arXiv:1905.05070v2 [math.NA] 21 May 2019

ERROR ANALYSIS OF AN L2-TYPE METHOD ON GRADED MESHES FOR A FRACTIONAL-ORDER PARABOLIC PROBLEM*

NATALIA KOPTEVA†

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 behaviour at an initial time. An L2-type discrete fractional-derivative operator of order $3-\alpha$ is considered on nonuniform temporal meshes. Sufficient conditions for the inverse-monotonicity of this operator are established, which yields sharp pointwise-in-time error bounds on quasi-graded temporal meshes with arbitrary degree of grading. In particular, those results imply that milder (compared to the optimal) grading yields optimal convergence rates in positive time. Semi-discretizations in time and full discretizations are addressed. The theoretical findings are illustrated by numerical experiments.

Key words. fractional-order parabolic equation, L2 scheme, graded temporal mesh, arbitrary degree of grading, pointwise-in-time error bounds.

AMS subject classifications. 65M15, 65M60.

1. Introduction. The Caputo fractional derivative in time, which will be denoted by D_t^α, is defined [3] 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.

The paper is devoted to the analysis of an L2-type discrete fractional-derivative operator for D_t^α from [10], based on piecewise-quadratic Lagrange interpolants. In [10], this operator is analysed on uniform temporal meshes, and the optimal convergence order $3-\alpha$ in time is established under strong regularity assumptions on the exact solution. (Similar L2-type discretizations of order $3-\alpha$ on uniform temporal meshes were considered, e.g., in articles [4, 13], the latter giving optimal error bounds in positive time taking into account more realistic low regularity of the exact solution.)

The purpose of this paper is consider this discrete fractional-derivative operator on more general quasi-graded temporal meshes. For this, we employ the framework from the recent paper [9] (which builds on the analysis of [8], and, to some degree, [2]). This approach is based on barrier functions for derivation of subtle stability properties, and allows, in a relatively simple way, to get sharp pointwise-in-time error bounds on quasi-graded temporal meshes with arbitrary degree of grading.

- However, compared to the two methods considered in [9], the L1 scheme and the Alikhanov L2-1_σ scheme, now we have a significantly more challenging case, as the considered discrete fractional-derivative operator is not associated with an M-matrix. So our main challenge in this paper will be to establish the inverse-monotonicity of the discrete operator on nonuniform meshes.
- For the same reason, the generalization of our error analysis to the parabolic case also becomes substantially more challenging.

Note that the inverse-monotonicity on uniform temporal meshes was established in [10]. However, the evaluations in the latter article are quite intricate, so it is not

*This research was supported by Science Foundation Ireland Grant SFI/12/IA/1683.
†Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland (natalia.kopteva@ul.ie).

clear whether they can be generalized to more general meshes. We take a very different route and employ a non-standard set of basis functions (see Fig. 2.1), which very naturally leads to a representation of the discrete operator as a product of two matrices. Then relatively simple sufficient inverse-monotoncity conditions are formulated and the required stability properties of the discrete fractional-derivative operator are established, which enables us to employ the error analysis framework from [9].

This error analysis will be applied for the fractional-order parabolic problem

$$\begin{aligned} &D_t^\alpha u + \mathcal{L}u = f(x,t) \quad \text{for } (x,t) \in \Omega \times (0,T], \\ &u(x,t) = 0 \quad \text{for } (x,t) \in \partial\Omega \times (0,T], \qquad u(x,0) = u_0(x) \quad \text{for } x \in \Omega. \end{aligned} \tag{1.2}$$

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 defined by

$$\mathcal{L}u := \sum_{k=1}^{d} \bigl\{-\partial_{x_k}(a_k(x)\,\partial_{x_k}u)\bigr\} + c(x)\,u, \tag{1.3}$$

with sufficiently smooth coefficients $\{a_k\}$ and c in $C(\bar\Omega)$, for which we assume that $a_k > 0$ and $c \ge 0$ in $\bar\Omega$.

The L2-type fractional-derivative operator that we consider, denoted δ_t^α, is defined as follows. On the temporal mesh $0 = t_0 < t_1 < \ldots < t_M = T$, $\forall\, m = 1, \ldots, M$ let

$$\delta_t^\alpha U^m := D_t^\alpha(\Pi^m U)(t_m), \qquad \Pi^m := \begin{cases} \Pi_{1,1} & \text{on } (0,t_1) \quad \text{for } m = 1, \\ \Pi_{2,j} & \text{on } (t_{j-1},t_j) \quad \text{for } 1 \le j < m, \\ \Pi_{2,j-1} & \text{on } (t_{j-1},t_j) \quad \text{for } j = m > 1, \end{cases} \tag{1.4a}$$

where $\Pi_{1,j}$ and $\Pi_{2,j}$ are the standard linear and quadratic Lagrange interpolation operators with the following interpolation points:

$$\Pi_{1,j} \;:\; \{t_{j-1}, t_j\}, \qquad\qquad \Pi_{2,j} \;:\; \{t_{j-1}, t_j, t_{j+1}\}. \tag{1.4b}$$

Similarly to [12, 8, 2], 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 (1.2). It should be noted that these three papers are concerned with global-in-time error bounds on graded meshes. There is also a lot of interest in the literature in optimal error bounds in positive time on uniform meshes; see, e.g. [5, 7, 8]. By contrast, here, following the recent paper [9], 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. In particular, our results imply that milder (compared to the optimal) grading yields optimal convergence rates in positive time; see Remarks 4.2 and 4.3.

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 problem (1.2), 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 [6, Table 1.1]). Indeed, [11, Theorem 2.1] shows that if a solution u of (1.2) 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 behaviour at $t = 0$.

Outline. Sufficient conditions for inverse-monotonicity of the discrete fractional-derivative operator are established in §2, which enables us to establish its stability properties on quasi-graded meshes in §3. Error analysis for a simplest example without spatial derivatives is given in §4, while semi-discretizations in time and full discretizations for the parabolic case are addressed in §5. Finally, our theoretical findings are illustrated by numerical experiments in §6.

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 positive 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 $k \ge 0$, we shall use the standard norms in the space $L_2(\Omega)$ and the related Sobolev spaces $W_2^k(\Omega)$, while $H_0^1(\Omega)$ is the standard space of functions in $W_2^1(\Omega)$ vanishing on $\partial\Omega$.

2. Inverse-monotonicity of the discrete fractional-derivative operator. In this section we shall establish sufficient conditions on the temporal mesh $\{t_j\}_{j=0}^M$ for the inverse-monotonicity of the discrete fractional-derivative operator δ_t^α. The latter is understood in the sense that the matrix associated with δ_t^α is inverse-monotone, i.e. all elements of the inverse of this matrix are non-negative.

The following notation for the temporal mesh will be used throughout the paper:

$$\tau_j := t_j - t_{j-1}, \quad \tilde\tau_j := \tfrac12(\tau_{j-1}+\tau_j) \quad \rho_j := \frac{\tau_j}{\tau_{j-1}} \quad \sigma_j := \frac{\tau_j-\tau_{j-1}}{\tau_j+\tau_{j-1}} = 1 - \frac{2}{1+\rho_j}. \tag{2.1}$$

2.1. Matrix product representation for the discrete fractional-derivative operator. Given a set of real numbers $\{\beta_j\}_{j=0}^M$ such that $\beta_j \in [0,1)$ and $\beta_0 = 0$, our first task will be to find a representation for δ_t^α in the form

$$\delta_t^\alpha U^m = \sum_{j=0}^m \kappa_{m,j} V^j \quad \forall\, m \ge 1, \quad \text{where} \quad V^j := \frac{U^j - \beta_j U^{j-1}}{1-\beta_j} \quad \forall\, j \ge 1, \quad V^0 := U^0, \tag{2.2a}$$

and then establish sufficient conditions for choosing a set $\{\beta_j\}$ such that

$$\kappa_{m,m} > 0 \quad \text{and} \quad \sum_{j=0}^m \kappa_{m,j} = 0 \quad \forall\, m \ge 1, \qquad \kappa_{m,j} \le 0 \quad \forall\, 0 \le j < m \le M. \tag{2.2b}$$

*Remark 2.1 (*Inverse monotonicity*). Set $F^m := \delta_t^\alpha U^m$ for $m = 1, \ldots, M$ and augment these equations by $F^0 = U^0$. Now* (2.2) *yields the representation $\vec F = A_1 \vec V$ with $\vec V = A_2 \vec U$, or simply $\vec F = A_1 A_2 \vec U$, where A_1 and A_2 are $(M+1)\times(M+1)$ matrices, and the notation of type $\vec U := \{U^j\}_{j=0}^M$ is used for the corresponding column vectors. Being M-matrices (i.e. diagonally dominant, with non-positive off-diagonal elements), both A_1 and A_2 are inverse-monotone, hence the product A_1A_2 is also inverse-monotone (i.e. the elements of its inverse are non-negative). Thus* (2.2) *implies that the operator δ_t^α is associated with an inverse-monotone matrix.*

To describe a representation of type (2.2a) in a simple way on an arbitrary temporary mesh, we shall employ a non-standard basis $\{\Phi^j(t_k)\}_{j=0}^M$ for functions in $\mathbb{R}^{M+1}$ associated with the mesh $\{t_k\}_{k=0}^M$, which is defined by

$$\Phi^j(t_k) := 0 \quad \text{for } k \le j-1, \qquad \Phi^j(t_j) := 1, \qquad \Phi^j(t_k) := \beta_k \Phi^j(t_{k-1}) \quad \text{for } k \ge j+1 \tag{2.3}$$

(see Fig. 2.1 (left)).

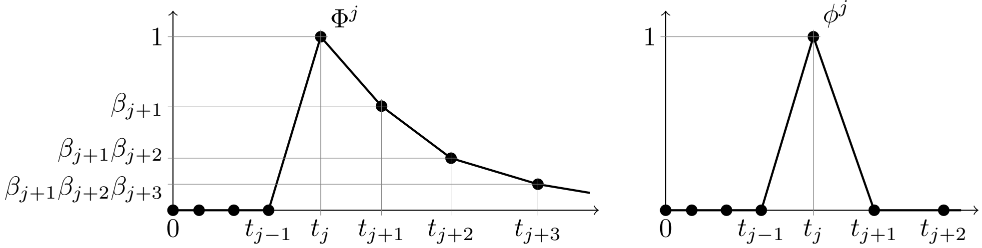

FIG. 2.1. *Non-standard basis* $\{\Phi^j\}$ *from* (2.3) *(left), hat-function basis* $\{\phi^j\}$ *(right).*

LEMMA 2.2. *Given a set* $\{\beta_j\}_{j=1}^M$ *with* $\beta_j \in [0,1)$ *and the basis* (2.3)*, the coefficients* $\kappa_{m,j}$ *in* (2.2a) *are described by*

$$\frac{\kappa_{m,j}}{1-\beta_j} = D_t^\alpha(\Pi^m \Phi^j)(t_m) \qquad \forall\, 0 \le j \le m \le M. \tag{2.4}$$

Proof. The definition of $\{V^j\}$ in (2.2a) is equivalent to the following basis expansion of $\{U^j\}$:

$$U^k = \sum_{j=0}^{M} V^j(1-\beta_j)\,\Phi^j(t_k) \qquad \forall\, k = 0, \dots M. \tag{2.5}$$

Indeed, by (2.3), for $k = 0$ this yields $U^0 = V^0(1-\beta_0) = V^0$, while for $k \ge 1$, in view of $\Phi^j(t_k) = 0$ for $j > k$, one can replace $\sum_{j=0}^M$ in (2.5) by $\sum_{j=0}^k$, so, indeed,

$$U^k = \underbrace{\sum_{j=0}^{k-1} V^j(1-\beta_j) \underbrace{\Phi^j(t_k)}_{=\beta_k \Phi^j(t_{k-1})}}_{=\beta_k U^{k-1}} + V^k(1-\beta_k) = \beta_k U^{k-1} + (1-\beta_k)V^k.$$

Next, (2.5) immediately implies that $\Pi^m U = \sum_{j=0}^M V^j(1-\beta_j)\Pi^m\Phi^j$ on $(0, t_m)$, where $\Pi^m \Phi^j = 0$ for $j > m$, so

$$\delta_t^m U^m = D_t^\alpha(\Pi^m U)(t_m) = \sum_{j=0}^{m} V^j(1-\beta_j)\, D_t^\alpha(\Pi^m\Phi^j)(t_m),$$

which, compared with (2.2a), immediately yields (2.4). ∎

It will be convenient to formulate sufficient conditions for (2.2b) in terms of the standard hat-function basis $\{\phi^j(t_k)\}_{j=0}^M$ for functions in $\mathbb{R}^{M+1}$ associated with the mesh $\{t_k\}_{k=0}^M$, i.e. $\phi^j(t_k)$ equals 1 if $k = j$ and 0 otherwise (see Fig. 2.1 (right)).

