SIAM J. NUMER. ANAL.
Vol. 58, No. 4, pp. 2212–2234

ERROR ANALYSIS FOR TIME-FRACTIONAL SEMILINEAR PARABOLIC EQUATIONS USING UPPER AND LOWER SOLUTIONS*

NATALIA KOPTEVA†

Abstract. A semilinear 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 L1-type discretizations of this problem, we employ the method of upper and lower solutions to obtain 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 the optimal convergence rate $2-\alpha$ in positive time, while quasi-uniform temporal meshes yield first-order convergence in positive time. Furthermore, under appropriate conditions on the nonlinearity, the method of upper and lower solutions immediately implies that, similarly to the exact solutions, the computed solutions lie within a certain range. Semidiscretizations in time and full discretizations using finite differences and finite elements in space are addressed. The theoretical findings are illustrated by numerical experiments.

Key words. fractional-order parabolic equation, semilinear, L1 scheme, graded mesh

AMS subject classifications. 65M06, 65M15, 65M60

DOI. 10.1137/20M1313015

1. Introduction. The method of upper and lower solutions is a very elegant technique frequently used in the analysis of semilinear parabolic and elliptic equations [1, 10, 28], as well as their discretizations [29, 30, 16, 20]. In this paper we shall generalize this approach to discretizations of semilinear fractional-parabolic equations. This, essentially, will enable us to seamlessly extend the error analysis of the recent paper [19] to the challenging semilinar case and thus obtain sharp pointwise-in-time error bounds for quasi-graded temporal meshes with arbitrary degree of grading. There are a few papers on the numerical analysis of similar nonlinear time-fractional equations [8, 15, 13, 23], but we are not aware of any such general results in the literature.

The following fractional-in-time semilinear parabolic problem is considered:

$$\begin{aligned} &D_t^\alpha u + \mathcal{L}u + f(x,t,u) = 0 \quad \text{for } (x,t) \in \Omega \times (0,T], \\ &u(x,t) = 0 \quad \text{for } (x,t) \in \partial\Omega \times (0,T], \qquad u(x,0) = u_0(x) \quad \text{for } x \in \Omega. \end{aligned} \tag{1.1}$$

This problem is posed in a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ (where $d \in \{1,2,3\}$). The operator D_t^α, for some $\alpha \in (0,1)$, is the Caputo fractional derivative in time defined [7] by

$$D_t^\alpha u(\cdot,t) := \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha}\, \partial_s u(\cdot,s)\, ds \qquad \text{for } 0 < t \le T, \tag{1.2}$$

where $\Gamma(\cdot)$ is the Gamma function, and ∂_s denotes the partial derivative in s. The spatial operator $\mathcal{L}$ here is a linear second-order elliptic operator:

*Received by the editors January 15, 2020; accepted for publication (in revised form) June 12, 2020; published electronically July 30, 2020.
https://doi.org/10.1137/20M1313015
Funding: The work of the 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).

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

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

This problem will be considered under the following assumptions on f.

A1. Let f be continuous in s and satisfy $f(\cdot,t,s) \in L_\infty(\Omega)$ $\forall t > 0$ and $s \in \mathbb{R}$, and the one-sided Lipschitz condition

$$f(x,t,s_1) - f(x,t,s_2) \ge -\lambda[s_1 - s_2] \qquad \forall s_1 \ge s_2, \quad x \in \Omega, \quad t > 0,$$

with some constant $\lambda \ge 0$.

A2. There exist constants $\sigma_1 \le 0 \le \sigma_2$ such that $f(\cdot,\cdot,\sigma_1) \le 0$ and $f(\cdot,\cdot,\sigma_2) \ge 0$, while $c = 0$ in (1.3).

Example 1 (negative diffusion coefficient). The linear $f = c^*(x,t)\,u + F(x,t)$, with a possibly negative diffusion coefficient $c^* \ge -\lambda$, clearly satisfies A1.

Example 2 (Allen–Cahn equation). The cubic $f = u^3 - u$ satisfies both A1 and A2 with, e.g., $-\sigma_1 = \sigma_2 = 1$. In particular, the recent papers [8, 13, 23] are devoted to this equation. Note that if $|u_0| \le 1$, then $|u| \le 1$ $\forall t$ [8, Theorem 2.4], while our results below imply a similar property for the computed solutions.

Example 3 (Fisher equation). The quadratic $f = u^2 - u$ satisfies A2 with, e.g., $\sigma_1 = 0$ and $\sigma_2 = 1$, but not A1. (To be more precise, A2 is satisfied for $s \ge -C$, where $C \ge 0$ is a fixed positive constant.) Such equations are addressed in section 8.1.

In this paper, we shall focus on popular L1-type schemes for problem (1.1). Thus, consider the discetization of $D_t^\alpha u$ defined, for $m = 1, \dots, M$, by

$$\delta_t^\alpha U^m := \frac{1}{\Gamma(1-\alpha)} \sum_{j=1}^{m} \delta_t U^j \int_{t_{j-1}}^{t_j} (t_m - s)^{-\alpha}\, ds, \qquad \delta_t U^j := \frac{U^j - U^{j-1}}{t_j - t_{j-1}}, \tag{1.4}$$

when associated with the temporal mesh $0 = t_0 < t_1 < \cdots < t_M = T$ on $[0,T]$. (Note that, similarly to [18, 19], the approach of the present paper may be extended to other discretizations that are monotone in time.)

An essential building block in our analysis is the following stability result. Given $\lambda \ge 0$ and $\gamma \in \mathbb{R}$ (where $\gamma \ne 0$ if $\lambda > 0$), as well as a temporal mesh $\{t_j\}_{j=0}^M$ on $[0,T]$ with $\tau := t_1$, under certain conditions on the mesh, the following is true for $\{V^j\}_{j=0}^M$:

(1.5)

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

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. Note that (1.5) is sharp in the sense that it is consistent with the analogous property for the continuous operator $D_t^\alpha - \lambda$ (similarly to [19, Remark 1.1]). It is worth mentioning that for $\lambda = 0$ it is obtained in [19] using barrier functions, while here we extend (1.5) to $\lambda > 0$ simply as a clever corollary of this property for $\lambda = 0$ (by constructing an appropriate upper solution for the operator $\delta_t^\alpha - \lambda$).

It should be noted that while the explicit inverse of $D_t^\alpha - \lambda$ is easily available, the proof of (1.5) 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 [21, 22, 13, 23]. However, the latter approach involves intricate evaluations and, furthermore, yields less sharp error bounds (see Remark 4.3 for a more detailed discussion). Our approach in [19] and here is entirely different and is substantially more concise as we obtain (1.5) essentially using clever barrier functions for δ_t^α, while the numerical results in [19] and section 9 indicate that our error bounds are sharp in the pointwise-in-time sense.

Similarly to [5, 6, 13, 23, 17, 19, 21, 25, 26, 34], 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 (1.1). At the same time, as a particular case, our results immediately apply to uniform temporal meshes.

A number of outstanding theoretical gaps in the error analysis for semilinear fractional-parabolic equations will be addressed.

- Under very general conditions A1 and A2, whenever the exact solution lies within a certain range (e.g., $[\sigma_1, \sigma_2]$, or it is positive), the method of discrete upper and lower solutions will easily yield a similar property for the computed solutions. Similar results have been obtained only for the Allen–Cahn equation using the specific form of f; see [8, Theorem 3.3], [13, Theorem 2.2], [23, Theorem 3.1].
- Combining the theory of upper and lower solutions with the subtle and sharp stability property (1.5) yields sharp pointwise-in-time error bounds for quasi-graded temporal meshes with arbitrary degree of grading. We are not aware of any such general results in the literature.
- A straightforward particular case of our error bounds is that the (quasi-)uniform temporal mesh yields the first-order convergence in positive time $t \gtrsim 1$ (see Remark 4.3). This is consistent with the error bounds in [12, 14, 17, 19] obtained for the linear case but appears to be a new result for the semilinear equations.
- Another particular case of our error bounds indicates that the optimal convergence rates of order $2-\alpha$ in positive time $t \gtrsim 1$ are attained using much milder (compared to the optimal) grading with $r > 2 - \alpha$ (see Remark 4.3). This is consistent with [19] but has not been proved before for the semilinear case.
- Note also that when the optimal grading parameter $r = (2-\alpha)/\alpha$ is used, as particular cases, we recover the optimal global convergence rate of order $2-\alpha$ (similarly to [13, Theorem 3.1]), while in the case of quasi-uniform temporal meshes we recover the global convergence rate of order α (similarly to [8, Theorem 4.2] and [15, Theorem 4.4]); see Remark 4.4.

Strictly speaking, Remarks 4.3 and 4.4, to which we have referred above, apply to the L1 discretizations of the initial-value problem of type (1.1). At the same time, the discussion there focuses on the term $\mathcal{E}^m$, which also appears in the error estimates for semidiscretizations of the initial-boundary-value problem (1.1), and its full discretizations using finite differences and finite elements (see Theorems 5.1, 6.1, 7.1, 7.4, and 6.1^*).

To be more precise with regard to the earlier literature, [8, 13, 23] are devoted to the Allen–Cahn equation, while [15] addresses a more general semilinear equation with a Lipschitz-continuous $f = f(u)$ (which is more restrictive compared to A1). In [8], the error is estimated globally in time in the $L_2(\Omega)$ norm for Grünwald–Letnikov-type semidiscretizations on uniform temporal meshes. In [15], similar error bounds are given for the L1 and the backward Euler convolution quadrature discretizations in

time combined with linear finite elements in space. In [13, 23], the error is estimated in the $L_\infty(\Omega)$ norm for, respectively, the L1 and Alikhanov schemes in time combined with standard finite differences in space for the case of periodic boundary conditions.

Throughout the paper, it is assumed that there exists a unique solution of (1.1) such that at least $u(\cdot,t) \in H_0^1(\Omega)\ \forall t > 0$, and $\|\partial_t^l u(\cdot,t)\|_{L_p(\Omega)} \lesssim 1 + t^{\alpha-l}$ for $l \le 2$ with $p \in \{2,\infty\}$. The latter is a realistic assumption (e.g., proved in [15, (3.1)] for $l = 1$, $p = 2$), in contrast to stronger assumptions of type $\|\partial_t^l u(\cdot,t)\|_{L_p(\Omega)} \lesssim 1$ frequently made in the literature. Indeed, [33, Theorem 2.1] clearly shows that this stronger assumption is too restrictive even in the linear case. When full discretizations are considered in sections 6–7, additional assumptions are required on $\|\partial_{x_k}^l u(\cdot,t)\|_{L_\infty(\Omega)}$ for $l = 3, 4$ in Theorem 6.1 and on $\|\partial_t^l u(\cdot,t)\|_{W_p^2(\Omega)}$ for $l = 0, 1$, $p \in \{2,\infty\}$ in Remarks 7.3 and 7.5. For bounds of this type in the linear case, see [24, (1.6) and (1.7)], [31], [34, section 2], [17, section 6]. For some existence, uniqueness, and regularity results for the semilinear case, we also refer the reader to [8, Theorem 2.3] and [15, Theorem 3.1].

