLean and K. Mustapha, *A second-order accurate numerical method for a fractional wave equation*, Numer. Math., 105 (2007), pp. 481–510.

[17] 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.

[18] 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.

[19] M. Stynes, *Too much regularity may force too much uniqueness*, Fract. Calc. Appl. Anal., 19 (2016), pp. 1554–1562.

[20] M. Stynes, E. O'Riordan and J. L. Gracia, *Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation*, SIAM J. Numer. Anal., 55 (2017), pp. 1057–1079.