LEMMA 2.3. *Let the temporal mesh satisfy* $\rho_j \ge \rho_{j+1} \ge 1\ \forall j \ge 2$. *Then representation* (2.2a) *satisfies* (2.2b) *if*

$$\delta_t^\alpha \phi^{m-1}(t_m) + \beta_m \delta_t^\alpha \phi^m(t_m) < 0 \qquad \text{for } m \ge 1, \tag{2.6a}$$

$$\delta_t^\alpha \phi^{m-2}(t_m) + \beta_{m-1}\Big[\delta_t^\alpha \phi^{m-1}(t_m) + \beta_m \delta_t^\alpha \phi^m(t_m)\Big] \le 0 \qquad \text{for } m \ge 2, \tag{2.6b}$$

where $\delta_t^\alpha \phi^k(t_m)\ \forall\, k$ *is understood as* $D_t^\alpha(\Pi^m \phi^k)(t_m)$*. Under the above conditions we also have*

$$-\,\kappa_{m,0} \gtrsim t_m^{-\alpha} \qquad \textit{for}\ \ m \geq 3. \tag{2.7}$$

Proof. First, by (2.2a), note that $V^j = 1\ \forall\, j$ implies that $U^j = 1\ \forall\, j$, which then, by (1.4), implies that $\delta_t^\alpha U^m = 0\ \forall\, m \geq 1$, so one gets $0 = \sum_{j=0}^m \kappa_{m,j} \cdot 1\ \forall\, m \geq 1$, which immediately yields the second relation in (2.2b).

Next, by (2.4) combined with $1 - \beta_j > 0\ \forall\, j$, we conclude that $\kappa_{m,j} \leq 0\ \forall\, j < m$ is equivalent to $D_t^\alpha(\Pi^m \Phi^j)(t_m) \leq 0\ \forall\, j < m$. To find sufficient conditions for the latter, note that (2.3) implies that

$$\Phi^m(t_k) = \phi^m(t_k) \quad \forall\, k \leq m, \qquad \Phi^j(t_k) = \phi^j(t_k) + \beta_{j+1}\Phi^{j+1}(t_k) \quad \forall\, j, k \geq 0. \tag{2.8}$$

In particular, $\forall\, t_k \leq t_m$ one has $\Phi^{m-1}(t_k) = \phi^{m-1}(t_k) + \beta_m \phi^m(t_k)$ and $\Phi^{m-2}(t_k) = \phi^{m-2}(t_k) + \beta_{m-1}\Phi^{m-1}(t_k)$, so conditions (2.6a) and (2.6b) are respectively equivalent to $D_t^\alpha(\Pi^m \Phi^{m-1})(t_m) < 0$ and $D_t^\alpha(\Pi^m \Phi^{m-2})(t_m) \leq 0$. Once the latter two inequalities hold true, an argument by induction shows that for $D_t^\alpha(\Pi^m \Phi^j)(t_m) \leq 0$ $\forall\, j \leq m-3$ it suffices to check that $\delta_t^\alpha \phi^j(t_m) \leq 0\ \forall\, j \leq m-3$. The latter is true under the condition $\rho_j \geq \rho_{j+1} \geq 1\ \forall\, j \geq 2$, by [2, Lemma 4] (see also Remark 2.4).

To complete the proof of (2.2b), note that one can replace $\kappa_{m,m} > 0$ in (2.2b) by $\kappa_{m,m-1} < 0$, the latter being satisfied due to the strict inequality in (2.6a).

For (2.7), let $m \geq 3$ and note that the above argument, in particular the second relation in (2.8) with $j = 0$, implies that $\delta_t^\alpha \Phi^0(t_m) \leq \delta_t^\alpha \phi^0(t_m) \simeq -t_m^{-\alpha}$ (where we also used $\delta_t^\alpha \phi^0(t_m) = D_t^\alpha(\Pi^m \phi^0)(t_m) \simeq -t_m^{-\alpha}$, which can be shown on an arbitrary mesh from (1.4)). Combining this bound with (2.4) immediately yields (2.7). ☐

Remark 2.4. In the statement of Lemma 2.3, the assumption that $\rho_j \geq \rho_{j+1} \geq 1$ $\forall\, j \geq 2$ *is only required for* $\delta_t^\alpha \phi^j(t_m) \leq 0\ \ \forall j \leq m-3$*. For the latter we use [2, Lemma 4], which is obtained for the Alikhanov scheme, but we rely on the fact that if, using the notation of [2],* $\sigma = 1$*, then the coefficients* $\kappa^*_{m,j}$ *in the representation of type* $\delta_t^\alpha U^m = \sum_{j=0}^m \kappa^*_{m,j} U^j$ *are the same for the Alikhanov scheme and our scheme* $\forall\, j \leq m-3$*, and, furthermore,* $\kappa^*_{m,j} = \delta_t^\alpha \phi^j(t_m)$*. Note also that the above assumption on* $\{\rho_j\}$ *may be replaced by a weaker assumption; see [2, (12), (16) and Remark 3].*

It is convenient to rewrite conditions (2.6) using the notation

$$\mathcal{A}_m := \tilde\tau_m^\alpha\, \Gamma(1-\alpha)\, 2^\alpha\, \delta_t^\alpha \phi^m(t_m), \tag{2.9a}$$

$$-\mathcal{B}_m := \tilde\tau_m^\alpha\, \Gamma(1-\alpha)\, 2^\alpha\, \delta_t^\alpha \phi^{m-1}(t_m) \qquad \text{for}\ \ m \geq 1, \tag{2.9b}$$

$$\mathcal{F}_m := \tilde\tau_m^\alpha\, \Gamma(1-\alpha)\, 2^\alpha\, \delta_t^\alpha [\phi^{m-2} + \phi^{m-1} + \phi^m](t_m) \qquad \text{for}\ \ m \geq 2, \tag{2.9c}$$

where $\tilde\tau_1 := \tau_1$ and $\tilde\tau_m = \frac{1}{2}(\tau_{m-1} + \tau_m)$ for $m \geq 2$ is from (2.1).

Corollary 2.5. *Let the temporal mesh satisfy* $\rho_j \geq \rho_{j+1} \geq 1\ \forall\, j \geq 2$*. Then representation* (2.2a) *satisfies* (2.2b) *if*

$$\mathcal{B}_m - \beta_m \mathcal{A}_m > 0 \qquad \textit{for}\ \ m \geq 1, \tag{2.10a}$$

$$(\mathcal{B}_m - \mathcal{A}_m + \mathcal{F}_m) - \beta_{m-1}(\mathcal{B}_m - \beta_m \mathcal{A}_m) \leq 0 \qquad \textit{for}\ \ m \geq 2. \tag{2.10b}$$

Remark 2.6. Combining (2.4) *with* (2.8) *and* (2.9)*, from the proof of Lemma 2.3 one gets*

$$\Gamma(1-\alpha)\, 2^\alpha\, \frac{\kappa_{m,m}}{1-\beta_m} = \tilde\tau_m^{-\alpha}\, \mathcal{A}_m, \qquad \frac{\kappa_{m,m}}{1-\beta_m} \cdot \frac{1-\beta_{m-1}}{|\kappa_{m,m-1}|} = \frac{\mathcal{A}_m}{|\mathcal{B}_m - \beta_m \mathcal{A}_m|}\,. \tag{2.11}$$

2.2. Uniform temporal mesh. We shall first estimate the quantities in (2.9) and check the inverse-monotonicity conditions (2.10) for the case of uniform temporal meshes.

LEMMA 2.7 (Uniform temporal mesh). *Let* $\tau_j = \tau = TM^{-1}\ \forall j \geq 1$. *Then for the quantities in* (2.9) *one has*

$$\mathcal{A}_1 = \mathcal{B}_1 > 0; \qquad \mathcal{A}_m = \mathcal{A}, \qquad \mathcal{B}_m = \mathcal{B}' - \mathcal{B}''_m \geq \nu\mathcal{B}', \qquad \mathcal{F}_m \leq 1 + \mathcal{B}''_m \quad \forall\, m \geq 2, \tag{2.12}$$

where

$$\mathcal{A} := \frac{\alpha+2}{(1-\alpha)(2-\alpha)}, \quad \mathcal{B}' := \frac{4\alpha}{(1-\alpha)(2-\alpha)}, \quad 0 \leq \mathcal{B}''_m \leq \frac{\alpha}{24}, \quad \nu := 1 - \tfrac{1}{48}(1-\alpha). \tag{2.13}$$

Proof. For $m = 1$, we have $\Pi^m\phi^0(s) = 1 - s/t_1$ and $\Pi^m\phi^1(s) = s/t_1$ on $(0, t_1)$ (as here $\Pi^m = \Pi_{1,1}$), so $\delta_t^\alpha\phi^1(t_1) = -\delta_t^\alpha\phi^0(t_1) > 0$, so $\mathcal{A}_1 = \mathcal{B}_1 > 0$.

Now let $m \geq 2$ and combine (2.9) with (1.4) and (1.1). Rewriting the resulting integrals in terms of a new variable $\hat{s} := (s - t_{m-1})/\tau$, so the interval (t_{m-2}, t_m) is mapped to $(-1, 1)$, while $\tilde{\tau}_m^\alpha(t_m - s)^{-\alpha} = (1 - \hat{s})^{-\alpha}$, a calculation shows that

$$\mathcal{A}_m = 2^\alpha \int_{-1}^{1} (\hat{s} + \tfrac{1}{2})(1 - \hat{s})^{-\alpha}\, d\hat{s} = \mathcal{A}, \tag{2.14a}$$

$$\mathcal{B}_m = 2^\alpha \int_{-1}^{1} 2\hat{s}(1 - \hat{s})^{-\alpha}\, d\hat{s} - \mathcal{B}''_m = \mathcal{B}' - \mathcal{B}''_m. \tag{2.14b}$$

Here we used the observations that $\Pi^m\phi^m(\hat{s})$ is $\frac{1}{2}\hat{s}(\hat{s}+1)$ on $(-1,1)$ and vanishes otherwise, while $\Pi^m\phi^{m-1}(\hat{s})$ is $1 - \hat{s}^2$ on $(-1,1)$ and vanishes for $\hat{s} > 1$. For $m = 2$ one has $\mathcal{B}''_2 = 0$, while $\mathcal{B}''_m$ for $m > 2$ corresponds to $\Pi^m\phi^{m-1}(\hat{s}) = \frac{1}{2}(\hat{s}+1)(\hat{s}+2) < 0$ on $(-2,-1)$, so, using integration by parts on this interval, we arrive at

$$\mathcal{B}''_m := -\alpha 2^\alpha \int_{-2}^{-1} \underbrace{\Pi^m\phi^{m-1}(\hat{s})}_{<0}\, \underbrace{(1-\hat{s})^{-\alpha-1}}_{<2^{-\alpha-1}}\, d\hat{s} \leq -\alpha 2^{-1} \int_{-2}^{-1} \Pi^m\phi^{m-1}(\hat{s})\, d\hat{s} \leq \frac{\alpha}{24}\,, \tag{2.14c}$$

in view of $\int_{-2}^{-1} \Pi^m\phi^{m-1}(\hat{s})\, d\hat{s} = -\frac{1}{12}$. Note also that $\mathcal{B}''_m/\mathcal{B}' \leq \frac{1}{96}(1-\alpha)(2-\alpha) \leq 1-\nu$, so we get another desired assertion $\mathcal{B}' - \mathcal{B}''_m \geq \nu\mathcal{B}'$.