Outline. We start by describing discrete upper and lower solutions and their properties in section 2; all these results are valid for arbitrarily large T (see Remark 2.6). Next, section 3 is devoted to the proof of the stability result (1.5) (a version of which for arbitrarily large T is discussed in section 3.3). This result is then employed to obtain pointwise-in-time error bounds for L1-type discretizations of the initial-value problem of type (1.1) in section 4, semidiscretizations of the initial-boundary-value problem (1.1) in section 5, and its full discretizations using finite differences in section 6 and finite elements in section 7 (where the consideration is restricted to $\mathcal{L} = -\triangle$). Generalizations of the above results, such as the treatment of other types of boundary conditions, are discussed in section 8. Finally, our theoretical findings are illustrated by numerical experiments in section 9.

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, f, and α, 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 $H^1(\Omega) = W_2^1(\Omega)$ and $H_0^1(\Omega)$ is the space of functions in $H^1(\Omega)$ vanishing on $\partial\Omega$.

2. Discrete upper and lower solutions. In this section we shall consider definitions and certain properties of discrete upper and lower solutions in the context of the semidiscretization of our original problem. Extensions for the operator $\delta_t^\alpha - \lambda$ without spatial derivatives and certain full discretizations will be given in sections 2.1–2.2.

Consider the semidiscretization of our problem (1.1) in time:

(2.1)
$$\delta_t^\alpha U^m + \mathcal{L}U^m + f(\cdot,t_m,U^m) = 0 \ \text{ in } \Omega, \quad U^m = 0 \ \text{ on } \partial\Omega \quad \forall\, m = 1,\dots,M; \quad U^0 = u_0.$$

DEFINITION. *The discrete function* $\{\bar U^j\}_{j=0}^M$ *is called an* upper solution *of problem* (2.1) *if it satisfies (possibly in a weak sense* [10, section 9.3]*) the following conditions:*

(2.2)
$$\delta_t^\alpha \bar U^m + \mathcal{L}\bar U^m + f(\cdot,t_m,\bar U^m) \ge 0 \ \text{ in } \Omega, \quad \bar U^m \ge 0 \ \text{ on } \partial\Omega \quad \forall\, m \ge 1; \quad \bar U^0 \ge u_0.$$

The discrete function $\{\underline U^j\}_{j=0}^M$ *is called a* lower solution *of problem* (2.1) *if it satisfies the reversed inequalities in* (2.2).

LEMMA 2.1. *Suppose that* f *satisfies* A1*, and* $\lambda\tau_j^\alpha \le \{\Gamma(2-\alpha)\}^{-1}\ \forall\, j \ge 1$.

(i) *If* $u_0 \in L_\infty(\Omega)$*, then problem* (2.1) *has a unique solution* $\{U^j\}_{j=0}^M$*, with* $U^j \in H_0^1(\Omega) \cap L_\infty(\Omega)\ \forall j \ge 1$.

(ii) *If* $\{\bar U^j\}_{j=0}^M$*, is an upper solution of problem* (2.1)*,* $\bar U^0 \in L_\infty(\Omega)$*, and* $\bar U^j \in H^1(\Omega)\cap L_\infty(\Omega)$ $\forall j \ge 1$*, then* $U^j \le \bar U^j$ $\forall j \ge 0$.

(iii) *If* $\{\underline U^j\}_{j=0}^M$ *is a lower solution of problem* (2.1)*,* $\underline U^0 \in L_\infty(\Omega)$*, and* $\underline U^j \in H^1(\Omega)\cap L_\infty(\Omega)$ $\forall j \ge 1$*, then* $\underline U^j \le U^j$ $\forall j \ge 0$.

Proof. A straightforward calculation shows that (1.4) can be represented as

(2.3)
$$\delta_t^\alpha U^m = \kappa_{m,m}U^m - \sum_{j=0}^{m-1}\kappa_{m,j}U^j, \quad \text{where} \quad \kappa_{m,m} = \frac{\tau_m^{-\alpha}}{\Gamma(2-\alpha)}, \quad \kappa_{m,j} > 0 \ \ \forall\, m \ge j.$$

(i) The proof is by induction. Assume that there exist desired $\{U^j\}_{j<m}$. Combining (2.1) with (2.3), one concludes that each U^m solves the semilinear elliptic equation

$$\mathcal{N}^m U^m := \mathcal{L}U^m + \big[f(\cdot,t_m,U^m) + \kappa_{m,m}U^m\big] = F^m \qquad \text{in } \Omega, \tag{2.4}$$

where $F^m := \sum_{j=0}^{m-1}\kappa_{m,j}U^j$ is a linear combination of $\{U^j\}_{j<m}$, so $F^m \in L_\infty(\Omega)$. As $\lambda\tau_m^\alpha \le \{\Gamma(2-\alpha)\}^{-1}$ is equivalent to $\lambda \le \kappa_{m,m}$, the part $[f(\cdot,\cdot,U^m) + \kappa_{m,m}U^m]$ in (2.4) is monotone with respect to U^m. Our assumptions on the elliptic operator $\mathcal{L}$ imply that it satisfies the maximum principle, so an application of the argument used in the proof of [9, Lemma 1] with [4, Lemma 16] (where a more general $\mathcal{L}$ is considered) to this elliptic equation, subject to $U^m = 0$ on $\partial\Omega$, yields existence of a unique solution $U^m \in H_0^1(\Omega)\cap L_\infty(\Omega)$. To be more precise, the argument in [9, Lemma 1] relies on $\|v\|_{L_\infty(\Omega)} \lesssim \|\mathcal{L}v\|_{L_\infty(\Omega)}$ $\forall\, v \in H_0^1(\Omega)\cap L_\infty(\Omega)$ (which follows from the maximum principle [11, Theorem 3.7 and section 8.1]) and essentially reduces (2.4) under assumption A1 (and, hence, with a monotone nonlinearity) to the case addressed in [4, Lemma 16]. The latter lemma yields existence of a solution in $H_0^1(\Omega)\cap L_\infty(\Omega)$ for the equation $\mathcal{L}U^m + g(x,U^m) = 0$ with an appropriate $g(x,s)$ uniformly bounded in $\Omega\times\mathbb{R}$, measurable in x and continuous in s.

(ii) The proof is again by induction. Assume that we have established $U^j \le \bar U^j$ for $j < m$. Then for $\bar U^m$ one gets a version of (2.4): $\mathcal{N}^m\bar U^m \ge \bar F^m := \sum_{j=0}^{m-1}\kappa_{m,j}\bar U^j$. Note that $\bar F^m \ge F^m$ (in view of $\kappa_{m,j} > 0$), so $\mathcal{N}^m\bar U^m \ge \mathcal{N}^m U^m$. From this in the domain $\widehat\Omega := \{U^m > \bar U^m\}$, one gets

$$\mathcal{N}^m U^m - \mathcal{N}^m\bar U^m \ge \mathcal{L}[U^m - \bar U^m] + \underbrace{(\kappa_{m,m}-\lambda)}_{\ge 0}\underbrace{[U^m - \bar U^m]}_{>0} \ge \mathcal{L}[U^m - \bar U^m].$$

Hence $\mathcal{L}[U^m - \bar U^m] \le 0$ in $\widehat\Omega$. Finally, with an application of the maximum principle for functions in $H^1(\widehat\Omega)$ [11, section 8.1], one concludes that $\sup_{\widehat\Omega}(U^m - \bar U^m) \le 0$. The desired bound $U^m \le \bar U^m$ in Ω follows.

(iii) Imitate the argument of part (ii). ∎

Corollary 2.2 (bounds for the computed solution). *Under the conditions of Lemma* 2.1*, suppose that* f *also satisfies* A2*, and* $\sigma_1 \le u_0 \le \sigma_2$*. Then for the unique solution of* (2.1) *one has* $\sigma_1 \le U^j \le \sigma_2$ $\forall j \ge 0$.

Proof. A2 implies that σ_1 and σ_2 are, respectively, lower and upper solutions of (2.1). Hence, Lemma 2.1(ii), (iii) yields the desired assertion. ∎

2.1. Extension to the operator $\delta_t^\alpha - \lambda$.

Remark 2.3. The above definitions of upper and lower solutions, as well as a version of Lemma 2.1, but under a stronger assumption $\lambda\tau_j^\alpha < \{\Gamma(2-\alpha)\}^{-1}\ \forall\, j \ge 1$, clearly apply to the simpler operator $\delta_t^\alpha - \lambda$ without spatial derivatives.

COROLLARY 2.4 (comparision principle for $\delta_t^\alpha - \lambda$). *Let the temporal mesh satisfy* $\lambda\tau_j^\alpha < \{\Gamma(2-\alpha)\}^{-1}\ \forall\, j \ge 1$. *Then* $V^0 \le B^0$ *and* $(\delta_t^\alpha - \lambda)V^m \le (\delta_t^\alpha - \lambda)B^m\ \forall\, m \ge 1$ *imply* $V^m \le B^m\ \forall\, m \ge 0$.

Proof. In view of Remark 2.3, the desired conclusion follows from a version of Lemma 2.1(ii) for the operator $\delta_t^\alpha - \lambda$. ∎

2.2. Extension to full discretizations. Let $\bar\Omega_h$ be a finite-dimensional set of points in $\bar\Omega$, comprising the nodes of a certain spatial mesh, and let $\Omega_h := \bar\Omega_h \backslash \partial\Omega$ denote the set of interior mesh nodes. Consider a fully discrete version of (2.1) in the form

$$\begin{aligned} &\delta_t^\alpha U^m(z) + \mathcal{L}_h U^m(z) + f(z, t_m, U^m(z)) = 0 \quad \text{for } z \in \Omega_h,\ \ m = 1, \dots, M, \\ &U^m = 0 \quad \text{in } \bar\Omega_h \cap \partial\Omega,\ \ m = 1, \dots, M, \qquad U^0 = u_0 \quad \text{in } \bar\Omega_h. \end{aligned} \tag{2.5}$$

Generalizing, in an obvious manner, the above definitions of upper and lower solutions to fully discrete problem (2.5), we formulate a version of Lemma 2.1.