As to $\mathcal{F}_m$, set $\chi^{m-2} := \phi^{m-2} + \phi^{m-1} + \phi^m$ and note that $\chi^{m-2}(t_j)$ is 0 for $j < m-2$ and 1 for $j \geq m-2$. So for $m = 2$ one has $\chi^{m-2} = 1$ on $(0, t_m)$ so $\mathcal{F}_m = 0$. Otherwise $\frac{d}{d\hat{s}}\Pi^m\chi^{m-2}(\hat{s})$ has support on $(-2,-1)$ for $m = 3$ and on $(-3,-1)$ for $m > 3$, so we spit $\mathcal{F}_m = \mathcal{F}'_m + \mathcal{F}''_m$ with $\mathcal{F}''_3 = 0$ and

$$\mathcal{F}'_m := 2^\alpha \int_{-2}^{-1} \underbrace{\tfrac{d}{d\hat{s}}\Pi^m\chi^{m-2}(\hat{s})}_{>0}\, \underbrace{(1-\hat{s})^{-\alpha}}_{\leq 2^{-\alpha}}\, d\hat{s} \leq \int_{-2}^{-1} (\Pi^m\chi^{m-2})'(\hat{s})\, d\hat{s} = 1.$$

For $m > 3$, we also need to estimate $\mathcal{F}''_m$, which involves $\Pi^m\chi^{m-2}(\hat{s}) = \frac{1}{2}(\hat{s}+2)(\hat{s}+3)$ on $(-3,-2)$, and is bounded similarly to $\mathcal{B}''_m$ in (2.14c), which yields $0 \leq \mathcal{F}''_m \leq \mathcal{B}''_m$. Hence, we get the final assertion $\mathcal{F}_m \leq 1 + \mathcal{B}''_m$. ◻

COROLLARY 2.8 (Uniform temporal mesh). *Let* $\tau_j = \tau = TM^{-1}\ \forall j \geq 1$ *and, using the notation* (2.13), *set* $\beta_j := \beta := \frac{\theta}{2}\nu\mathcal{B}'/\mathcal{A}\ \forall\, j \geq 1$ *with any* $\theta \in [\frac{1}{2}, 1]$. *Then* $\beta \in (0, \frac{2}{3})$, *and the operator* δ_t^α *enjoys the inverse-monotone representation* (2.2).

Proof. By (2.13), one has $\beta=\theta\nu\frac{2\alpha}{\alpha+2}\in(0,\frac{2}{3})\ \forall\,\alpha\in(0,1),\ \forall\,\theta\in(0,1]$.

By Corollary 2.5, for (2.2) it suffices to check conditions (2.10). For $m\ge 1$ condition (2.10a) is straightforward in view of $\mathcal{A}_1=\mathcal{B}_1>0$ from (2.12). For $m\ge 2$, (2.12) yields $\mathcal{A}_m=\mathcal{A}$ and $\mathcal{B}_m-\mathcal{A}_m+\mathcal{F}_m\le\mathcal{B}'-\mathcal{A}+1$, while $\mathcal{B}_m\ge\nu\mathcal{B}'$ implies $\mathcal{B}_m-\beta_m\mathcal{A}_m=\mathcal{B}_m-\frac{\theta}{2}\nu\mathcal{B}'\ge(1-\frac{\theta}{2})\,\nu\mathcal{B}'>0$. So (2.10a) follows, while for (2.10b) it suffices to show that

$$(\mathcal{B}'-\mathcal{A}+1)-\beta\,(1-\tfrac{\theta}{2})\,\nu\mathcal{B}'<0.$$

Recall that $\beta=\frac{\theta}{2}\nu\mathcal{B}'/\mathcal{A}$, so multiplying the above inequality by $4\nu^{-2}\mathcal{A}/\mathcal{B}'^2$, one gets

$$\theta(2-\theta)>\underbrace{4\,(\mathcal{A}/\mathcal{B}')\left(1-\mathcal{A}/\mathcal{B}'+1/\mathcal{B}'\right)}_{=\frac14(\alpha+2)\ \text{ by (2.13)}}\cdot\nu^{-2}. \tag{2.15}$$

The latter, and hence (2.10b), is satisfied if

$$|\theta-1|<\sqrt{1-\tfrac14(\alpha+2)\,\nu^{-2}}\quad\Leftarrow\quad\theta\in(\theta_0(\alpha),1],\quad\theta_0(\alpha):=1-\sqrt{1-\tfrac14(\alpha+2)\,\nu^{-2}}.$$

Here $\theta_0(\alpha)<\frac12$ follows from $\alpha+2<3\,\nu^2\ \forall\,\alpha\in(0,1)$. ∎

2.3. General temporal meshes. Now we shall estimate the quantities in (2.9) and check the inverse-monotonicity conditions (2.10) for more general meshes.

LEMMA 2.9 (General temporal mesh). *Suppose that $\tau_j\le\tau_{j+1}\ \forall j\ge1$. Then for the quantities in* (2.9) *one has $\mathcal{A}_1=\mathcal{B}_1>0$ and $\forall\,m\ge2$*

$$\mathcal{A}_m=\mathcal{A}-\frac{\sigma_m}{2(1+\sigma_m)}\mathcal{B}'>\frac{2\mathcal{A}}{3},\qquad\mathcal{B}_m=\frac{\mathcal{B}'}{1-\sigma_m^2}-\mathcal{B}''_m\ge\nu\frac{\mathcal{B}'}{1-\sigma_m^2},\qquad\mathcal{F}_m\le1+\mathcal{B}''_m, \tag{2.16}$$

where we use the notation (2.13) *and $\sigma_m\in[0,1)$ from* (2.1).

Proof. We shall imitate the proof of Lemma 2.7 making appropriate changes for $m\ge2$. Rewrite all integrals in terms of the variable $\hat s:=(s-\frac12[t_{m-2}+t_m])/\tilde\tau_m$, so the interval (t_{m-2},t_m) is mapped to $(-1,1)$, but $s=t_{m-1}$ is now mapped to $\hat s=-\sigma_m$.

The evaluation of $\mathcal{A}_m$ is similar to (2.14a), but now (to ensure $\Pi^m\phi^m=0$ at $\hat s=-\sigma_m$) one has $\Pi^m\phi^m(s)=\frac12\hat s(\hat s+1)+\frac12(1-\hat s^2)\sigma_m/(1+\sigma_m)$ on $(-1,1)$, which yields the desired assertion for $\mathcal{A}_m$.

Next, similarly to (2.14b), split $\mathcal{B}_m=\mathcal{B}'_m-\mathcal{B}''_m$, where now $\Pi^m\phi^{m-1}(\hat s)=(1-\hat s^2)/(1-\sigma_m^2)$ on $(-1,1)$ (so that $\Pi^m\phi^{m-1}=1$ at $\hat s=-\sigma_m$), so we get a version of (2.14b) with $\mathcal{B}'$ replaced by $\mathcal{B}'_m=\mathcal{B}'/(1-\sigma_m^2)$. As to $\mathcal{B}''_m$ for $m>2$, it is estimated exactly as in (2.14c), only now the support of $\Pi^m\phi^{m-1}(\hat s)$ for $\hat s<-1$ is limited to a certain subset of $(\sigma_m-2,-1)$ (in view of $\tau_j\le\tau_{j+1}\ \forall j\ge1$), so $\int_{-2}^{-1}|\Pi^m\phi^{m-1}(\hat s)|d\hat s\le\frac{1}{12}$, which leads to the same upper bound for $\mathcal{B}''_m$ as in Lemma 2.7.

The estimation of $\mathcal{F}_m$ remains as in the proof of the proof of Lemma 2.7; in particular, we again enjoy $\mathcal{F}''_m\le\mathcal{B}''_m$ in view of $\tau_j\le\tau_{j+1}\ \forall j\ge1$.

Finally, $\mathcal{A}_m>\frac23\mathcal{A}$ for $m\ge2$ follows from $\frac{\sigma_m}{2(1+\sigma_m)}\le\frac14\ \forall\,\sigma_m\in[0,1)$ combined with the definitions of $\mathcal{A}$ and $\mathcal{B}'$ in (2.13). ∎

COROLLARY 2.10 (General temporal mesh). *Let the temporal mesh satisfy $\sigma_j\ge\sigma_{j+1}\ge0\ \forall\,j\ge2$, and for any $\theta\in[\frac12,1]$ set*

$$\eta(\sigma):=(1-\sigma^2)\big[\mathcal{A}/\mathcal{B}'-\tfrac{\sigma}{2(1+\sigma)}\big],\qquad\beta_1:=\beta_2,\qquad\beta_j:=\tfrac{\theta}{2}\nu/\eta(\sigma_j)\quad\forall\,j\ge2, \tag{2.17}$$

where we use the notation (2.13) *and* $\sigma_j \in [0,1)$ *from* (2.1). *Then* $\beta_j \ge \beta_{j+1} > 0$ $\forall\, j \ge 1$. *Furthermore, for any* $\theta \in [\frac12, 1]$ *there exists* $\bar\sigma = \bar\sigma(\alpha,\theta) \in (0,1)$ *such that if* $\sigma_j \in [0,\bar\sigma]$ $\forall j \ge 2$, *then* $\beta_j \in (0,1)$ $\forall j \ge 1$ *and the operator* δ_t^α *enjoys the inverse-monotone representation* (2.2).

Proof. Note that $\eta(\sigma) > 0\ \forall\,\sigma \in [0,1)$, in view of $\mathcal{A}_m > 0\ \forall\,\sigma_m \in [0,1)$ in (2.16). Hence $\beta_j > 0\ \forall j \ge 1$. Also η is a decreasing function of σ, so $\sigma_j \ge \sigma_{j+1} \ge 0\ \forall\, j \ge 2$ implies $\beta_j \ge \beta_{j+1} > 0\ \forall\, j \ge 1$.

Next, note that, by (2.1), $\sigma_j \ge \sigma_{j+1} \ge 0$ implies $\rho_j \ge \rho_{j+1} \ge 1\ \forall\, j \ge 2$. So, by Corollary 2.5, for (2.2) it suffices to check conditions (2.10). For $m \ge 1$ condition (2.10a) is straightforward in view of $\mathcal{A}_1 = \mathcal{B}_1 > 0$ (provided that $\beta_1 = \beta_2 < 1$, which will be shown below). For $m \ge 2$, (2.16) yields $\mathcal{A}_m = \eta(\sigma_m)\frac{\mathcal{B}'}{1-\sigma_m^2}$, so $\beta_m\mathcal{A}_m = \frac{\theta}{2}\nu\frac{\mathcal{B}'}{1-\sigma_m^2}$, while $\mathcal{B}_m \ge \nu\frac{\mathcal{B}'}{1-\sigma_m^2}$ implies $\mathcal{B}_m - \beta_m\mathcal{A}_m \ge (1-\frac{\theta}{2})\,\nu\frac{\mathcal{B}'}{1-\sigma_m^2} > 0$, so (2.10a) follows. For (2.10b) also using $\mathcal{B}_m - \mathcal{A}_m + \mathcal{F}_m \le [1-\eta(\sigma_m)]\frac{\mathcal{B}'}{1-\sigma_m^2} + 1$, we conclude that it suffices to show that

$$[1-\eta(\sigma_m)]\,\tfrac{\mathcal{B}'}{1-\sigma_m^2} + 1 - \underbrace{\beta_{m-1}}_{\ge\beta_m}(1-\tfrac{\theta}{2})\,\nu\,\tfrac{\mathcal{B}'}{1-\sigma_m^2} \le 0.$$

Dividing this by $\frac{\mathcal{B}'}{1-\sigma_m^2}$ and multiplying by $4\eta(\sigma_m)\,\nu^{-2}$, and also using $\beta_m = \frac{\theta}{2}\nu/\eta(\sigma_m)$, we find that (2.10b) is satisfied if

$$\theta(2-\theta) \ge 4\eta(\sigma_m)\Big(1-\eta(\sigma_m) + (1-\sigma_m^2)/\mathcal{B}'\Big)\cdot\nu^{-2}. \tag{2.18}$$

Comparing this to (2.15) and also noting that $\eta(0) = \mathcal{A}/\mathcal{B}'$, we see that if $\sigma_m = 0$, then a strict version of (2.18) becomes (2.15), so, as was shown in the proof of Corollary 2.8, it is satisfied $\forall\,\theta \in [\frac12, 1]$. Also, if $\sigma_m = 0$, then $\beta_m = \beta < \frac23$ (where β is defined in Corollary 2.8). Consequently, $\forall\,\theta \in [\frac12,1]$ there exists $\bar\sigma(\alpha,\theta) \in (0,1)$ such that both (2.18) and $\beta_m < 1$ are satisfied $\forall\, m \ge 2$ if $\sigma_m \in [0,\bar\sigma]$ $\forall\, m \ge 2$. (The computation of $\bar\sigma(\alpha,\theta)$ is discussed in Remark 2.12 below.) ∎

Remark 2.11. Under the conditions of Corollary 2.10, $\forall\, m \ge 3$, *one has*

$$\beta_m\frac{\kappa_{m,m}}{1-\beta_m}\cdot\frac{1-\beta_{m-1}}{|\kappa_{m,m-1}|} = \frac{\beta_m\mathcal{A}_m}{\mathcal{B}_m - \beta_m\mathcal{A}_m} \le \frac{\theta}{2-\theta}\,,$$

$$\frac{\kappa_{m,m}}{1-\beta_m}\cdot\frac{1-\beta_{m-1}}{\kappa_{m-1,m-1}} = \frac{\tilde\tau_m^{-\alpha}\,\mathcal{A}_m}{\tilde\tau_{m-1}^{-\alpha}\,\mathcal{A}_{m-1}} \ge \frac{\tilde\tau_{m-1}^{\alpha}}{\tilde\tau_m^{\alpha}}\,,$$

where we used (2.11) *and the observations on* $\beta_m\mathcal{A}_m$ *and* $\mathcal{B}_m - \beta_m\mathcal{A}_m$ *made in the proof of Corollary 2.10. For the second relation, we also relied on* $\{\mathcal{A}_m\}_{m=2}^M$ *being a decreasing function of* σ_m, *in view of* (2.16).