LEMMA 2.5. *Suppose the spatial operator* $\mathcal{L}_h$ *in* (2.5) *is associated with an M-matrix,* f *satisfies* A1, *and* $\lambda\tau_j^\alpha \le \{\Gamma(2-\alpha)\}^{-1}\ \forall\, j \ge 1$.

(i) *If* $u_0 \in L_\infty(\Omega)$, *then problem* (2.5) *has a unique solution* $\{U^j\}_{j=0}^M$.

(ii) *If* $\{\bar U^j\}_{j=0}^M$, *is an upper solution of problem* (2.5), *then* $U^j \le \bar U^j\ \forall j \ge 0$.

(iii) *If* $\{\underline U^j\}_{j=0}^M$ *is a lower solution of problem* (2.5), *then* $\underline U^j \le U^j\ \forall j \ge 0$.

Proof. For part (i), we imitate the proof of Lemma 2.1(i). For any $m \ge 1$, the solution U^m of (2.5) satisfies the following version of (2.4): $\mathcal{N}_h^m U^m := \mathcal{L}_h U^m + [f(\cdot, t_m, U^m) + \kappa_{m,m} U^m] = F^m$ in Ω_h. This is a system of $\dim(\Omega_h)$ nonlinear equations, and, in view of condition A1 on f, the part $[f(\cdot,\cdot,U^m) + \kappa_{m,m}U^m]$ is monotone in U^m. Consequently, the mapping $\mathcal{N}_h^m$ satisfies the conditions in [27, section 13.5.6], which yields existence of a unique solution of this equation U^m in Ω_h.

For parts (ii) and (iii), we start by imitating the proof of Lemma 2.1(ii) and, assuming that $U^j \le \bar U^j$ in Ω_h for $j < m$, conclude that $\mathcal{N}_h^m \bar U^m \ge \mathcal{N}_h^m U^m$ in Ω_h. In view of [27, section 13.5.6], the mapping $\mathcal{N}_h^m$ is inverse isotone, which immediately yields $U^m \le \bar U^m$ in Ω_h. ∎

Remark 2.6 ($T \gg 1$). An inspection of the above proofs shows that all results of section 2, as well as Theorems 6.1(ii) and 7.1(ii), are valid for arbitrarily large T. Additionally, a version of the stability property (1.5) for arbitrarily large T will be discussed in section 3.3.

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

3.1. Quasi-graded temporal meshes. Main stability result for $\delta_t^\alpha - \lambda$. Throughout the paper, we shall assume that the temporal mesh is quasi-graded in the sense that, with some $r \ge 1$,

$$\tau := t_1 \simeq M^{-r}, \qquad \tau_j := t_j - t_{j-1} \lesssim \tau^{1/r} t_j^{1-1/r} \qquad \forall\, j = 1, \dots, M. \tag{3.1}$$

Importantly, the results from [19], which we shall employ, apply to this mesh in view of [19, Lemma 2.7].

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 (3.1), in view of $\tau_j \simeq M^{-1}\, t_{j-1}^{1-1/r}$ and $t_j \le 2^r t_{j-1}$ for $j \ge 2$.

The key in our error analysis is the following stability property, which is also the main result of this section.

THEOREM 3.1 (stability of $\delta_t^\alpha - \lambda$). *Let $\lambda\tau_j^\alpha < \{\Gamma(2-\alpha)\}^{-1}\ \forall\, j \ge 1$.*

(i) *Additionally, let the temporal mesh satisfy* (3.1) *with $1 \le r \le (2-\alpha)/\alpha$. Given $\{V^j\}_{j=0}^M$, the stability property* (1.5) *holds true for any fixed $\lambda \ge 0$ and $\gamma \ne 0$.*

(ii) *If $\gamma \le \alpha - 1$, then one has the above result without assuming* (3.1).

(iii) *The above results remain valid if $|(\delta_t^\alpha - \lambda)V^j| \lesssim (\tau/t_j)^{\gamma+1}$ in* (1.5) *is replaced by $(\delta_t^\alpha - \lambda)|V^j| \lesssim (\tau/t_j)^{\gamma+1}$.*

Note that the above result is a generalization of the following particular case, addressed in [19].

THEOREM 3.1* (see [19, Theorem 2.1]). *If $\lambda = 0$, then Theorem* 3.1 *holds true for any fixed $\gamma \in \mathbb{R}$.*

3.2. Proof of Theorem 3.1. To prove Theorem 3.1, we shall employ its particular case, Theorem 3.1* already established in [19], and the following lemma.

LEMMA 3.2. *For any fixed positive constant $c_0 < \frac12\{\lambda\Gamma(2-\alpha)\}^{-1/\alpha}$ such that $\bar\tau := \max\{\tau_j\} \le \frac12 c_0$, and any fixed mesh point $t_m \in \{t_j\}_{j=0}^M$, there exists $\{B^j\}_{j=0}^M$ such that*

$$B^j = 0\ \ \forall\, j \le m, \quad 0 \le B^j \lesssim 1 \quad \text{and} \quad (\delta_t^\alpha - \lambda)\, B^j \gtrsim \begin{cases} 0 & \text{for } t_j < t_m + c_0 \\ 1 & \text{for } t_j \ge t_m + c_0 \end{cases} \quad \forall\, j \ge 1.$$

Next, we proceed to the proof of Theorem 3.1, which will be followed by the proof of Lemma 3.2.

Proof of Theorem 3.1. (i) In view of the comparison principle given by Corollary 2.4, it suffices to show that under the conditions of Theorem 3.1(i)/(ii), there exists a function $\{\mathcal{W}^j\}$ such that

$$\left.\begin{aligned} (\delta_t^\alpha - \lambda)\mathcal{W}^j &\gtrsim (\tau/t_j)^{1+\gamma} \\ \forall j \ge 1,\ \ \mathcal{W}^0 &= 0 \end{aligned}\right\} \text{ and } 0 \le \mathcal{W}^j \lesssim \mathcal{V}_\gamma^j = \tau t_j^{\alpha-1}(\tau/t_j)^{\min\{0,\,\gamma\}}\ \ \forall j \ge 1. \tag{3.2}$$

Note that here the representation of $\mathcal{V}_\gamma^j$, defined in (1.5), relies on $\gamma \ne 0$.

For any $\gamma \ne 0$, Theorem 3.1*(i), the conditions of which are also satisfied, yields

$$\left.\begin{aligned} \delta_t^\alpha \mathcal{B}_\gamma^j &= (\tau/t_j)^{1+\gamma} \\ \forall j \ge 1, \quad \mathcal{B}_\gamma^0 &= 0 \end{aligned}\right\} \ \Rightarrow\ 0 \le \mathcal{B}_\gamma^j \lesssim \mathcal{V}_\gamma^j = \tau^\alpha (\tau/t_j)^{1+\min\{0,\,\gamma\}-\alpha} \quad \forall j \ge 1. \tag{3.3}$$

Here the representation of $\mathcal{V}_\gamma^j$ is different from (but equivalent to) the one in (3.2) and will be more convenient in what follows.

Set

$$\gamma^* := \min\{0,\, \gamma\} - \alpha < 0.$$

Now, (3.3) implies that, for a sufficiently large constant C,

$$0 \le \mathcal{B}_\gamma^j \lesssim \tau^\alpha (\tau/t_j)^{1+\gamma^*} \quad \Rightarrow \quad (\delta_t^\alpha - \lambda)\, \mathcal{B}_\gamma^j \gtrsim (\tau/t_j)^{1+\gamma} - C\tau^\alpha (\tau/t_j)^{1+\gamma^*},$$

$$\begin{aligned} 0 \le \mathcal{B}_{\gamma^*}^j \lesssim \tau^\alpha (\tau/t_j)^{1+\gamma^*-\alpha} \ \Rightarrow\ (\delta_t^\alpha - \lambda)\, \mathcal{B}_{\gamma^*}^j &\gtrsim (\tau/t_j)^{1+\gamma^*} - C\tau^\alpha (\tau/t_j)^{1+\gamma^*-\alpha} \\ &\gtrsim (\tau/t_j)^{1+\gamma^*}\big[1 - Ct_j^\alpha\big], \end{aligned}$$

where we also used $\gamma^{**} := \min\{0,\, \gamma^*\} - \alpha = \gamma^* - \alpha$. Consequently, for a sufficiently large constant $\bar c$ and a sufficiently small constant c_1, one obtains

$$(\delta_t^\alpha - \lambda)\left[\mathcal{B}_\gamma^j + \bar c\,\tau^\alpha\,\mathcal{B}_{\gamma^*}^j\right] \gtrsim (\tau/t_j)^{1+\gamma} - C\begin{cases} 0 & \text{for } t_j < c_1,\\ \tau^{1+\min\{0,\,\gamma\}} & \text{for } t_j \ge c_1.\end{cases}$$

Here, for the case $t_j \ge c_1$, we also employed $\tau^\alpha\,\tau^{1+\gamma^*} = \tau^{1+\min\{0,\,\gamma\}}$. Note also that

$$\mathcal{B}_\gamma^j + \bar c\,\tau^\alpha\,\mathcal{B}_{\gamma^*}^j \lesssim \tau^\alpha (\tau/t_j)^{1+\gamma^*} = \mathcal{V}_\gamma^j\,. \tag{3.4}$$

Finally, let

$$\mathcal{W}^j := \mathcal{B}_\gamma^j + \bar c\,\tau^\alpha\,\mathcal{B}_{\gamma^*}^j + \bar c^2\,\tau^{1+\min\{0,\,\gamma\}}\,B^j,$$

where $\{B^j\}$ is from Lemma 3.2 with $c_0 := \frac{1}{2}c_1$ and $t_m \simeq 1$ such that $t_m + c_0 \le c_1$. Then $(\delta_t^\alpha - \lambda)\mathcal{W}^j \gtrsim (\tau/t_j)^{1+\gamma}$, in agreement with (3.2), while the required upper bound $\mathcal{W}^j \lesssim \mathcal{V}_\gamma^j$ follows from (3.4) combined with $B^j = 0$ for $j \le m$ and $\tau^{1+\min\{0,\,\gamma\}} B_j \lesssim \tau t_j^{\alpha-1}(\tau/t_j)^{\min\{0,\,\gamma\}} = \mathcal{V}_\gamma^j$ for $t_j \ge t_m \gtrsim 1$. Thus, (3.4) is established.

(ii) Note that as now the conditions of Theorem 3.1*(ii) are satisfied, one gets (3.3) only for any $\gamma \le \alpha - 1$. Importantly, if γ satisfies the latter restriction, so does γ^*. Hence, the proof of part (i) applies to this case.

(iii) Let $W^0 = 0$ and $(\delta_t^\alpha - \lambda)W^j \simeq (\tau/t_j)^{\gamma+1} \ge (\delta_t^\alpha - \lambda)|V^j|\ \forall\, j \ge 1$. Then $0 \le |V^j| \le W^j\ \forall\, j \ge 1$ (in view of Corollary 2.4), while the results of parts (i) and (ii) apply to $\{W^j\}$. ∎

It remains to prove the auxiliary Lemma 3.2 (which we used in the above proof).

Proof of Lemma 3.2. First, consider the case $t_m = 0$. Let

$$B(t) := \sum_{k=0}^{K} \bar c^k\, B_k(t), \quad B_k(t) := \max\{0,\, t - q_k\}, \quad q_0 := 0, \quad q_k \in [c_0 k - \bar\tau,\, c_0 k].$$

Here $0 \le K \lesssim 1$ is chosen so that $T \in (q_{K+1}, q_{K+2}]$, i.e., $K + 2 = \lceil T/c_0 \rceil \lesssim 1$ (unless $c_0 \ge T$, in which case $K := 0$).

Applying the continuous operator $D_t^\alpha - \lambda$ to $B_0 = t$, one easily gets

$$(D_t^\alpha - \lambda)\,B_0(t) = t^{1-\alpha}\big(\{\Gamma(2-\alpha)\}^{-1} - \lambda t^\alpha\big) \gtrsim \begin{cases} 0 & \text{for } t \in (0, c_0) \supset (0, q_1),\\ 1 & \text{for } t \in [c_0 - \bar\tau, 2c_0] \supset [q_1, q_2],\\ -1 & \text{for } t \in (q_2, T].\end{cases}$$

In a similar manner, $\forall\, m$ one gets

$$(D_t^\alpha - \lambda)\,B_m(t) \gtrsim \begin{cases} 0 & \text{for } t \in (0, q_{m+1}),\\ 1 & \text{for } t \in [q_{m+1}, q_{m+2}],\\ -1 & \text{for } t \in (q_{m+2}, T].\end{cases}$$

Note that $T \in (q_{K+1}, q_{K+2}]$ implies $(q_{m+2}, T] = \emptyset$ for $m = K$. As $K \lesssim 1$, choosing $\bar c \lesssim 1$ sufficiently large in the definition of $B(t)$, one can obtain

$$(D_t^\alpha - \lambda)\,B(t) \gtrsim \begin{cases} 0 & \text{for } t \in (0, q_1),\\ 1 & \text{for } t \in [q_1, T].\end{cases}$$

It remains to choose $\{q_m\}_{m=0}^K \subset \{t_j\}_{j=0}^M$, e.g., by letting each q_m be the maximal mesh point subject to $q_m \in [c_0 m - \bar\tau, c_0 m]$. Then $(\delta_t^\alpha - \lambda)B(t_j) = (D_t^\alpha - \lambda)B(t_j)$, and the desired result follows for the discrete function $B^j := B(t_j)$.

Finally, consider the case $t_m > 0$. On the submesh $\{t_j\}_{j=m}^M$, construct a discrete function as above. Augmenting this function by zeros on the remaining submesh $\{t_j\}_{j=0}^{m-1}$, one gets the desired $\{B^j\}_{j=0}^M$. ∎

3.3. Long-time stability of $\delta_t^\alpha - \lambda$ for $\gamma + 1 = \alpha$. The stability property (1.5) is established in Theorem 3.1 under the assumption that $T \lesssim 1$, i.e., the terminal time is bounded. At the same time, long-time solutions are frequently of interest for semilinear problems. However, the analysis of long-term stability and convergence is considerably more challenging even for classical parabolic equations. Here we shall extend (1.5) to the case of arbitrarily large T and arbitrary temporal meshes for one particular case of $\gamma + 1 = \alpha$ (which corresponds to the optimal grading parameter $r = (2-\alpha)/\alpha$ in the convergence analysis of section 4).

In the remainder of this section, all constants, including those used in the definition of the notation of type $\lesssim$, will be understood as independent of T.

LEMMA 3.3. *Suppose that $\gamma + 1 = \alpha$, and T is arbitrarily large.*

(i) *If $\lambda = 0$, then (1.5) remains valid on an arbitrary mesh independently of T.*

(ii) *For any $\lambda' > \lambda > 0$, there exist $c_0 = c_0(\lambda')$ and $\bar{C} = \bar{C}(\lambda')$, independent of T and m, such that if $\tau_j \le c_0\ \forall j \le m$, then*

(3.5)
$$V^0 = 0,\ \ |(\delta_t^\alpha - \lambda)V^j| \le (\tau/t_j)^\alpha\ \ \forall\, j \ge 1\ \ \Rightarrow\ \ |V^j| \le \bar{C}\tau^\alpha\, E_\alpha(\lambda' t_j^\alpha)\ \ \forall\, j \le m,$$

where $E_\alpha(s) := \sum_{k=0}^\infty \frac{s^k}{\Gamma(k\alpha+1)}$ is the Mittag–Leffler function of order α.

Proof. (i) For $\lambda = 0$, (1.5) involves $\mathcal{V}_{\alpha-1}^j = \tau^\alpha$ and remains valid on an arbitrary mesh for an arbitrarily large T, as can be shown by an inspection of the proof of [17, Lemma 2.1]. The latter proof may be interpreted as using the barrier $\{\mathbb{1}^j\}_{j\ge0}$ defined by $\mathbb{1}^0 = 0$ and $\mathbb{1}^j = 1$ for $j \ge 1$, which, combined with (1.4), yields $\delta_t^\alpha \mathbb{1}^j \gtrsim t_j^{-\alpha}$ independently of T. The desired bound follows.

(ii) Let $\lambda' > \lambda > 0$ and, to simplify the presentation, first consider a uniform mesh $\{t_j = j\tau\}_{j=0}^\infty$. Clearly, $(\delta_t^\alpha - \lambda)\mathbb{1}^j \ge Ct_j^{-\alpha} - \lambda$. Set $\mathbb{E}(t) := E_\alpha(\lambda' t^\alpha) - 1$, for which [7, Theorem 4.3] yields $(D_t^\alpha - \lambda')\mathbb{E}(t) = \lambda'$. Next, splitting $\sum_{k=1}^\infty$ in the definition of E_α leads to $\mathbb{E} = \hat{\mathbb{E}} + \breve{\mathbb{E}} =: \sum_{1\le k\le 1/\alpha} + \sum_{k>1/\alpha}$, where $\hat{\mathbb{E}}$ is concave and $\breve{\mathbb{E}}$ is convex. For a generic function v, let v^I be its piecewise-linear interpolant, and $v^*(t) := v(t-\tau)$ for $t \ge \tau$ with $v^*(t) := 0$ otherwise.

We shall show below that if τ is sufficiently small, then

$$D_t^\alpha(\hat{\mathbb{E}}^* + \breve{\mathbb{E}})^I \ge D_t^\alpha(\hat{\mathbb{E}}^* + \breve{\mathbb{E}}^*) = D_t^\alpha \mathbb{E}^* = \lambda'(\mathbb{E}^* + 1) \qquad \forall\, t_j,\ j \ge 2, \tag{3.6a}$$

$$\lambda'(\mathbb{E}^* + 1) \ge \lambda\left(\mathbb{E} + \frac{1}{2}\right) \ge \lambda(\hat{\mathbb{E}}^* + \breve{\mathbb{E}}) + \frac{1}{2}\lambda \qquad \forall\, t > 0. \tag{3.6b}$$

Now, for the barrier $B^j := \mathbb{1}^j + 2(\hat{\mathbb{E}}^* + \breve{\mathbb{E}})(t_j)$, the bounds (3.6) imply that $(\delta_t^\alpha - \lambda)B^j \ge \delta_t^\alpha \mathbb{1}^j \gtrsim t_j^{-\alpha}$, while $B^j \le 1 + 2\mathbb{E}(t_j) \le 2E_\alpha(\lambda' t_j^\alpha)$. The desired bound (3.5) follows.

It remains to establish (3.6). For (3.6a), note that $D_t^\alpha((\hat{\mathbb{E}}^*)^I - \hat{\mathbb{E}}^*)(t_j) \ge 0$ follows from $\hat{\mathbb{E}}^* - (\hat{\mathbb{E}}^*)^I \ge 0$ (the latter in view of $\hat{\mathbb{E}}$ being concave) and can be shown by recalling (1.2) and then applying integration by parts (similarly to the truncation error representations used in the proofs of [17, Lemma 2.3] and [19, Lemma 3.4]). We also use $D_t^\alpha(\breve{\mathbb{E}}^I - \breve{\mathbb{E}}^*)(t_j) \ge 0$, which follows from (1.2) combined with the observation

that on each (t_{j-1},t_j) one has $\frac{d}{dt}\breve{\mathbb{E}}^I(t) \ge \frac{d}{dt}\breve{\mathbb{E}}(t_{j-1}) = \frac{d}{dt}\breve{\mathbb{E}}^*(t_j) \ge \frac{d}{dt}\breve{\mathbb{E}}^*(t)$ (as $\breve{\mathbb{E}}$ is convex). The final inequality in (3.6b) follows from $\mathbb{E} = \hat{\mathbb{E}} + \breve{\mathbb{E}}$ combined with $\hat{\mathbb{E}} \ge \hat{\mathbb{E}}^*$. Finally, for the first bound in (3.6b), note that the asymptotic representation [3, (1.8.27)] of $E_\alpha(s)$ implies that, with a sufficiently large constant c_2, for $t-\tau \ge c_2$ one has $|\mathbb{E}(t) + 1 - \alpha^{-1}\exp(\lambda'^{1/\alpha}t)| \le \frac{1}{4}\lambda/\lambda'$. Using the latter to estimate $\lambda'(\mathbb{E}^* + 1) - \lambda(\mathbb{E}+1)$ yields (3.6b) for $t \ge c_2$ provided that $\tau \le \lambda'^{-1/\alpha}\ln(\lambda'/\lambda)$. Additionally, $\mathbb{E}(t) \le \frac{1}{2}$ for $t \le c_1$, with a sufficiently small c_1, again yields (3.6b) for $t \le c_1$. For the remaining $t \in (c_1,c_2)$, using $|\mathbb{E}'| \lesssim 1$ one gets $\mathbb{E}^* - \mathbb{E} \gtrsim -\tau \ge -\frac{1}{2}$, so (3.6b) is proved $\forall\, t > 0$.