*Remark 2.12 (*Computation of $\bar\sigma$*). Using the notation* $\eta_m = \eta(\sigma_m)$, *one can rewrite* (2.18) *as*

$$4\eta_m\big(1+a-\eta_m\big) \le b, \quad \textit{where} \quad a := (1-\sigma_m^2)/\mathcal{B}' > 0, \quad b := \nu^2\theta(2-\theta) < 1, \tag{2.19}$$

which is equivalent to

$$2\eta_m \ge (1+a) + \sqrt{(1+a)^2 - b} > 1. \tag{2.20}$$

Importantly, this also ensures that $\beta_m < (2\eta_m)^{-1} < 1$. *Note that the remaining solutions of the quadratic inequality in* (2.19) *are described by*

$$2\eta_m \le (1+a) - \sqrt{(1+a)^2 - b} = \frac{b}{(1+a) + \sqrt{(1+a)^2 - b}} < \frac{\theta\nu(2-\theta)}{1 + \sqrt{1 - \theta(2-\theta)}} = \theta\nu,$$

which corresponds to $\theta\nu\beta_m^{-1} = 2\eta_m < \theta\nu$ *or* $\beta_m > 1$, *so such solutions are of no interest. Going back to* (2.20), *in which we use the definitions of* $\eta(\sigma)$ *from* (2.17) *and* a *from* (2.19), *we arrive at*

$$(1-\sigma_m^2)\big[(2\mathcal{A}-1)/\mathcal{B}' - \tfrac{\sigma_m}{(1+\sigma_m)}\big] \ge 1 + \sqrt{(1+a)^2 - b}.$$

Consequently, we impose $\sigma_m \in [0, \bar\sigma]$, *where* $\bar\sigma \in (0,1)$ *is the minimal solution of the equation (in which* g *is from* (2.19)*)*

$$\underbrace{(1-\bar\sigma)\big[c(1+\bar\sigma) - \bar\sigma\big]}_{=:g_L(\bar\sigma)} = \underbrace{1 + \sqrt{(1 + (1-\bar\sigma^2)/\mathcal{B}')^2 - b}}_{=:g_R(\bar\sigma)}, \quad c := \tfrac{2\mathcal{A}-1}{\mathcal{B}'} = \tfrac{2+5\alpha-\alpha^2}{4\alpha} > \tfrac{3}{2}.$$

Note that $g_L(\sigma)$ *is a parabola with zeros at* 1 *and* $-1 - \frac{1}{c-1}$, *so it is decreasing for positive* σ *while* $g_R(\sigma)$ *is also decreasing, so for each fixed* α *and* θ, *starting with* $\bar\sigma^{[0]} := 0$, *the iterative procedure* $g_L(\bar\sigma^{[q+1]}) = g_R(\bar\sigma^{[q]})$ *will generate an increasing sequence* $\bar\sigma^{[q]} \in (0,1)$ *converging to* $\bar\sigma$. *Finally, note that* $\theta = 1$ *will produce the least restrictive* $\bar\sigma$.

3. Stability properties for the discrete fractional-derivative operator. In this section we shall combine the inverse-monotonicity of the operator δ_t^α established in §2 with the barrier-function stability analysis developed in [9] for quasi-graded temporal meshes.

THEOREM 3.1 (Discrete comparison principle). *Let the temporal mesh satisfy* $\sigma_j \ge \sigma_{j+1} \ge 0\ \forall j \ge 2$. *There exists* $\bar\sigma = \bar\sigma(\alpha) \in (0,1)$ *such that if, additionally,* $\sigma_j \in [0,\bar\sigma]\ \forall j \ge 2$, *then the following statements are true.*

(i) If $U^0 \ge 0$ *and* $\delta_t^\alpha U^m \ge 0\ \forall m \ge 1$, *then* $U^j \ge 0$ *for* $\forall j \ge 0$.

(ii) If for a certain barrier function $\{B^j\}_{j=0}^M$ *one has* $|U^0| \le B^0$ *and* $|\delta_t^\alpha U^m| \le \delta_t^\alpha B^m\ \forall m \ge 1$, *then* $|U^j| \le B^j\ \forall j \ge 0$.

(iii) If $U^0 = 0$, *then* $|U^m| \lesssim \max_{j=1,\dots,m}\big\{t_j^\alpha\,|\delta_t^\alpha U^j|\big\}\ \forall m \ge 1$.

Proof. Let $\bar\sigma$ be from Corollary 2.10 (for any $\theta \in [\frac{1}{2}, 1]$, e.g., $\theta = 1$). Then the operator δ_t^α enjoys the inverse-monotone representation (2.2), which will play the crucial role in our proof.

(i) For $\{V^j\}$ from (2.2), one has $V^0 = U^0 \ge 0$, so $\delta_t^\alpha U^m \ge 0\ \forall m \ge 1$ implies $V^j \ge 0\ \forall j \ge 0$, from which we then conclude that $U^j \ge 0$ for $\forall j \ge 0$. (Alternatively, the proof may directly employ the inverse monotonicity of the matrix associated with δ_t^α; see Remark 2.1.)

(ii) As the operator δ_t^α is linear, the result follows from part (i).

(iii) For $\{V^j\}$ from (2.2), we claim that $|V^m| \lesssim \max_{j=1,\dots,m}\big\{t_j^\alpha\,|\delta_t^\alpha U^j|\big\}$. To show this, note that $V^0 = U^0 = 0$, so $|V^1| = \kappa_{1,1}^{-1}|\delta_t^\alpha U^1|$ and $|V^2| \le |V^1| + \kappa_{2,2}^{-1}|\delta_t^\alpha U^2|$, where, by (2.11), (2.16), $\kappa_{1,1} \simeq t_1^{-\alpha}$ and $\kappa_{2,2} \simeq \tilde\tau_2^{-\alpha} \simeq t_2^{-\alpha}$, so for $m = 1, 2$ the desired bound on $|V^m|$ follows. If $|V^n| = \max_{j\le m}|V^j|$ for some $3 \le n \le m$, then $\sum_{j=1}^n \kappa_{n,j}|V^n| \le |\delta_t^\alpha U^n|$, where $\sum_{j=1}^n \kappa_{n,j} = -\kappa_{n,0} \gtrsim t_n^{-\alpha}$, in view of (2.7), so again $|V^m| \le |V^n| \lesssim t_n^\alpha|\delta_t^\alpha U^n| \le \max_{j\le m}\big\{t_j^\alpha\,|\delta_t^\alpha U^j|\big\}$.

Next, a similar argument shows that if $\max_{j\le m}|U^j| = |U^k|$ for some $k \le m$, then $|U^k| \le |V^k|$. Consequently, $|U^m| \le |U^k| \lesssim \max_{j=1,\dots,k}\{t_j^\alpha\,|\delta_t^\alpha U^j|\}$. □

THEOREM 3.2 (Quasi-graded temporal grid). *Given $\gamma \in \mathbb{R}$, let the temporal mesh satisfy*

$$\tau_1 \simeq M^{-r}, \qquad \tau_j \simeq t_j/j, \qquad t_j \simeq \tau_1 j^r \qquad \forall\, j = 1,\dots,M \tag{3.1}$$

for some $1 \le r \le (3-\alpha)/\alpha$ if $\gamma > \alpha - 1$ or for some $r \ge 1$ if $\gamma \le \alpha - 1$. Additionally, let the temporal mesh satisfy $\sigma_j \ge \sigma_{j+1} \ge 0$ $\forall\, j \ge 2$ and $\sigma_j \in [0,\bar\sigma]$ $\forall j \ge K+1$, where $\bar\sigma \in (0,1)$ is from Theorem 3.1, and $1 \le K \lesssim 1$ (i.e. K is sufficiently large, but independent of M). Then for $\{U^j\}_{j=0}^M$ one has

$$\left.\begin{array}{c}|\delta_t^\alpha U^j| \lesssim (\tau_1/t_j)^{\gamma+1}\\ \forall j\ge 1,\quad U^0=0\end{array}\right\} \Rightarrow |U^j| \lesssim \mathcal{U}^j(\tau_1;\gamma) := \tau_1 t_j^{\alpha-1}\begin{cases}1 & \text{if } \gamma>0\\ 1+\ln(t_j/\tau_1) & \text{if } \gamma=0\\ (\tau_1/t_j)^\gamma & \text{if } \gamma<0\end{cases} \quad \forall j\ge 1. \tag{3.2}$$

Proof. (i) First, consider the case $K = 1$. If $1 \le r \le (3-\alpha)/\alpha$, note that mesh assumptions (3.1) are equivalent to those in [9, (2.1)], so the desired assertion is obtained by an application of Theorem 3.1(ii) with the barrier function $\{B^j\}$ from [9, proofs of Theorems 2.1(i) and 4.2(i)]. If $\gamma \le \alpha - 1$, then (3.2) can be shown (without assuming (3.1)) by an application of Theorem 3.1(iii) imitating the proof of [9, Theorem 2.1(ii)].

(ii) Next, consider the case $K > 1$. As $K \lesssim 1$, by (3.1), one has $\tau_j \simeq \tau_1$ $\forall j \le K$. So for $m \le K$ a calculation yields $|U^m| \lesssim \sum_{j=0}^{m-1}|U^j| + \tau_1^\alpha|\delta_t^\alpha U^m|$ (in particular, $\kappa_{m,m} \simeq \tau_1^{-\alpha}$ follows from (2.11), (2.16)). As $|\delta_t^\alpha U^m| \lesssim 1$, so one gets $|U^m| \lesssim \tau_1^\alpha \simeq \mathcal{U}^m$ $\forall\, m \le K$.

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 $|\delta_t^\alpha \mathring{U}^m| \lesssim 1$ for $m = K+1, K+2$. Consider $m > K+2$. By (1.4), 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}, 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 $t_{K+1} \simeq \tau_1$). Consequently, for $m \ge K+1$ 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, let $\mathring{\delta}_t^\alpha$ be the operator of type δ_t^α, but associated with the mesh $\{t_j\}_{j=K-1}^M$, i.e. for any $\{W^j\}_{j=K-1}^M$, set $\mathring{\delta}_t^\alpha W^K := \int_{t_{K-1}}^{t_K}(\Pi_{1,K}W)(t_K-s)^{-\alpha}ds$ and $\mathring{\delta}_t^\alpha W^m := \int_{t_{K-1}}^{t_m}(\Pi^m W)(t_m-s)^{-\alpha}ds$ for $m > K$. Then $\mathring{\delta}_t^\alpha \mathring{U}^K = 0$, while $|\mathring{\delta}_t^\alpha \mathring{U}^m| = |\delta_t^\alpha \mathring{U}^m|$ for $m > K$. Importantly, the bound of type (3.2), which we already proved for δ_t^α for the case $K = 1$, applies to $\mathring{\delta}_t^\alpha$. In the latter bound, $j \ge K$ and t_j is replaced by $t_j - t_{K-1} \simeq t_j$. In particular, we conclude that if $\gamma \le \alpha$, then $|\mathring{U}^j| \lesssim \mathcal{U}^j(\tau_1,\gamma)$, while if $\gamma > \alpha$, then $|\mathring{U}^j| \lesssim \mathcal{U}^j(\tau_1,\alpha) = \mathcal{U}^j(\tau_1,\gamma)$. Combining our findings, one gets $|U^j| = |\mathring{U}^j| \lesssim \mathcal{U}^j(\tau_1,\gamma)$ $\forall\, j \ge K+1$, and hence (3.2) $\forall\, j \ge 1$. □

COROLLARY 3.3 (Graded temporal grid). *Given $\gamma \in \mathbb{R}$, let the temporal mesh be defined by $\{t_j = T(j/M)^r\}_{j=0}^M$ for some $1 \le r \le (3-\alpha)/\alpha$ if $\gamma > \alpha - 1$ or for some $r \ge 1$ if $\gamma \le \alpha - 1$. Then the conditions of Theorem 3.2 on the mesh are satisfied, and so (3.2) holds true for any $\{U^j\}_{j=0}^M$ with $U^0 = 0$.*

Proof. Clearly, the mesh satisfies (3.1), as well as $\sigma_j \ge \sigma_{j+1} \ge 0$ $\forall\, j \ge 2$. So it remains to find $K \lesssim 1$ such that $\sigma_j \in [0,\bar\sigma]$ $\forall j \ge K+1$. For the latter, in view of (2.1), the sequence $\{\sigma_j\}$, as well as the related sequence $\{\rho_j\}$, is decreasing, so it suffices to satisfy

$$\rho_{K+1} = \frac{\tau_{K+1}}{\tau_K} = \frac{(K+1)^r - K^r}{K^r - (K-1)^r} = \frac{(1+1/K)^r - 1}{1-(1-1/K)^r} \le \bar\rho := \frac{2}{1-\bar\sigma} - 1. \tag{3.3}$$

As $\bar\sigma$ is independent of M, clearly, one can always choose such sufficiently large $K = K(r, \bar\sigma)$ independently of M. ∎

*Remark 3.4 (*Modified graded mesh*). Although, as shown by Corollary 3.3, the result of Theorem 3.2 applies to the standard graded mesh, but it may still be desirable for the operator δ_t^α to enjoy the inverse-monotonicity property of type* (2.2) $\forall\, j \ge 1$ *(rather than $\forall\, j \ge K+1$). This can be easily ensured by a simple modification of the graded scheme as follows. Let*

$$t_j := T\,\hat t_j/\hat t_M, \qquad \textit{where} \qquad \hat t_j := \big(\tfrac{j+K'}{M}\big)^r - \big(\tfrac{K'}{M}\big)^r, \qquad K' := K-1, \tag{3.4}$$

with K from (3.3)*. To compute $K = K(r,\bar\sigma)$, note that $\bar\sigma$ can be computed, as described in Remark 2.12. Note also that if $K=1$, one gets the standard graded mesh, while $K>1$ implies that $\hat t_M = (1+K'/M)^r - (K'/M)^r \approx 1 + rK'/M$. Clearly, Corollary 3.3 also applies to the modified graded mesh.*

*Remark 3.5 (*Inverse-monotone modification of δ_t^α*). Consider the standard graded temporal mesh $\{t_j = T(j/M)^r\}_{j=0}^M$ for some $r \ge 1$. As an alternative to modifying this mesh, as described in Remark 3.4, one can ensure the inverse-monotonicity* (2.2) *$\forall\, j \ge 1$ by tweaking the definition of δ_t^α in* (1.4) *for $m \le K$ only as follows. Reset $\Pi^m := \Pi_{1,j}$ on (t_{j-1}, t_j) $\forall j \le m \le K$ (i.e. the inverse-monotone L1 discretization is used for $m \le K$). With this modification, also reset $\beta_j := \beta_{K+1}$ $\forall\, j = 1, \ldots, K$ in* (2.17)*. Then all results of this paper, that are valid for the graded mesh, also hold true for the modified discrete fractional-derivative operator (as can be shown by only minor modifications in the relevant proofs).*

We finish this section with a more subtle version of Theorem 3.2, which will be useful when considering the fractional-derivative parabolic case in §5.

THEOREM 3.2*. *Let $\bar\sigma$ and the set $\{\beta_j\}_{j=1}^M$ be from Corollary 2.10 (for any $\theta \in [\frac12, 1]$), and $\{\kappa_{m,j}\}$ be the unique set of the coefficients in the corresponding representation* (2.2a) *for the operator δ_t^α. Also, given $\gamma \in \mathbb{R}$, let the temporal mesh satisfy the conditions of Theorem 3.2 with $K=1$. Then for $\{U^j\}_{j=0}^M$ and $\{W^j\}_{j=0}^M$ with $U^0 = W^0 = 0$ the following is true:*

$$\left\{\begin{array}{ll} \displaystyle\sum_{j=0}^m \kappa_{m,j} W^j \lesssim (\tau_1/t_m)^{\gamma+1} & \forall\, m \ge 1 \\ \displaystyle\frac{|U^j| - \beta_j |U^{j-1}|}{1-\beta_j} \lesssim W^j & \forall\, j \ge 1 \end{array}\right. \quad \Rightarrow \quad |U^j| \lesssim \mathcal{U}^j(\tau_1;\gamma), \tag{3.5}$$

where $\mathcal{U}^j$ is defined in (3.2)*.*

Proof. Note that the choice of $\bar\sigma$ and $\{\beta_j\}_{j=1}^M$ in Corollary 2.10 ensures that the corresponding representation (2.2a) for the operator δ_t^α satisfies (2.2b), i.e. δ_t^α is associated with an inverse-monotone matrix; see Remark 2.1. Using the notation of this remark, the assumptions in (3.5) become $A_1\vec W \lesssim \vec F$ and $A_2|\vec U| \lesssim \vec W$, where $F^m := (\tau_1/t_m)^{\gamma+1}$. As A_1 and A_2 are inverse-monotone, so $\vec V \lesssim A_1^{-1}\vec F$, and then $|\vec U| \le A_2^{-1}\vec W \lesssim A_2^{-1}A_1^{-1}\vec F$. On the other hand, Theorem 3.2 implies that $0 \le A_2^{-1}A_1^{-1}\vec F \lesssim \vec{\mathcal{U}}$, which yields the desired assertion. ∎

4. Error estimation for a simplest example (without spatial derivatives). Consider a fractional-derivative problem without spatial derivatives together with its discretization of type (1.4):

$$D_t^\alpha u(t) = f(t) \qquad \text{for } t \in (0,T], \qquad\qquad u(0) = u_0, \tag{4.1a}$$
$$\delta_t^\alpha U^m = f(t_m) \qquad \text{for } m = 1,\dots,M, \qquad\qquad U^0 = u_0. \tag{4.1b}$$

Throughout this subsection, with slight abuse of notation, ∂_t will be used for $\frac{d}{dt}$.

The main result of this section is the following theorem, to the proof of which we shall devote the remainder of the section.

THEOREM 4.1. *Let the temporal mesh satisfy* (3.1) *for some* $r \ge 1$*, and also* $\sigma_j \ge \sigma_{j+1} \ge 0\ \forall\, j \ge 2$ *and* $\sigma_j \in [0,\bar\sigma]\ \forall j \ge K+1$*, where* $\bar\sigma \in (0,1)$ *is from Theorem 3.1, and* $1 \le K \lesssim 1$*. Suppose that* u *and* $\{U^m\}$ *satisfy* (4.1)*, and* $|\partial_t^l u| \lesssim 1 + t^{\alpha-l}$ *for* $l = 1,3$ *and* $t \in (0,T]$*. Then* $\forall\, m = 1,\dots,M$ *one has*

$$|u(t_m) - U^m| \lesssim \mathcal{E}^m := \begin{cases} M^{-r}\, t_m^{\alpha-1} & \text{if } 1 \le r < 3-\alpha, \\ M^{\alpha-3}\, t_m^{\alpha-1}[1+\ln(t_m/t_1)] & \text{if } r = 3-\alpha, \\ M^{\alpha-3}\, t_m^{\alpha-(3-\alpha)/r} & \text{if } r > 3-\alpha \end{cases} \tag{4.2}$$

*Remark 4.2 (*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 4.3 (*Global convergence*). Note that* $\max_{m\ge1}\mathcal{E}^m \simeq \mathcal{E}^1 \simeq \tau_1^\alpha \simeq M^{-\alpha r}$ *for* $\alpha \le (3-\alpha)/r$*, while* $\max_{m\ge1}\mathcal{E}^m \simeq \mathcal{E}^M \simeq M^{\alpha-3}$ *otherwise. Consequently, Theorem 4.1 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* $r = (3-\alpha)/\alpha$*.*

Remark 4.4. Theorem 4.1 applies to the standard graded mesh $\{t_j = T(j/M)^r\}_{j=0}^M$ *for any* $r \ge 1$ *(in view of Corollary 3.3), as well as to the modified graded mesh* (3.4)*. Furthermore, the proof of this theorem can be easily extended to the case of the modified discrete fractional-derivative operator described in Remark 3.5.*

To prove Theorem 4.1, we first get an auxiliary result.

LEMMA 4.5 (Truncation error). *For a sufficiently smooth function* u*, let* $r^m := \delta_t^\alpha u(t_m) - D_t^\alpha u(t_m)\ \forall\, m \ge 1$*, and*

$$\psi^1 := \sup_{s\in(0,t_2)}\big(s^{1-\alpha}|\partial_s u(s)|\big) + t_2^{-\alpha}\operatorname{osc}\big(u,[0,t_2]\big), \tag{4.3a}$$
$$\psi^j := t_j^{3-\alpha}\sup_{s\in(t_{j-1},t_{j+1})}|\partial_s^3 u(s)| \qquad \forall\, 2 \le j \le M-1. \tag{4.3b}$$

Then, under conditions (3.1) *on the temporal mesh, one has*

$$|r^m| \lesssim (\tau_1/t_m)^{\min\{\alpha+1,\,(3-\alpha)/r\}} \max_{j\le\max\{1,m-1\}}\{\psi^j\} \qquad \forall\, m \ge 1. \tag{4.4}$$

Proof. We closely imitate the proof of [9, Lemma 4.7], so some details will be skipped here. From (1.4), recall that $\delta_t^\alpha u(t_m) = D_t^\alpha(\Pi^m u)(t_m)$. Next, recalling the definition (1.1) of D_t^α, with the auxiliary function $\chi := u - \Pi^m u$, we arrive at

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

Split the above integral to intervals $(0,t_1)$ and (t_1,t_m). On $(0,t_1)$ note that $\chi(t_1)=0$ implies $\chi(s)=-\int_s^{t_1}\chi'(\zeta)d\zeta$, where $|\chi'|\le|\partial_s u|+|\partial_s(\Pi^m u)|$, while $|\partial_s(\Pi^m u)|\lesssim t_2^{-1}\mathrm{osc}(u,[0,t_2])\le s^{\alpha-1}t_2^{-\alpha}\mathrm{osc}(u,[0,t_2])$ (in view of $\tau_1\simeq\tau_2$), so a calculation yields $|\chi(s)|\lesssim s^{\alpha-1}(t_1-s)\psi^1$. Next, on any (t_{j-1},t_j) for $1<j<m$ one has $|\chi|\lesssim\tau_j^3t_j^{\alpha-3}\psi^j$. Finally, on (t_{m-1},t_m), if $m>2$, then $|\chi|\lesssim\tau_m^2(t_m-s)t_m^{\alpha-3}\psi^{m-1}$, while if $m=2$, then we imitate the estimation on $(0,t_1)$ and again get $|\chi(s)|\lesssim s^{\alpha-1}(t_2-s)\psi^1\lesssim\tau_2^2(t_2-s)t_2^{\alpha-3}\psi^1$.

Combining our findings on χ, a calculation shows that we get the following version of [9, (4.8)]:

$$|r^m|\lesssim\mathring{\mathcal{J}}^m\,(\tau_1/t_m)^{\alpha+1}\,\psi^1+\mathcal{J}^m\max_{j=2,\dots,m}\bigl\{\nu_{m,j}(\tau_j/t_j)^{3-\alpha}(t_j/t_m)^{\alpha+1}\,\psi^{j^*}\bigr\}. \tag{4.5}$$

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 (4.5) is as follows:

$$\mathring{\mathcal{J}}^m:=(t_m/\tau_1)^{\alpha+1}\int_0^{t_1}s^{\alpha-1}(t_1-s)\,(t_m-s)^{-\alpha-1}ds\lesssim1,$$

$$\mathcal{J}^m:=\tau_m^{\alpha}\,t_m^{\alpha/r+1}\int_{t_1}^{t_m}s^{-\alpha/r-1}\,(t_m-s)^{-\alpha-1}\,\min\{1,(t_m-s)/\tau_m\}\,ds\lesssim1,$$

$$\nu_{m,j}:=(\tau_j/\tau_m)^{\alpha}\,(t_j/t_m)^{-\alpha(1-1/r)}\simeq1,$$

$$j^*:=\min\{j,m-1\}.$$

Here the bound on $\nu_{m,j}$ follows from $\tau_j/\tau_m\simeq(t_j/t_m)^{1-1/r}$ (in view of (3.1)). For the estimation of quantities of type $\mathring{\mathcal{J}}^m$ and $\mathcal{J}^m$, we refer the reader to [8]. 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 (3.1)).