If the mesh is nonuniform, a version of the above argument applies with $v^*(t) := v(t-t_l^*)$, where t_l^* is the minimal node in $\{t_j\}$ such that $t_l^* \ge \bar\tau_m := \max_{j\le m}\tau_j$ (hence, $t_l^* \le t_{l-1} + \bar\tau_m \le 2\bar\tau_m$ is sufficiently small). ∎

Remark 3.4. If $\lambda' > \lambda$ in the stability result (3.5) is replaced by λ, then it becomes consistent with the analogous property for the continuous Caputo operator D_t^α. Indeed, for $\mathbb{1}(t) := \min\{t/\tau, 1\}$ and $\bar{\mathbb{E}}(t) := E_\alpha(\lambda t^\alpha) - 1$, a calculation shows that if $t \ge \tau$, then $(D_t^\alpha - \lambda)[\mathbb{1} + \bar{\mathbb{E}}](t) = D_t^\alpha \mathbb{1}(t) \simeq \tau^{-1}[t^{1-\alpha} - (t-\tau)^{1-\alpha}] \simeq t^{-\alpha}$ and $[\mathbb{1} + \bar{\mathbb{E}}](t) = E_\alpha(\lambda t^\alpha)$.

4. 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 semilinear fractional-derivative problem without spatial derivatives together with its discretization:

$$D_t^\alpha u(t) + f(t,u) = 0 \quad \text{for } t \in (0,T], \qquad u(0) = u_0, \tag{4.1a}$$

$$\delta_t^\alpha U^m + f(t_m, U^m) = 0 \quad \text{for } m = 1,\dots,M, \qquad U^0 = u_0. \tag{4.1b}$$

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

The main result here is the following error estimate.

THEOREM 4.1.

(i) *Let the temporal mesh satisfy* (3.1) *with* $r \ge 1$, *and let* $\lambda\tau_j^\alpha < \{\Gamma(2-\alpha)\}^{-1}$ $\forall\, j \ge 1$. *Suppose that* u *is a unique solution of* (4.1a), *in which* f *satisfies a version of* A1, *and* $|\partial_t^l u| \lesssim 1 + t^{\alpha-l}$ *for* $l = 1,2$ *and* $t \in (0,T]$. *Then there exists a unique solution* $\{U^m\}$ *of* (4.1b), *and* $\forall\, m \ge 1$

$$|u(t_m) - U^m| \lesssim \mathcal{E}^m := \begin{cases} M^{-r}\, t_m^{\alpha-1} & \text{if } 1 \le r < 2-\alpha, \\ M^{-r(1-\epsilon)}\, t_m^{\alpha-(1-\epsilon)} & \text{if } r = 2-\alpha, \\ M^{\alpha-2}\, t_m^{\alpha-(2-\alpha)/r} & \text{if } r > 2-\alpha, \end{cases} \tag{4.2}$$

where ϵ *is an arbitrarily small positive constant.*

(ii) *If, additionally,* f *satisfies a version of* A2, *and* $\sigma_1 \le u_0 \le \sigma_2$, *then* $\sigma_1 \le U^m \le \sigma_2$ $\forall\, m \ge 0$.

Remark 4.2 (case $r = 2-\alpha$). Note that for the case $\lambda = 0$ in A1, one can easily get a slightly sharper version of (4.2) with

$$\mathcal{E}^m := M^{\alpha-2}\, t_m^{\alpha-1}[1 + \ln(t_m/t_1)] \qquad \text{if } r = 2-\alpha \ \text{ and } \ \lambda = 0$$

(similarly to the results for the linear case in [19]). In comparison, (4.2) gives a slightly less optimal bound because we have established (1.5) for $\gamma = 0$ only when $\lambda = 0$ (see Theorem 3.1*).

Remark 4.3 (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 $\mathcal{E}^m \simeq M^{(\alpha-2)(1-\epsilon)}$ with an arbitrarily small $\epsilon > 0$.

Note also that for $r = 1$ (i.e., for the quasi-uniform temporal mesh), we have $\mathcal{E}^m \simeq M^{-1}$. This is consistent with the error bounds in [12, 14, 17, 19] obtained for the linear case but appears to be a new result for the semilinear equations.

By contrast, [13, Theorem 3.1] (obtained by means of a discrete Grönwall inequality for the time-fractional Allen–Cahn equation) gives a somewhat similar but considerably less sharp error bound for graded meshes, as (in our notation) it involves the term $O(\tau^\alpha) = O(M^{-\alpha r})$, so it requires (in our notation) $r = (2-\alpha)/\alpha$ to attain the optimal convergence rate in positive time. In fact, for any $r < (2-\alpha)/\alpha$, our error bound is sharper than the pointwise-in-time bound from [13, Theorem 3.1]. (Note also that a similar term $O(\tau^\alpha) = O(M^{-\alpha r})$ appears in the error bound of [23, Theorem 4.1] for the higher-order Alikhanov scheme.)

Remark 4.4 (global convergence). Note that $\max_{m\ge1} \mathcal{E}^m \simeq \mathcal{E}^1 \simeq \tau_1^\alpha \simeq M^{-\alpha r}$ for $\alpha \le (2-\alpha)/r$, while $\max_{m\ge1} \mathcal{E}^m \simeq \mathcal{E}^M \simeq M^{\alpha-2}$ otherwise. Consequently, Theorem 4.1 yields the global error bound $|u(t_m) - U^m| \lesssim M^{-\min\{\alpha r, 2-\alpha\}}$.

This immediately implies that the optimal grading parameter for global accuracy is $r = (2-\alpha)/\alpha$. Note that similar global error bounds were obtained in [21, 17, 34] for the linear case and in [13, Theorem 3.1] for the Allen–Cahn equation.

For $r = 1$, our global error bound becomes $|u(t_m) - U^m| \lesssim M^{-\alpha}$, which is consistent with the bounds of [8, Theorem 4.2] and [15, Theorem 4.4], respectively obtained for Grünwald–Letnikov-type schemes and for L1-type schemes, as well as for the backward Euler convolution quadrature.

Proof of Theorem 4.1. In view of Remark 2.3, the existence of a unique solution $\{U^m\}$ follows from Lemma 2.1(i), while part (ii) follows from Corollary 2.2.

It remains to establish (4.2). Consider the error $e^m := u(t_m) - U^m$ and the truncation error $r^m := \delta_t^\alpha u(t_m) - D_t^\alpha u(t_m)$ $\forall\, m \ge 1$. A standard calculation using (4.1) yields $e^0 = 0$ and

$$\delta_t^\alpha e^m + [f(t_m, U^m + e^m) - f(t_m, U^m)] = r^m \qquad \forall\, m \ge 1. \tag{4.3}$$

Multiply this equation by $\varsigma^m := \operatorname{sign}(e^m)$ and note that $\varsigma^m e^m = |e^m|$ so

$$\varsigma^m(\delta_t^\alpha e^m) \ge \kappa_{m,m}|e^m| - \sum_{j=0}^{m-1} \underbrace{\kappa_{m,j}}_{>0} |e^j| = \delta_t^\alpha |e^m|,$$

$$\varsigma^m[f(t_m, U^m + e^m) - f(t_m, U^m)] \ge -\lambda|e^m|,$$

where we used (2.3) and condition A1 on f. Hence, we arrive at

$$(\delta_t^\alpha - \lambda)\,|e^m| \le |r^m| \qquad \forall\, m \ge 1. \tag{4.4}$$

For the truncation error, recall from [19, Lemma 3.4 and proof of Theorem 3.1] that

$$|r^m| \lesssim (\tau/t_m)^{\gamma+1} \quad \forall\, m \ge 1, \qquad \text{where } \gamma + 1 := \min\{\alpha+1, (2-\alpha)/r\}. \tag{4.5}$$

Hence, we can apply Theorem 3.1 to (4.4) (in particular, note part (iii) of this theorem).

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 3.1(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 $\gamma=0$. As our stability result does not apply to this case, we note that now $|r^m| \lesssim \tau/t_m \lesssim (\tau/t_m)^{1-\epsilon}$ for an arbitrarily small $\epsilon>0$. An application of Theorem 3.1(i) yields $|e^m| \lesssim \tau\, t_m^{\alpha-1}(\tau/t_m)^{-\epsilon} \simeq \tau^{1-\epsilon}\, t_m^{\alpha-(1-\epsilon)}$, where $\tau \simeq M^{-r}$, so $\tau^{1-\epsilon} \simeq M^{-r(1-\epsilon)}$

Case $r > 2-\alpha$. Then $(2-\alpha)/r < 1$, while $\alpha+1>1$, so $\gamma+1 = (2-\alpha)/r < 1$. An application of Theorem 3.1(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 $\tau^{(2-\alpha)/r} \simeq M^{\alpha-2}$. ∎

5. Error analysis for the L1 semidiscretization in time. Recall the semidiscretization of our problem (1.1) in time, given by (2.1).

THEOREM 5.1.

(i) *Let the temporal mesh satisfy* (3.1) *with* $r \ge 1$, *and let* $\lambda\tau_j^\alpha < \{\Gamma(2-\alpha)\}^{-1}$ $\forall\, j \ge 1$. *Suppose that* u *is a unique solution of* (1.1), (1.3) *with the initial condition* $u_0 \in L_\infty(\Omega)$ *and under assumption* A1 *on* f. *Also, given* $p \in \{2,\infty\}$, *suppose that* $u(\cdot,t) \in H_0^1(\Omega)$ *for* $t \in (0,T]$ *and* $\|\partial_t^l u(\cdot,t)\|_{L_p(\Omega)} \lesssim 1 + t^{\alpha-l}$ *for* $l = 0,1,2$ *and* $t \in (0,T]$. *Then there exists a unique solution* $\{U^m\}$ *of* (2.1), *and*

$$\|u(\cdot,t_m) - U^m\|_{L_p(\Omega)} \lesssim \mathcal{E}^m \qquad \forall\, m = 1,\dots,M, \tag{5.1}$$

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

(ii) *If, additionally,* f *satisfies* A2, *and* $\sigma_1 \le u_0 \le \sigma_2$, *then* $\sigma_1 \le U^m \le \sigma_2$ $\forall\, m \ge 0$.

Proof. We imitate the proof of Theorem 4.1.

The existence of a unique solution $U^m \in H_0^1(\Omega) \cap L_\infty(\Omega)$ for $m \ge 1$ follows from Lemma 2.1(i), while part (ii) follows from Corollary 2.2.