Combining the above observations with (4.5) yields

$$|r^m|\lesssim\max_{j\le\max\{1,m-1\}}\Bigl\{\underbrace{(\tau_j/t_j)^{3-\alpha}}_{\simeq(\tau_1/t_j)^{(3-\alpha)/r}}\,(t_j/t_m)^{\alpha+1}\,\psi^j\Bigr\},$$

where we also used $\tau_j/t_j\simeq(\tau_1/t_j)^{1/r}$ (in view of (3.1)). The desired bound (4.4) follows as $\tau_1\le t_j\le t_m$. □

Proof of Theorem 4.1. Consider the error $e^m:=u(t_m)-U^m$, for which (4.1) implies $e^0=0$ and $\delta_t^\alpha e^m=r^m\ \forall m\ge1$, where the truncation error r^m is from Lemma 4.5 and hence satisfies (4.4). Furthermore, combining (4.3) with (3.1) yields $\psi^1\lesssim1$ (in view of $|\mathrm{osc}(u,[0,t_2])|\le\int_0^{t_2}|\partial_s u|\,ds\lesssim t_2^\alpha$) and $\psi^j\lesssim1$ for $j\ge2$ (in view of $s\simeq t_j$ for $s\in(t_{j-1},t_{j+1})$ for this case). Consequently, we arrive at

$$|r^m|\lesssim(\tau/t_m)^{\gamma+1}\quad\forall\,m\ge1,\qquad\text{where }\ \gamma+1:=\min\{\alpha+1,(3-\alpha)/r\}. \tag{4.6}$$

Next we apply (3.2) from Theorem 3.2 to bound $e^m=u(t_m)-U^m$. Consider three cases.

Case $1\le r<3-\alpha$. Then both $(3-\alpha)/r>1$ and $\alpha+1>1$, so $\gamma>0$. An application of (3.2) for this case yields $|e^m|\lesssim\tau_1\,t_m^{\alpha-1}$, where $\tau_1\simeq M^{-r}$.

Case $r = 3-\alpha$. Then $(3-\alpha)/r = 1$, while $\alpha+1 > 1$, so $\gamma = 0$. An application of (3.2) yields $|e^m| \lesssim \tau_1\, t_m^{\alpha-1}[1+\ln(t_m/t_1)]$, where $\tau_1 \simeq M^{-r} = M^{\alpha-3}$.

Case $r > 3-\alpha$. Then $(3-\alpha)/r < 1$, while $\alpha+1 > 1$, so $\gamma+1 = (3-\alpha)/r < 1$. An application of (3.2) (where, importantly, unless $r \le (3-\alpha)/\alpha$, one has $\gamma \ge \alpha-1$) yields $|e^m| \lesssim \tau_1\, t_m^{\alpha-1}(\tau_1/t_m)^{(3-\alpha)/r-1} \simeq \tau_1^{(3-\alpha)/r} t_m^{\alpha-(3-\alpha)/r}$, where $\tau_1^{(3-\alpha)/r} \simeq M^{\alpha-3}$. $\square$

5. Error analysis for the parabolic case. In this section, we shall generalize the analysis of §4 to problems with variable coefficients and spatial derivatives. Both semidiscretizations in time and fully discrete methods will be addressed.

5.1. Error analysis for semidiscretizations in time. Consider the semidiscretization of our problem (1.2) in time using the discrete fractional-derivative operator δ_t^α from (1.4):

$$\delta_t^\alpha U^j + \mathcal{L}U^j = f(\cdot,t_j) \;\text{ in }\Omega,\quad U^j = 0 \;\text{ on }\partial\Omega \quad \forall\, j = 1,\dots,M; \qquad U^0 = u_0. \tag{5.1}$$

LEMMA 5.1 (Stability for parabolic case). *Given $\gamma \in \mathbb{R}$, let the temporal mesh satisfy* (3.1) *for some $1 \le r \le (3-\alpha)/\alpha$ if $\gamma > \alpha-1$ or for some $r \ge 1$ if $\gamma \le \alpha-1$. There exists $\bar\sigma^* = \bar\sigma^*(\alpha) \in (0,1)$ such that if, additionally, the temporal mesh satisfies $\sigma_j \ge \sigma_{j+1} \ge 0$ $\forall\, j \ge 2$ and $\sigma_j \in [0,\bar\sigma^*]$ $\forall j \ge K+1$, where $1 \le K \lesssim 1$ (i.e. K is sufficiently large, but independent of M), then for $\{U^j\}_{j=0}^M$ from* (5.1) *one has*

$$\left.\begin{aligned} \|f(\cdot,t_j)\|_{L_2(\Omega)} \lesssim (\tau_1/t_j)^{\gamma+1} \\ \forall j \ge 1, \quad U^0 = 0 \;\text{ in }\bar\Omega \end{aligned}\right\} \quad\Rightarrow\quad \|U^j\|_{L_2(\Omega)} \lesssim \mathcal{U}^j(\tau_1;\gamma) \;\; \forall j \ge 1, \tag{5.2}$$

where $\mathcal{U}^j$ is defined in (3.2).

Proof. Fix any $\theta \in [\frac12, 0)$ and let $\bar\sigma = \bar\sigma(\alpha,\theta)$ be from Corollary 2.10.

(i) First, we shall prove that there exists $\bar\sigma^* \in (0,\bar\sigma]$ such that $\sigma_j \in [0,\bar\sigma^*]$ $\forall j \ge K+1$ implies

$$\left(|\kappa_{m,m-1}|^{-1}\frac{\beta_m}{1-\beta_m}\right)^2 \le \frac{\kappa_{m,m}^{-1}}{1-\beta_m}\cdot\frac{\kappa_{m-1,m-1}^{-1}}{1-\beta_{m-1}} \qquad \forall m \ge K+2, \tag{5.3}$$

where β_m are defined by (2.17) for $m > K$ and are equal to β_{K+1} for $m \le K$, while $\{\kappa_{m,j}\}$ is the unique set of the coefficients in the corresponding representation (2.2a) for the operator δ_t^α. To check this, rewrite (5.3) as

$$\left(\beta_m\frac{\kappa_{m,m}}{1-\beta_m}\cdot\frac{1-\beta_{m-1}}{|\kappa_{m,m-1}|}\right)^2 \le \frac{\kappa_{m,m}}{1-\beta_m}\cdot\frac{1-\beta_{m-1}}{\kappa_{m-1,m-1}} \;\Leftarrow\; \left(\frac{\theta}{2-\theta}\right)^{2/\alpha} \le \frac{\tilde\tau_{m-1}}{\tilde\tau_m} \;\forall m \ge K+2,$$

where the implication follows from Remark 2.11. The sequence $\{\sigma_j\}$ is decreasing, and hence, in view of (2.1), the related sequence $\{\rho_j\}$ is also decreasing, so it suffices to check that

$$\frac{\tilde\tau_{K+2}}{\tilde\tau_{K+1}} = \frac{\rho_{K+1}\tau_K + \rho_{K+2}\tau_{K+1}}{\tau_K + \tau_{K+1}} \le \rho_{K+1} \le \bar\rho^* := \left(\frac{2-\theta}{\theta}\right)^{2/\alpha} \in (1, 3^{2/\alpha}].$$

From this, $\bar\sigma^* := \min\{\bar\sigma,\, 1+\frac{2}{1-\bar\rho^*}\} > 0$ will yield (5.3).

(ii) Next, suppose that $K = 1$, i.e. $\sigma_j \in [0,\bar\sigma^*]$ $\forall j \ge 2$. Then (5.3) holds true $\forall\, m \ge 3$, while, in view of Corollary 2.10, $\bar\sigma^* \le \bar\sigma$ implies that the operator δ_t^α

enjoys the inverse-monotone representation (2.2). Now, using (2.2a) and the notation $V^m = \frac{1}{1-\beta_m}U^m - \frac{\beta_m}{1-\beta_m}U^{m-1}$ and $f^m := f(\cdot, t_m)$, we can rewrite (5.1) as

$$\kappa_{m,m}V^m + \mathcal{L}U^m = |\kappa_{m,m-1}|V^{m-1} + \sum_{j=1}^{m-2}|\kappa_{m,j}|V^j + f^m.$$

Consider the inner product of the above and V^m using the notation

$$w^m := \sqrt{\|V^m\|^2_{L_2(\Omega)} + \tfrac{\kappa^{-1}_{m,m}}{1-\beta_m}\langle \mathcal{L}U^m, U^m\rangle}.$$

Then

$$\kappa_{m,m}(w^m)^2 =$$
$$= \underbrace{|\kappa_{m,m-1}|\langle V^{m-1}, V^m\rangle + \frac{\beta^m}{1-\beta_m}\langle \mathcal{L}U^m, U^{m-1}\rangle}_{=:|\kappa_{m,m-1}|\,Q^m} + \sum_{j=1}^{m-2}|\kappa_{m,j}|\underbrace{\langle V^m, V^j\rangle}_{\le w^m w^j} + \underbrace{\langle V^m, f^m\rangle}_{\le w^m\|f^m\|_{L_2(\Omega)}}.$$

Here $Q^1 = 0$ in view of $U^0 = V^0 = 0$ and $Q^m \le w^m w^{m-1}$ $\forall m \ge 3$ in view of (5.3). For $m = 2$ there is a sufficiently large constant $1 \le \bar{C} \lesssim 1$ such that $Q^2 \le \bar{C}w^2w^1$. (For example, using a version of Remark 2.11 for $m = 2$ and imitating the argument in part (i), one can choose $\bar{C} = (\tilde{\tau}_2^{-\alpha}\,\mathcal{A}_2)/(\tau_1^{-\alpha}\,\mathcal{A}_1)$; see also (2.9a) and (3.1).) Now dividing by w^m and recalling that, by (2.2b), $\kappa_{m,j} \le 0$ $\forall j < m$, we get

$$\kappa_{1,1}w^1 \le \|f^1\|_{L_2(\Omega)}, \qquad \kappa_{2,2}w^2 + \kappa_{2,1}(\bar{C}w^1) \le \|f^2\|_{L_2(\Omega)}, \qquad \sum_{j=1}^{m}\kappa_{m,j}w^m \le \|f^m\|_{L_2(\Omega)}.$$

Set $W^1 := \bar{C}w^1$ and $W^j := w^j$ otherwise. Then, in view of $\bar{C} \ge 1$ and $\kappa_{m,1} \le 1$ $\forall m \ge 3$, we arrive at $\sum_{j=1}^m \kappa_{m,j}W^j \lesssim \|f^m\|_{L_2(\Omega)} \lesssim (\tau_1/t_m)^{\gamma+1}$ $\forall m \ge 1$, while $W^0 = 0$. Note also that $\frac{1}{1-\beta_j}\|U^j\|_{L_2(\Omega)} - \frac{\beta_j}{1-\beta_j}\|U^{j-1}\|_{L_2(\Omega)} \le \|V^j\|_{L_2(\Omega)} \le w^j \le W^j$ $\forall j \ge 1$. Thus we conclude that the assumptions in (3.5) are satisfied with $|U^j|$ replaced by $\|U^j\|_{L_2(\Omega)}$. So an application of Theorem 3.2* yields the desired assertion $\|U^j\|_{L_2(\Omega)} \lesssim \mathcal{U}^j(\tau_1;\gamma)$ $\forall j \ge 1$.

(iii) It remains to consider the case $K > 1$, which will be reduced to the case $K = 1$ by imitating part (ii) in the proof of Theorem 3.2. In particular, for $m \le K$ we now get $\|U^m\|_{L_2(\Omega)} + \tau_1^\alpha\langle \mathcal{L}U^m, U^m\rangle \lesssim \sum_{j=0}^{m-1}\|U^j\|_{L_2(\Omega)} + \tau_1^\alpha\|f^m\|_{L_2(\Omega)}$. Here $\|f^m\|_{L_2(\Omega)} \lesssim 1$, so $\|U^m\|_{L_2(\Omega)} \lesssim \tau_1^\alpha \simeq \mathcal{U}^m$ $\forall m \le K$. For $m > K$, we proceed exactly as in part (ii) in the proof of Theorem 3.2 and employ $\{\mathring{U}^j\}$ and $\mathring{\delta}_t^\alpha$. □

THEOREM 5.2. *Let the temporal mesh satisfy* (3.1) *for some* $r \ge 1$, *and also* $\sigma_j \ge \sigma_{j+1} \ge 0$ $\forall j \ge 2$ *and* $\sigma_j \in [0, \bar{\sigma}^*]$ $\forall j \ge K+1$, *where* $\bar{\sigma}^* \in (0,1)$ *is from Lemma 5.1, and* $1 \le K \lesssim 1$. *Suppose that* u *from* (1.2) *satisfies* $\|\partial_t^l u(\cdot,t)\|_{L_2(\Omega)} \lesssim 1 + t^{\alpha-l}$ *for* $l = 1,3$ *and* $t \in (0,T]$. *Then for* $\{U^m\}$ *from* (5.1), *one has*

$$\|u(\cdot,t_m) - U^m\|_{L_2(\Omega)} \lesssim \mathcal{E}^m \qquad \forall m = 1,\dots,M,$$

where $\mathcal{E}^m$ *is from* (4.2).

Remark 5.3. Theorem 5.2 applies to the standard graded mesh $\{t_j = T(j/M)^r\}_{j=0}^M$ *for any* $r \ge 1$ *(in view of Corollary 3.3), as well as to the modified graded mesh* (3.4).

Furthermore, the proof of this theorem can be easily extended to the case of the modified discrete fractional-derivative operator described in Remark 3.5.

Proof. Consider the error $e^m := u(\cdot,t_m) - U^m$, for which (1.2) and (5.1) imply $\delta_t^\alpha e^m + \mathcal{L}e^m = r^m \;\forall\, m \ge 1$ and $e^0 = 0$, where the truncation error $r^m := \delta_t^\alpha u(\cdot,t_m) - D_t^\alpha u(\cdot,t_m)$ is estimated in Lemma 4.5 and hence satisfies (4.4). In the latter $\psi^j = \psi^j(x)$ is defined by (4.3), in which $u(\cdot)$ is understood as $u(x,\cdot)$ when evaluating $\partial_s u$, $\partial_s^2 u$, etc. Furthermore, combining (4.3) with (3.1) yields $\|\psi^1\|_{L_2(\Omega)} \lesssim 1$ (in view of $\|\mathrm{osc}(u(\cdot,t),[0,t_2])\|_{L_2(\Omega)} \le \int_0^{t_2} \|\partial_s u\|_{L_2(\Omega)} ds \lesssim t_2^\alpha$) and $\|\psi^j\|_{L_2(\Omega)} \lesssim 1$ for $j \ge 2$ (in view of $s \simeq t_j$ for $s \in (t_{j-1}, t_{j+1})$ for this case). Consequently, we get a version of (4.6): $\|r^m\|_{L_2(\Omega)} \lesssim (\tau/t_m)^{\gamma+1} \;\forall\, m \ge 1$, where $\gamma + 1 := \min\{\alpha+1, (3-\alpha)/r\}$. It remains to apply the estimate of type (5.2) from Lemma 5.1 to $\{e^j\}$ considering the three cases for r as in the proof of Theorem 4.1. ∎

5.2. Error analysis for full discretizations. In this section, we discretize (1.2)–(1.3), posed in a general bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, by applying a standard finite element spatial approximation to the temporal semidiscretization (5.1). Let $S_h \subset H_0^1(\Omega) \cap C(\bar\Omega)$ be a Lagrange finite element space of fixed degree $\ell \ge 1$ relative to a quasiuniform simplicial triangulation $\mathcal{T}$ of Ω. (To simplify the presentation, it will be assumed that the triangulation covers Ω exactly.) Now, $\forall\, m = 1,\ldots,M$, let $u_h^m \in S_h$ satisfy

$$\langle \delta_t^\alpha u_h^m, v_h\rangle + A(u_h^m, v_h) = \langle f(\cdot,t_m), v_h\rangle \qquad \forall v_h \in S_h \tag{5.4}$$

with some $u_h^0 \approx u_0$. Here $\langle\cdot,\cdot\rangle$ is the $L_2(\Omega)$ inner product, while A is the standard symmetric bilinear form associated with the elliptic operator $\mathcal{L}$ (i.e. $A(v,w) = \langle \mathcal{L}v, w\rangle$ for smooth v and w in $H_0^1(\Omega)$).

LEMMA 5.4 (Stability for full discretizations). *Under the conditions of Lemma 5.1 on the temporal mesh, for* $\{u_h^j\}_{j=0}^M$ *from* (5.4) *one has*

$$u_h^0 = 0 \;\text{ in } \bar\Omega, \qquad \|f(\cdot,t_j)\|_{L_2(\Omega)} \lesssim (\tau_1/t_j)^{\gamma+1} \;\;\forall j \ge 1 \quad \Rightarrow \quad \|u_h^j\|_{L_2(\Omega)} \lesssim \mathcal{U}^j(\tau_1;\gamma), \tag{5.5}$$

where $\mathcal{U}^j$ *is defined in* (3.2).

Proof. We closely imitate the proof of Lemma 5.1 replacing $\{U^j\}$ everywhere by $\{u_h^j\}$, and also employing (5.4) with $v_h := V^m = \frac{1}{1-\beta_m} u_h^m - \frac{\beta_m}{1-\beta_m} u_h^{m-1}$ instead of (5.1). ∎

Our error analysis will invoke the Ritz projection $\mathcal{R}_h u(t) \in S_h$ of $u(\cdot,t)$ associated with our discretization of the operator $\mathcal{L}$ and defined by $A(\mathcal{R}_h u, v_h) = \langle \mathcal{L}u, v_h\rangle \;\forall v_h \in S_h$ and $t \in [0,T]$. Assuming that the domain is such that $\|v\|_{W_2^2(\Omega)} \lesssim \|\mathcal{L}v\|_{L_2(\Omega)}$ whenever $\mathcal{L}v \in L_2(\Omega)$, for the error of the Ritz projection $\rho(\cdot,t) = \mathcal{R}_h u(t) - u(\cdot,t)$ one has

$$\|\partial_t^l \rho(\cdot,t)\|_{L_2(\Omega)} \lesssim h \inf_{v_h \in S_h} \|\partial_t^l u(\cdot,t) - v_h\|_{W_2^1(\Omega)} \qquad \text{for } \; l = 0,1, \;\; t \in (0,T]. \tag{5.6}$$

For $l = 0$, see, e.g., [1, Theorem 5.7.6]. A similar result for $l = 1$ follows as $\partial_t \rho(\cdot,t) = \mathcal{R}_h \dot u(t) - \dot u(\cdot,t)$, where $\dot u := \partial_t u$.

THEOREM 5.5. *Let the temporal mesh satisfy* (3.1) *for some* $r \ge 1$, *and also* $\sigma_j \ge \sigma_{j+1} \ge 0 \;\forall\, j \ge 2$ *and* $\sigma_j \in [0, \bar\sigma^*] \;\forall j \ge K+1$, *where* $\bar\sigma^* \in (0,1)$ *is from Lemma 5.1, and* $1 \le K \lesssim 1$. *Suppose that* u *from* (1.2) *satisfies* $\|\partial_t^l u(\cdot,t)\|_{L_2(\Omega)} \lesssim 1 + t^{\alpha-l}$ *for* $l = 1,3$ *and* $t \in (0,T]$. *Then for* $\{u_h^m\}$ *from* (5.4), *subject to* $u_h^0 = \mathcal{R}_h u_0$, *one has*

$$\|u(\cdot,t_m) - u_h^m\|_{L_2(\Omega)} \lesssim \mathcal{E}^m + \|\rho(\cdot,t_m)\|_{L_2(\Omega)} + t_m^\alpha \sup_{t\in(0,t_m)} \bigl\{ t^{1-\alpha} \|\partial_t \rho(\cdot,t)\|_{L_2(\Omega)} \bigr\} \tag{5.7}$$

$\forall\, m \ge 1$, *where* $\rho(\cdot,t) := \mathcal{R}_h u(t) - u(\cdot,t)$, *and* $\mathcal{E}^m$ *is from* (4.2).

Proof. Let $e_h^m := \mathcal{R}_h u(t_m) - u_h^m \in S_h$. Then $u(\cdot,t_m) - u_h^m = e_h^m - \rho(\cdot,t_m)$, so it suffices to prove the desired bounds for e_h^m. Note that $e_h^0 = 0$, while a standard calculation using (5.4) and (1.2) yields

$$\langle \delta_t^\alpha e_h^m, v_h\rangle + A(e_h^m, v_h) = \langle \delta_t^\alpha \underbrace{\mathcal{R}_h u}_{=\rho+u}(t_m), v_h\rangle + \underbrace{A(\mathcal{R}_h u(t_m), v_h)}_{=\langle \mathcal{L}u(\cdot,t_m), v_h\rangle} - \langle f(\cdot,t_m), v_h\rangle$$

$$= \langle \delta_t^\alpha \rho(\cdot,t_m) + r^m, v_h\rangle \qquad \forall v_h \in S_h \;\; \forall\, m \ge 1. \tag{5.8}$$

Here $r^m = \delta_t^\alpha u(\cdot,t_m) - D_t^\alpha u(\cdot,t_m)$ is from the proof of Theorem 5.2, where it was shown that $\|r^m\|_{L_2(\Omega)} \lesssim (\tau/t_m)^{\gamma+1}\ \forall\, m \ge 1$ with $\gamma + 1 := \min\{\alpha+1, (3-\alpha)/r\}$.

Suppose that $\delta_t^\alpha \rho(\cdot,t_m) = 0\ \forall\, m$ in (5.8). Then an application of the estimate of type (5.5) from Lemma 5.4 to $\{e_h^j\}$, with the three cases for r considered separately as in the proof of Theorem 4.1, yields $\|e_h^m\|_{L_2(\Omega)} \lesssim \mathcal{E}^m$.

Next, suppose that $r^m = 0\ \forall\, m$ in (5.8), and $\sup_{t\in(0,T)}\{t^{1-\alpha}\|\partial_t \rho(\cdot,t)\|_{L_2(\Omega)}\} = 1$. Then, by (1.1), $\|D_t^\alpha \rho(\cdot,t_m)\|_{L_2(\Omega)} \lesssim 1$. For $r_\rho^m := \delta_t^\alpha \rho(\cdot,t_m) - D_t^\alpha \rho(\cdot,t_m)$, a version of the truncation error estimation in Lemma 4.5 yields

$$|r_\rho^m| \lesssim (\tau/t_m)^{\min\{\alpha+1,\,(1-\alpha)/r\}} \max_{j=1,\dots,m-1}\{\psi_\rho^j\},$$

where $\{\psi_\rho^j\}$ are defined by versions of (4.3) with u replaced by ρ, and 3 in two places in (4.3b) replaced by 1. So we conclude that $\|r_\rho^m\|_{L_2(\Omega)} \lesssim 1$, and hence $\|\delta_t^\alpha \rho(\cdot,t_m)\|_{L_2(\Omega)} \lesssim 1\ \forall\, m \ge 1$. Now an application of the estimate of type (5.5) from Lemma 5.4 to $\{e_h^j\}$, with $\gamma+1 = 0$, yields $\|e_h^m\|_{L_2(\Omega)} \lesssim \mathcal{U}^m(\tau_1; -1) = t_m^\alpha$, where we also used the definition of $\mathcal{U}^m$ from (3.2).

As (5.8) is a linear problem for $\{e_h^m\}$, combining our findings yields (5.7). □

Recalling the error bounds (5.6) for the the Ritz projection, one immediately gets the following result.

COROLLARY 5.6. *Under the conditions of Theorem 5.2, let* $\|\partial_t^l u(\cdot,t)\|_{W_2^{\ell+1}(\Omega)} \lesssim 1 + t^{\alpha-l}$ *for* $l = 0, 1$ *and* $\|\partial_t^3 u(\cdot,t)\|_{L_2(\Omega)} \lesssim 1 + t^{\alpha-3}$, *where* $t \in (0,T]$. *Then there exists a unique solution* $\{u_h^m\}_{m=1}^M$ *of* (5.4) *subject to* $u_h^0 = \mathcal{R}_h u_0$, *and*

$$\|u(\cdot,t_m) - u_h^m\|_{L_2(\Omega)} \lesssim \mathcal{E}^m + h^{\ell+1} \qquad \forall\, m \ge 1,$$

where $\mathcal{E}^m$ *is from* (4.2).

Remark 5.7. The above Theorem 5.2 and Corollary 5.6 apply to the standard graded mesh $\{t_j = T(j/M)^r\}_{j=0}^M$ *for any* $r \ge 1$ *(in view of Corollary 3.3), as well as to the modified graded mesh* (3.4). *Furthermore, the proofs can be easily extended to the case of the modified discrete fractional-derivative operator described in Remark 3.5.*

Remark 5.8. The assumptions on the derivatives of u made in Corollary 5.6, as well as in Theorems 4.1 and 5.2, are realistic; see examples in [8, §6].