It remains to establish (5.1). For the error $e^m := u(\cdot,t_m) - U^m \in H_0^1(\Omega) \cap L_p(\Omega)$, using (1.1) and (2.1), one gets $e^0 = 0$ and

$$\delta_t^\alpha e^m + \mathcal{L}e^m + [f(\cdot,t_m,U^m+e^m) - f(\cdot,t_m,U^m)] = r^m \qquad \forall\, m \ge 1 \tag{5.2}$$

(which is a version of (4.3)). Here $r^m := \delta_t^\alpha u(\cdot,t_m) - D_t^\alpha u(\cdot,t_m)$, and, similarly to (4.5), it satisfies

$$\|r^m\|_{L_p(\Omega)} \lesssim (\tau/t_m)^{\gamma+1} \quad \forall\, m \ge 1, \quad \text{where } \gamma+1 := \min\{\alpha+1, (2-\alpha)/r\}. \tag{5.3}$$

Hence, to get the desired bound (5.1) it suffices to prove

$$(\delta_t^\alpha - \lambda)\|e^m\|_{L_p(\Omega)} \le \|r^m\|_{L_p(\Omega)} \qquad \forall\, m \ge 1, \tag{5.4}$$

which is a version of (4.4), so one then proceeds as in the proof of the error bound (4.2) in Theorem 4.1. The cases $p=2$ and $p=\infty$ of (5.4) will be addressed separately.

For $p=2$, consider the $L_2(\Omega)$ inner product (denoted $\langle\cdot,\cdot\rangle$) of (5.2) with e^m. Clearly, $c - \frac{1}{2}\sum_{k=1}^d \partial_{x_k} b_k \ge 0$ implies $\langle \mathcal{L}e^m, e^m\rangle \ge 0$, and we also have $\langle r^m, e^m\rangle \le \|r^m\|_{L_2(\Omega)}\|e^m\|_{L_2(\Omega)}$. Furthermore, recalling (2.3) and condition A1 on f, one concludes that

(5.5a)
$$\big\langle \delta_t^\alpha e^m,\, e^m\big\rangle = \kappa_{m,m}\|e^m\|^2_{L_2(\Omega)} - \sum_{j=0}^{m-1} \underbrace{\kappa_{m,j}}_{>0}\big\langle e^j,\, e^m\big\rangle \ge \big(\delta_t^\alpha \|e^m\|_{L_2(\Omega)}\big)\|e^m\|_{L_2(\Omega)}\,,$$

(5.5b)
$$\big\langle f(\cdot,t_m,U^m+e^m) - f(\cdot,t_m,U^m),\, e^m\big\rangle \ge -\lambda\|e^m\|^2_{L_2(\Omega)}\,.$$

Combining these findings, one gets (5.4) for $p=2$.

For $p=\infty$, in view of our assumptions, (5.2) implies that $\mathcal{L}e^m \in L_\infty(\Omega)$ (where the bounds of type $f(\cdot,t_m,U^m) \le f(\cdot,t_m,\|U^m\|_{L_\infty(\Omega)}) + 2\lambda\|U^m\|_{L_\infty(\Omega)}$ were used). So $e^m \in H_0^1(\Omega)\cap C(\bar\Omega)$. Now let $\max_{x\in\Omega}|e^m(x)| = |e^m(x^*)|$ for some $x^*\in\Omega$ (where x^* depends on m). Also, let $\varsigma^m := \operatorname{sign}\big(e^m(x^*)\big)$ and note that $\varsigma^m e^m(x^*) = |e^m(x^*)| = \|e^m\|_{L_\infty(\Omega)}$. Now multiply (5.2) at $x=x^*$ by ς^m and note that

(5.6a)
$$\varsigma^m(\delta_t^\alpha e^m)\big|_{x=x^*} \ge \kappa_{m,m}|e^m(x^*)| - \sum_{j=0}^{m-1}\underbrace{\kappa_{m,j}}_{>0}\|e^j\|_{L_\infty(\Omega)} = \delta_t^\alpha\|e^m\|_{L_\infty(\Omega)},$$

(5.6b)
$$\varsigma^m[f(\cdot,t_m,U^m+e^m) - f(\cdot,t_m,U^m)]\big|_{x=x^*} \ge -\lambda|e^m(x^*)| = -\lambda\|e^m\|_{L_\infty(\Omega)},$$

where we used (2.3) and condition A1 on f. Hence, one gets

(5.7)
$$\varsigma^m\mathcal{L}e^m(x^*) + (\delta_t^\alpha - \lambda)\|e^m\|_{L_\infty(\Omega)} \le \|r^m\|_{L_\infty(\Omega)}\,.$$

Furthermore, if e^m is sufficiently smooth and $\mathcal{L}e^m(x^*)$ is defined in the classical sense, then $c\ge 0$ implies $\varsigma^m\mathcal{L}e^m(x^*)\ge 0$, so (5.4) for $p=\infty$ follows. For less smooth e^m, see Remark 5.2. ∎

Remark 5.2 ((5.4) for $p=\infty$). For less smooth $e^m \in H_0^1(\Omega)\cap C(\bar\Omega)$ in the proof of Theorem 5.1, split Ω into disjoint sets Ω^+ and Ω^- such that $\Omega^\pm := \{\pm e^m > 0\}$. Next, instead of (5.7), use similar, but more general relations

$$\pm[\mathcal{L}+\kappa_{m,m}]e(x) \le \sum_{j=0}^{m-1}\kappa_{m,j}\|e^j\|_{L_\infty(\Omega)} + \lambda\|e^m\|_{L_\infty(\Omega)} + \|r^m\|_{L_\infty(\Omega)} \quad \forall\, x\in\Omega^\pm.$$

(The above are obtained using (5.6) with x^* replaced by $x\in\Omega^\pm$ and ς^m by $\pm$.) The desired bound (5.4) for $p=\infty$ follows in view of

(5.8)
$$\kappa_{m,m}\|e^m\|_{L_\infty(\Omega^\pm)} \le \sup_{\Omega^\pm}\big\{\pm[\mathcal{L}+\kappa_{m,m}]e^m\big\}.$$

The latter is obtained using the maximum principle for functions in $H_0^1(\Omega^\pm)$ [11, section 8.1]. To be more precise, note that $\|e^m\|_{L_\infty(\Omega^\pm)} = \sup_{\Omega^\pm}(\pm e^m)$, while the operator $\mathcal{L}+\kappa_{m,m}$ is linear, so it suffices to get (5.8) only for Ω^+. Set $M := \sup_{\Omega^+}\big\{[\mathcal{L}+\kappa_{m,m}]e\big\}$. Then by the maximum principle in Ω^+, $M\ge 0$. Consequently, $c\ge 0$ implies that $[\mathcal{L}+\kappa_{m,m}](M-\kappa_{m,m}e^m)\ge 0$ in Ω^+, so another application of the maximum principle yields $M-\kappa_{m,m}e^m\ge 0$, which immediately yields (5.8) for Ω^+. Thus, (5.4) is proved.

6. Maximum norm error analysis for finite difference discretizations. Consider our problem (1.1), (1.3) in the spatial domain $\Omega = (0,1)^d \subset \mathbb{R}^d$. Let $\bar\Omega_h$ be the tensor product of d uniform meshes $\{ih\}_{i=0}^N$, with $\Omega_h := \bar\Omega_h\backslash\partial\Omega$ denoting the set of interior mesh nodes. Now, consider a finite difference discretization in the form (2.5), where δ_t^α is defined by (1.4). Let the discrete spatial operator $\mathcal{L}_h$ in (2.5) be

a standard finite difference operator defined, using the standard orthonormal basis $\{\mathbf{i}_k\}_{k=1}^d$ in $\mathbb{R}^d$ (such that $z = (z_1, \dots, z_d) = \sum_{k=1}^d z_k\, \mathbf{i}_k$ for any $z \in \mathbb{R}^d$), by

$$\begin{aligned}&\mathcal{L}_h V(z)\\ &:= \sum_{k=1}^d h^{-2} \left\{ a_k \left(z + \frac{1}{2} h \mathbf{i}_k \right) \left[V(z) - V(z + h\mathbf{i}_k) \right] + a_k \left(z - \frac{1}{2} h \mathbf{i}_k \right) \left[V(z) - V(z - h\mathbf{i}_k) \right] \right\}\\ &\quad + \sum_{k=1}^d \frac{1}{2} h^{-1}\, b_k(z) \left[V(z + h\mathbf{i}_k) - V(z - h\mathbf{i}_k) \right] + c(z)\, V(z) \qquad \text{for } z \in \Omega_h.\end{aligned}$$

(Here the terms in the first and second sums respectively discretize $-\partial_{x_k}(a_k\, \partial_{x_k} u)$ and $b_k\, \partial_{x_k} u$ from (1.3).) The error of this method will be bounded in the nodal maximum norm, denoted $\| \cdot \|_{L_\infty(\Omega_h)} := \max_{\Omega_h} | \cdot |$.

We shall assume that h is sufficiently small so that $\mathcal{L}_h$ satisfies the discrete maximum principle:

$$h^{-1} \geq \max_{k=1,\dots,d} \left\{ \frac{1}{2} \|b_k\|_{L_\infty(\Omega)}\, \|a_k^{-1}\|_{L_\infty(\Omega)} \right\}. \tag{6.1}$$

Hence, the spatial discrete operator $\mathcal{L}_h$ is associated with an M-matrix, so Lemma 2.5 applies to our discretization.

THEOREM 6.1.

(i) *Let the temporal mesh satisfy* (3.1) *with* $r \geq 1$, *and let* $\lambda \tau_j^\alpha < \{\Gamma(2-\alpha)\}^{-1}$ $\forall\, j \geq 1$. *Suppose that* u *is a unique solution of* (1.1), (1.3) *in* $\Omega = (0,1)^d$ *with the initial condition* $u_0 \in L_\infty(\Omega)$ *and under assumption* A1 *on* f. *Also, suppose that* $\|\partial_t^l u(\cdot, t)\|_{L_\infty(\Omega)} \lesssim 1 + t^{\alpha - l}$ *for* $l = 1, 2$ *and* $t \in (0, T]$, *and* $\|\partial_{x_k}^l u(\cdot, t)\|_{L_\infty(\Omega)} \lesssim 1$ *for* $l = 3, 4$, $k = 1, \dots, d$ *and* $t \in (0, T]$. *Then, under condition* (6.1) *on the above* $\mathcal{L}_h$, *there exists a unique solution* $\{U^j\}_{j=0}^M$ *of* (2.5), *and*

$$\|u(\cdot, t_m) - U^m\|_{L_\infty(\Omega_h)} \lesssim \mathcal{E}^m + t_m^\alpha\, h^2 \qquad \forall\, m = 1, \dots, M, \tag{6.2}$$

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

(ii) *If, additionally,* f *satisfies* A2, *and* $\sigma_1 \leq u_0 \leq \sigma_2$, *then* $\sigma_1 \leq U^m \leq \sigma_2$ $\forall\, m \geq 0$.