6. Numerical results. Our fractional-order parabolic test problem is (1.2) with $\mathcal{L} = -(\partial_{x_1}^2 + \partial_{x_2}^2)$, posed in the domain $\Omega \times [0,1]$ (see Fig. 6.1, left) with $\partial\Omega$ parameterized by $x_1(l) := \frac{2}{3} R\cos\theta$ and $x_2(l) := R\sin\theta$, where $R(l) := 0.4 + 0.5\cos^2 l$ and $\theta(l) := l + e^{(l-5)/2}\sin(l/2)\sin l$ for $l \in [0, 2\pi]$; see [8, §7]. We choose f, as well as the initial and non-homogeneous boundary conditions, so that the unique exact solution $u = t^\alpha \cos(xy)$. This problem is discretized by (5.4) (with an obvious modification

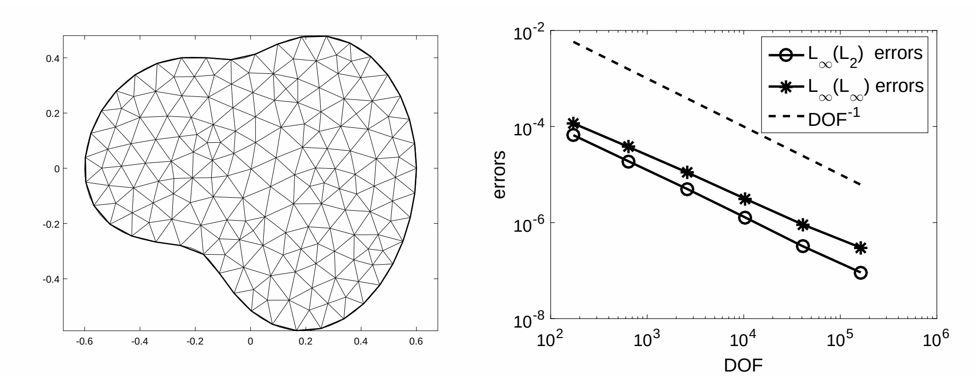

FIG. 6.1. *Fractional-order parabolic test problem: Delaunay triangulation of Ω with DOF=172 (left), maximum $L_2(\Omega)$ and $L_\infty(\Omega)$ errors for $\alpha = 0.5$, $r = (3-\alpha)/\alpha$ and $M = 2048$.*

TABLE 6.1
Fractional-order parabolic test problem: maximum $L_2(\Omega)$ errors (odd rows) and computational rates q in M^{-q} (even rows) for $r = (3-\alpha)/\alpha$ and spatial DOF=255435

	$M=32$	$M=64$	$M=128$	$M=256$	$M=512$	$M=1024$
$\alpha = 0.3$	4.909e-2	8.794e-3	1.427e-3	2.239e-4	3.470e-5	5.353e-6
	2.481	2.624	2.672	2.690	2.697	
$\alpha = 0.5$	2.318e-3	4.807e-4	9.014e-5	1.631e-5	2.911e-6	5.164e-7
	2.270	2.415	2.466	2.486	2.495	
$\alpha = 0.7$	7.752e-4	2.036e-4	4.565e-5	9.774e-6	2.030e-6	4.163e-7
	1.929	2.157	2.224	2.267	2.286	

for the case of non-homogeneous 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 in the maximum $L_2(\Omega)$ norm are shown in Fig. 6.1 (right) and Table 6.1 for, respectively, a large fixed M and DOF. In the latter case, we also give computational rates of convergence. The errors were computed as $\max_{m=1,\dots,M} \|u_h - u^I\|_{L_2(\Omega)}$, where $u^I \in S_h$ is the piecewise-linear interpolant in Ω. (Fig. 6.1 (right) also shows the errors in the maximum $L_\infty(\Omega)$ norm.) The graded temporal mesh $\{t_j = T(j/M)^r\}_{j=0}^M$ was used with the optimal $r = (3-\alpha)/\alpha$; see Remark 4.3. In view of the latter remark, by Corollary 5.6, the errors are expected to be $\lesssim M^{-(3-\alpha)} + h^2$, where $h^2 \simeq \mathrm{DOF}^{-1}$. Our numerical results clearly confirm the sharpness of this corollary for the considered case.

6.1. Pointwise sharpness of error estimate for the initial-value problem. Here, to demonstrate the sharpness of the error estimate (4.2) given by Theorem 4.1, we consider the simplest initial-value fractional-derivative test problem (4.1) with the simplest typical exact solution $u(t) := t^\alpha$. Table 6.2 shows the errors and the corresponding convergence rates at $t = 1$, which agree with (4.2), in view of Remark 4.2. 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 6.3 clearly confirm the conclusions of Remark 4.3, which predicts from the pointwise bound (4.2) that the global errors are $\lesssim M^{-\min\{\alpha r,3-\alpha\}}$. Furthermore, in Fig. 6.2, the pointwise errors for various r are compared with the pointwise theoretical error bound (4.2), and again, with the exception of a few initial mesh nodes, we observe remarkably good agreement. Note that Fig. 6.2 only addresses the case $\alpha = 0.5$, but for other values of α we observed similar consistency of (4.2) with the actual pointwise errors.

Table 6.2
Initial-value test problem: errors at $t = 1$ (odd rows) and computational rates q in M^{-q} (even rows) for $r = 1$, $r = (3-\alpha)/.95$ and $r = (3-\alpha)/\alpha$

		$M = 2^5$	$M = 2^7$	$M = 2^9$	$M = 2^{11}$	$M = 2^{13}$	$M = 2^{15}$
$r = 1$	$\alpha = 0.3$	3.324e-3	8.297e-4	2.073e-4	5.182e-5	1.296e-5	3.239e-6
		1.001	1.000	1.000	1.000	1.000	
	$\alpha = 0.5$	4.557e-3	1.141e-3	2.852e-4	7.132e-5	1.783e-5	4.457e-6
		0.999	1.000	1.000	1.000	1.000	
	$\alpha = 0.7$	4.501e-3	1.127e-3	2.818e-4	7.047e-5	1.762e-5	4.405e-6
		0.999	1.000	1.000	1.000	1.000	
$r = \frac{3-\alpha}{.95}$	$\alpha = 0.3$	1.570e-4	3.435e-6	7.601e-8	1.701e-9	3.843e-11	8.771e-13
		2.757	2.749	2.741	2.734	2.727	
	$\alpha = 0.5$	5.440e-4	1.828e-5	6.038e-7	1.972e-8	6.384e-10	2.053e-11
		2.447	2.460	2.468	2.474	2.480	
	$\alpha = 0.7$	9.278e-4	4.524e-5	2.101e-6	9.477e-8	4.191e-9	1.827e-10
		2.179	2.214	2.235	2.249	2.260	
$r = \frac{3-\alpha}{\alpha}$	$\alpha = 0.3$	8.360e-4	1.481e-5	2.950e-7	6.248e-9	1.373e-10	3.088e-12
		2.910	2.825	2.781	2.754	2.737	
	$\alpha = 0.5$	7.448e-4	1.973e-5	5.839e-7	1.788e-8	5.541e-10	1.726e-11
		2.619	2.539	2.515	2.506	2.503	
	$\alpha = 0.7$	9.391e-4	3.381e-5	1.320e-6	5.339e-8	2.188e-9	9.009e-11
		2.398	2.340	2.314	2.304	2.301	

Table 6.3
Initial-value test problem: maximum nodal errors (odd rows) and computational rates q in M^{-q} (even rows) for $r = 1$, $r = 3-\alpha$ and $r = (3-\alpha)/\alpha$

		$M = 2^5$	$M = 2^7$	$M = 2^9$	$M = 2^{11}$	$M = 2^{13}$	$M = 2^{15}$
$r = 1$	$\alpha = 0.3$	6.524e-2	4.304e-2	2.840e-2	1.873e-2	1.236e-2	8.155e-3
		0.300	0.300	0.300	0.300	0.300	
	$\alpha = 0.5$	3.794e-2	1.897e-2	9.484e-3	4.742e-3	2.371e-3	1.186e-3
		0.500	0.500	0.500	0.500	0.500	
	$\alpha = 0.7$	1.631e-2	6.180e-3	2.342e-3	8.874e-4	3.363e-4	1.274e-4
		0.700	0.700	0.700	0.700	0.700	
$r = 3-\alpha$	$\alpha = 0.3$	2.131e-2	6.934e-3	2.256e-3	7.339e-4	2.388e-4	7.768e-5
		0.810	0.810	0.810	0.810	0.810	
	$\alpha = 0.5$	6.185e-3	1.093e-3	1.933e-4	3.417e-5	6.040e-6	1.068e-6
		1.250	1.250	1.250	1.250	1.250	
	$\alpha = 0.7$	1.867e-3	2.004e-4	2.151e-5	2.308e-6	2.477e-7	2.659e-8
		1.610	1.610	1.610	1.610	1.610	
$r = \frac{3-\alpha}{\alpha}$	$\alpha = 0.3$	6.510e-2	1.542e-3	3.652e-5	8.648e-7	2.048e-8	4.851e-10
		2.700	2.700	2.700	2.700	2.700	
	$\alpha = 0.5$	3.142e-3	9.820e-5	3.069e-6	9.590e-8	2.997e-9	9.365e-11
		2.500	2.500	2.500	2.500	2.500	
	$\alpha = 0.7$	1.273e-3	5.247e-5	2.164e-6	8.922e-8	3.679e-9	1.517e-10
		2.300	2.300	2.300	2.300	2.300	

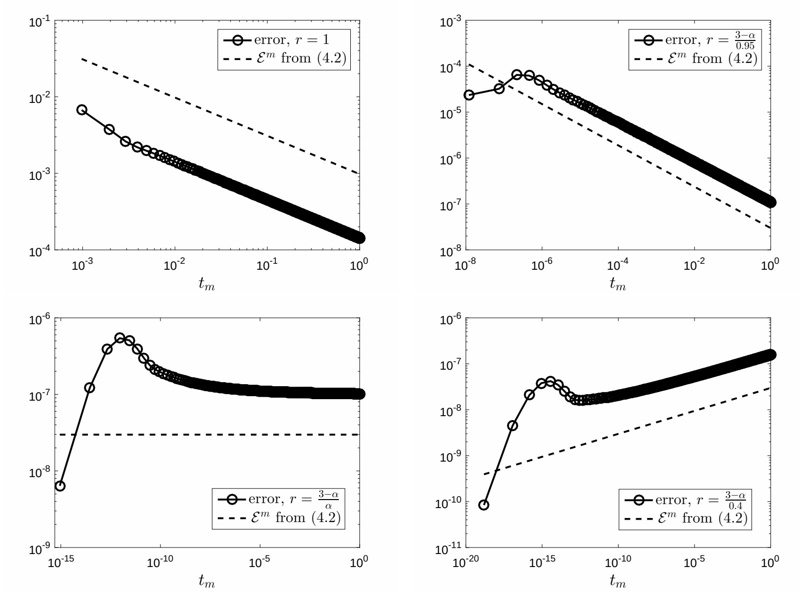

FIG. 6.2. *Initial-value test problem: pointwise errors for $\alpha = 0.5$ and $M = 1024$, cases $r = 1$, $r = (3-\alpha)/0.95$, $r = (3-\alpha)/\alpha$ and $r = (3-\alpha)/0.4$.*

REFERENCES

[1] S. C. BRENNER AND L. R. SCOTT, *The mathematical theory of finite element methods*, Springer-Verlag, New York, third ed., 2008.

[2] 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

[3] K. DIETHELM, *The analysis of fractional differential equations*, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2010.

[4] G. GAO, Z. SUN, AND H. ZHANG, *A new fractional numerical differentiation formula to approximate the Caputo fractional derivative and its applications*, J. Comput. Phys., 259 (2014), pp. 33–50.

[5] 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. 3342

[6] 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.

[7] 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.

[8] 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., 2019, published electronically 23-Jan-2019; doi: 10.1090/mcom/3410.

[9] N. KOPTEVA AND X. MENG, *Error analysis for a fractional-derivative parabolic problem on quasi-graded meshes using barrier functions*, arXiv:1905.07426, (2019).

[10] 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.

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

[12] 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.

[13] Y. XING AND Y. YAN, *A higher order numerical method for time fractional partial differential equations with nonsmooth data*, J. Comput. Phys., 357 (2018), pp. 305–323.