Proof. We imitate the proof of Theorem 5.1. The existence of a unique solution $\{U^m\}$ follows from Lemma 2.5(i), while part (ii) follows from Lemma 2.5(ii), (iii).

It remains to establish (6.2). For the error $e^m := u(\cdot, t_m) - U^m$, we get a version of (5.2) in Ω_h (instead of Ω) with $\mathcal{L}$ replaced by $\mathcal{L}_h$ and r^m replaced by $r^m + r_h^m$, where $r_h^m := (\mathcal{L}_h - \mathcal{L}) u(\cdot, t_m)$ is the truncation error associated with the spatial discretization. For the latter, a standard calculation yields $|r_h^m| \lesssim h^2$.

Next, we get the following version of (5.4):

$$(\delta_t^\alpha - \lambda) \|e^m\|_{L_\infty(\Omega_h)} \leq \|r^m + r_h^m\|_{L_\infty(\Omega_h)} \qquad \forall\, m \geq 1. \tag{6.3}$$

The proof of the latter closely imitates the proof of (5.4) for $p = \infty$, only now $x^* \in \Omega_h$ is such that $\max_{x \in \Omega_h} |e^m(x)| = |e^m(x^*)|$, and a version of (5.6) holds true with Ω replaced by Ω_h, $\mathcal{L}$ by $\mathcal{L}_h$, and r^m by $r^m + r_h^m$. Finally, in view of (6.1) combined with $c \geq 0$, one gets $\varsigma^m \mathcal{L}_h e^m(x^*) \geq 0$, and hence (6.3).

Let $E^0 = E_h^0 = 0$ and also $(\delta_t^\alpha - \lambda) E^m = \|r^m\|_{L_\infty(\Omega_h)}$ and $(\delta_t^\alpha - \lambda) E_h^m = \|r_h^m\|_{L_\infty(\Omega_h)} \lesssim h^2$. Then, applying Corollary 2.4 to (6.3), one gets $\|e^m\|_{L_\infty(\Omega_h)} \leq E^m + E_h^m$. Also, exactly as in the proof of Theorem 4.1, $E^m \lesssim \mathcal{E}^m$. For E_h^m, in view of Theorem 3.1, the stability property (1.5) with $\gamma = -1$ yields $E_h^m \lesssim t_m^\alpha\, h^2$. Combining these findings, one gets (6.2). ∎

7. Error analysis for finite element discretizations. Throughout this section, we restrict our consideration to the case $\mathcal{L} = -\triangle u = -\sum_{k=1}^{d} \partial_{x_k}^2$ (i.e., $a_k = 1$, $b_k = 0$ for $k = 1, \dots, d$, and $c = 0$ in (1.3)). Then we discretize (1.1), posed in a general bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, by applying a standard finite element spatial approximation to the temporal semidiscretization (2.1). Let $S_h \subset H_0^1(\Omega) \cap C(\bar{\Omega})$ be a Lagrange finite element space of fixed degree $\ell \geq 1$ 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, for $m = 1, \dots, M$, let $u_h^m \in S_h$ satisfy

$$\langle \delta_t^\alpha u_h^m, v_h \rangle_h + \langle \nabla u_h^m, \nabla v_h \rangle + \langle f(\cdot, t_m, u_h^m), v_h \rangle_h = 0 \qquad \forall v_h \in S_h \tag{7.1}$$

with $u_h^0 = u_0$. Here $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle_h$ respectively denote the exact $L_2(\Omega)$ inner product and, possibly, its quadrature approximation.

Our error analysis will employ the standard Ritz projection $\mathcal{R}_h u(t) \in S_h$ of $u(\cdot, t)$ defined by

$$\langle \nabla \mathcal{R}_h u, \nabla v_h \rangle = \langle -\triangle u, v_h \rangle_h \qquad \forall v_h \in S_h, \quad t \in [0, T]. \tag{7.2}$$

7.1. Lumped-mass linear finite elements: Error analysis in the $L_\infty(\Omega)$ norm. First, we consider lumped-mass linear finite element discretizations, i.e., $\ell = 1$ and $\langle \cdot, \cdot \rangle_h$ in (7.1) is defined using the quadrature rule $Q_T[v] := \int_T v^I$, where v^I is the standard linear Lagrange interpolant.

Let $\mathcal{N}$ denote the set of interior mesh nodes, with the corresponding piecewise-linear basis hat functions $\{\phi_z\}_{z \in \mathcal{N}}$. Then, using $v_h = \phi_z$ in (7.1), our discretization can be represented in the form of the discrete problem (2.5) for the nodal values of the computed solution $U^m(z) := u_h^m(z)$, with

$$\mathcal{L}_h U^m(z) := \frac{\langle \nabla u_h^m, \nabla \phi_z \rangle}{\langle 1, \phi_z \rangle_h} \qquad \forall\, z \in \mathcal{N}. \tag{7.3}$$

We shall additionally assume that the spatial triangulation is such that $\mathcal{L}_h$ is associated with an M-matrix (sufficient conditions for this are discussed in Remark 7.2). Hence, Lemma 2.5 applies to our finite element discretization.

Our main result for this discretization is the following.

THEOREM 7.1 (lumped-mass linear elements).

(i) *Let the temporal mesh satisfy* (3.1) *with* $r \geq 1$, *and let* $\lambda \tau_j^\alpha < \{\Gamma(2-\alpha)\}^{-1}$ $\forall\, j \geq 1$. *Suppose that* u *is a unique solution of* (1.1), (1.3) *with the initial condition* $u_0 \in L_\infty(\Omega)$ *and under assumption* A1 *on* f. *Also, suppose that* $\|\partial_t^l u(\cdot, t)\|_{L_\infty(\Omega)} \lesssim 1 + t^{\alpha - l}$ *for* $l = 1, 2$ *and* $t \in (0, T]$. *Then, if the operator* $\mathcal{L}_h$ *from* (7.3) *is associated with an M-matrix, there exists a unique solution* $\{u_h^j\}_{j=0}^M$ *of* (7.1), *and, for* $m = 1, \dots, M$,

$$\|u(\cdot, t_m) - u_h^m\|_{L_\infty(\Omega)} \lesssim \mathcal{E}^m + \max_{t \in \{0, t_m\}} \|\rho(\cdot, t)\|_{L_\infty(\Omega)} + \int_0^{t_m} \|\partial_t \rho(\cdot, t)\|_{L_\infty(\Omega)}\, dt, \tag{7.4}$$

where $\mathcal{E}^m$ *is defined in* (4.2), *and* $\rho(\cdot, t) := \mathcal{R}_h u(t) - u(\cdot, t)$ *is the error of the Ritz projection* (7.2).

(ii) *If, additionally,* f *satisfies* A2, *and* $\sigma_1 \leq u_0 \leq \sigma_2$, *then* $\sigma_1 \leq u_h^m \leq \sigma_2$ $\forall\, m \geq 0$.

Proof. We imitate the proofs of Theorems 5.1 and 6.1. First, since our discretization can be represented in the form of the discrete problem (2.5) for the nodal values of the computed solution $U^m(z) = u_h^m(z)$, the existence of a unique solution $\{U^m\}$ follows from Lemma 2.5(i), while part (ii) follows from Lemma 2.5(ii), (iii).

It remains to establish (7.4). Note that $u(\cdot,t_m)-u_h^m = [\mathcal{R}_h u(\cdot,t_m)-u_h^m]-\rho(\cdot,t_m)$, where $\mathcal{R}_h u(\cdot,t_m) - u_h^m \in S_h$. Hence, it suffices to prove the desired bound for the nodal values of the latter, which will be denoted by $e^m := \mathcal{R}_h u(\cdot,t_m) - U^m \ \forall\, z \in \mathcal{N}$.

In view of (7.3), one has $\mathcal{L}_h \mathcal{R}_h u(z,t_m) = -\triangle u(z,t_m)$. Or, equivalently, using (1.1) and the truncation error $r^m = \delta_t^\alpha u(\cdot,t_m) - D_t^\alpha u(\cdot,t_m)$, one can rewrite it as

$$\delta_t^\alpha u(z,t_m) + \mathcal{L}_h \mathcal{R}_h u(z,t_m) + f(z,t_m,u(z,t_m)) = r^m \qquad \forall\, z \in \mathcal{N},\ \forall\, m \ge 1.$$

Subtracting the nodal representation (2.5) of our discretization, one gets $e^0 = \rho^0$ and

$$\delta_t^\alpha [e^m-\rho^m]+\mathcal{L}_h e^m+[f(\cdot,t_m,u(\cdot,t_m))-f(\cdot,t_m,U^m)] = r^m \quad \forall\, z \in \mathcal{N},\ \forall\, m \ge 1 \tag{7.5}$$

(which is a version of (4.3)), where we used the notation $\rho^m := \rho(\cdot,t_m)$ at any $z \in \mathcal{N}$. Next, using the constant $\lambda \ge 0$ from assumption A1 on f, set

$$p^m := \lambda + \begin{cases} \frac{f(\cdot,t_m,u(\cdot,t_m))-f(\cdot,t_m,U^m)}{u(\cdot,t_m)-U^m} & \text{if } u(\cdot,t_m) \ne U^m, \\ 0 & \text{otherwise,} \end{cases} \qquad \forall\, z \in \mathcal{N},\ \forall\, m \ge 1.$$

Then, in view of A1, $p^m \ge 0$. Also, $f(\cdot,t_m,u(\cdot,t_m))-f(\cdot,t_m,U^m) = (p^m-\lambda)[e^m-\rho^m]$, so (7.5) can be rewritten as

$$(\delta_t^\alpha + \mathcal{L}_h + p^m - \lambda)e^m = r^m + (p^m-\lambda)\rho^m + \delta_t^\alpha \rho^m \qquad \forall\, z \in \mathcal{N},\ \forall\, m \ge 1. \tag{7.6}$$

This is a linear version of (2.5), so, on the one hand, in view of Lemma 2.5(ii), (iii), we can construct upper and lower solutions to estimate e^m. On the other hand, we can separately estimate the components of the error that correspond to the three terms in the right-hand side of (7.6).

First, suppose that the right-hand side of (7.6) equals r^m and $e^0 = 0$. Then for E^m such that $E^0 = 0$ and also $(\delta_t^\alpha - \lambda)E^m = \|r^m\|_{L_\infty(\Omega)}$, exactly as in the proof of Theorem 4.1, one gets $E^m \lesssim \mathcal{E}^m$. Also, by (7.3), $(\mathcal{L}_h + p^m)E^m = p^m E^m \ge 0$. Hence, the pair $\pm E^m$ gives discrete upper and lower solutions for (7.6) in this case. So $|e^m| \le E^m \lesssim \mathcal{E}^m$, and the desired bound of type (7.4) on $\|e^m\|_{L_\infty(\Omega)}$ follows.

Next, suppose that $e^0 = \rho^0$ and the right-hand side of (7.6) equals $(p^m - \lambda)\rho^m$ (where no upper bound on p^m is available). Let $B^0 = 0$ and $(\delta_t^\alpha - \lambda)B^m = 1$, so, in view of Theorem 3.1, $0 \le B^m \lesssim t_m^\alpha$. Next, note that $(\delta_t^\alpha - \lambda)[2\lambda B^m + 1] = \lambda$, while $(\mathcal{L}_h + p^m)[2\lambda B^m + 1] = p^m[2\lambda B^m + 1] \ge p^m$. Consequently, the pair of functions $\pm[2\lambda B^m + 1] \sup_{[0,t_M]} \|\rho\|_{L_\infty(\Omega)}$ gives discrete upper and lower solutions for (7.6) in this case. Hence, $|e^m| \le [2\lambda B^m + 1] \sup_{[0,t_M]} \|\rho\|_{L_\infty(\Omega)}$, so one immediately gets $\|e^M\|_{L_\infty(\Omega)} \lesssim \sup_{[0,t_M]} \|\rho\|_{L_\infty(\Omega)}$. As a similar argument applies for any $M \ge 1$, we deduce the desired bound of type (7.4) on $\|e^m\|_{L_\infty(\Omega)}$.

In a similar manner, consider (7.6) with the right-hand side equal to $\delta_t^\alpha \rho^m$ and $e^0 = 0$. Let $\bar\rho^m := \int_0^{t_m} \|\partial_t \rho(\cdot,s)\|_{L_\infty(\Omega)}\, ds$, for which, in view of (1.4), one gets $|\delta_t \rho^m| \le \delta_t \bar\rho^m$, and so $|\delta_t^\alpha \rho^m| \le \delta_t^\alpha \bar\rho^m$. Consequently, the pair of functions $\pm[\bar\rho^m + \lambda \bar\rho^M B^m]$ gives discrete upper and lower solutions for (7.6) in this case. Hence, $\|e^M\|_{L_\infty(\Omega)} \lesssim \bar\rho^M$. Applying a similar argument for any $M \ge 1$, we again deduce the desired bound of type (7.4) on $\|e^m\|_{L_\infty(\Omega)}$. ∎

Remark 7.2 ($\mathcal{L}_h$ associated with an M-matrix). The operator $\mathcal{L}_h$ from (7.3) is associated with a normalized stiffness matrix for $-\triangle$. The latter is an M-matrix under the following conditions on the triangulation. For $\Omega \subset \mathbb{R}^2$, let $\mathcal{T}$ be a Delaunay triangulation, i.e., the sum of the angles opposite to any interior edge is less than or equal to π. In the case $\Omega \subset \mathbb{R}^3$, it is sufficient, but not necessary, for the triangulation to be nonobtuse (i.e., with no interior angle in any mesh element exceeding $\frac{\pi}{2}$). For weaker necessary and sufficient conditions, we refer the reader to [35, Lemma 2.1].

Remark 7.3 (Ritz projection). The error bound (7.4) involves ρ, the error of the Ritz projection. For the latter, assuming that the spatial domain Ω is polygonal, convex polyhedral, or smooth, for the considered lumped-mass discretization, one has [17, (5.6)]

$$\|\partial_t^l \rho(\cdot,t)\|_{L_\infty(\Omega)} \lesssim h^{2-q}|\ln h|\Big\{\|\partial_t^l u(\cdot,t)\|_{W_\infty^{2-q}(\Omega)} + \|\partial_t^l \mathcal{L}u(\cdot,t)\|_{W_{d/2}^{2-q}(\Omega)}\Big\},$$

where $l = 0,1$, $q = 0,1$, and $t \in (0,T]$. Thus, under certain realistic assumptions on u (see, e.g., [17, Corollary 5.7 and Remark 5.8]), the error bound (7.4) yields $\|u(\cdot,t_m) - u_h^m\|_{L_\infty(\Omega)} \lesssim \mathcal{E}^m + h^2|\ln h|$.

7.2. Finite elements without quadrature: Error analysis in the $L_2(\Omega)$ norm. Next, consider finite elements of fixed degree $\ell \ge 1$ without quadrature, i.e., with $\langle\cdot,\cdot\rangle_h = \langle\cdot,\cdot\rangle$ in (7.1). We shall need an additional assumption on f.

A1*. Let f satisfy the one-sided Lipschitz condition

$$|f(x,t,s_1) - f(x,t,s_2)| \le \bar\lambda|s_1 - s_2| \qquad \forall s_1,\, s_2 \in \mathbb{R},\ \ x \in \Omega,\ \ t > 0,$$

with some constant $\bar\lambda \ge 0$. (Clearly, $\bar\lambda \ge \lambda$ for λ from A1.)

THEOREM 7.4. *Let the temporal mesh satisfy* (3.1) *with* $r \ge 1$, *and let* $\lambda\tau_j^\alpha < \{\Gamma(2-\alpha)\}^{-1}\ \forall\, j \ge 1$. *Suppose that* u *is a unique solution of* (1.1), (1.3) *with the initial condition* $u_0 \in L_\infty(\Omega)$ *and under assumptions* A1 *and* A1* *on* f. *Also, suppose that* $\|\partial_t^l u(\cdot,t)\|_{L_2(\Omega)} \lesssim 1 + t^{\alpha-l}$ *for* $l = 1,2$ *and* $t \in (0,T]$. *Then, under the condition* $\langle\cdot,\cdot\rangle_h = \langle\cdot,\cdot\rangle$, *there exists a unique solution* $\{u_h^j\}_{j=0}^M$ *of* (7.1), *and, for* $m = 1,\dots,M$,

$$\|u(\cdot,t_m) - u_h^m\|_{L_2(\Omega)} \lesssim \mathcal{E}^m + \max_{t\in[0,t_m]}\|\rho(\cdot,t)\|_{L_2(\Omega)} + \int_0^{t_m}\|\partial_t\rho(\cdot,t)\|_{L_2(\Omega)}\,dt, \tag{7.7}$$

where $\mathcal{E}^m$ *is defined in* (4.2), *and* $\rho(\cdot,t) := \mathcal{R}_h u(t) - u(\cdot,t)$ *is the error of the Ritz projection* (7.2).

Proof. The existence of a unique solution u_h^m is established noting that, in view of A1 and the upper bound on $\lambda\tau_j^\alpha$, at each time level t_m we have a finite element discretization of type (7.1) for the monotone elliptic equation (2.4) (as discussed in the proof of Lemma 2.1(i)). Hence, the latter finite element discretization is equivalent to the minimization of a uniformly convex and continuously differentiable functional on a finite-dimensional space, so the existence of a unique computed solution follows (see, e.g., [27, section 4.3.9]).

It remains to obtain the error bound (7.7), for which we shall partially imitate the proofs of Theorems 5.1 and 7.1. Let $e_h^m := \mathcal{R}_h u(t_m) - u_h^m \in S_h$ and $\rho^m := \rho(\cdot,t_m)$. Then $u(\cdot,t_m) - u_h^m = e_h^m - \rho^m$, so it suffices to prove the desired bounds for e_h^m. Now, a standard calculation using (7.1) (in which $\langle\cdot,\cdot\rangle_h = \langle\cdot,\cdot\rangle$) and (1.1) yields

$$\langle\delta_t^\alpha e_h^m, v_h\rangle + \langle\nabla e_h^m, \nabla v_h\rangle + \langle f(\cdot,t_m,u(\cdot,t_m)) - f(\cdot,t_m u_h^m), v_h\rangle = \langle\delta_t^\alpha\rho^m + r^m, v_h\rangle \tag{7.8}$$

$\forall\, v_h \in S_h$. Here we again use the truncation error $r^m = \delta_t^\alpha u(\cdot,t_m) - D_t^\alpha u(\cdot,t_m)$, for which we again have (5.3) with $p=2$. Next, note that $u(\cdot,t_m) = u_h^m + e_h^m - \rho^m$. So, setting $v_h := e_h^m$ and recalling A1*, we arrive at

$$\langle \delta_t^\alpha e_h^m, e_h^m\rangle + \langle f(\cdot,t_m,u_h^m+e_h^m) - f(\cdot,t_m u_h^m), e_h^m\rangle \le \langle r^m + \delta_t^\alpha \rho^m, e_h^m\rangle + \bar\lambda\langle |\rho^m|, |e_h^m|\rangle.$$

The left-hand side here is estimated using a version of (5.5) (with e^m replaced by e_h^m and U^m replaced by u_h^m). Hence, we get the following version of (5.4):

$$(\delta_t^\alpha - \lambda)\|e^m\|_{L_2(\Omega)} \le \|r^m\|_{L_2(\Omega)} + \|\delta_t^\alpha \rho^m\|_{L_2(\Omega)} + \bar\lambda\|\rho^m\|_{L_2(\Omega)} \qquad \forall\, m \ge 1, \tag{7.9}$$

subject to $e_h^0 = \rho^0$.

Let $E^0 = B^0 = 0$, and also $(\delta_t^\alpha - \lambda)E^m = \|r^m\|_{L_2(\Omega)}$ and $(\delta_t^\alpha - \lambda)B^m = 1$. Then, exactly as in the proof of Theorem 4.1, one gets $E^m \lesssim \mathcal{E}^m$. Also, in view of Theorem 3.1, $0 \le B^m \lesssim t_m^\alpha$. Additionally, consider $\bar\rho^m := \int_0^{t_m} \|\partial_t \rho(\cdot,s)\|_{L_2(\Omega)}\, ds$, for which, in view of (1.4), one gets $\|\delta_t \rho^m\|_{L_2(\Omega)} \le \delta_t \bar\rho^m$, and so $\|\delta_t^\alpha \rho^m\|_{L_2(\Omega)} \le \delta_t^\alpha \bar\rho^m$. Consequently, $(\delta_t^\alpha - \lambda)\bar\rho^m \ge \|\delta_t^\alpha \rho^m\|_{L_2(\Omega)} - \lambda\bar\rho^M$. Combining these findings, one concludes that the function

$$\|\rho^0\|_{L_2(\Omega)} + E^m + \bar\rho^m + \Big(\lambda\|\rho^0\|_{L_2(\Omega)} + \lambda\bar\rho^M + \bar\lambda \max_{j=0,\dots,M} \|\rho^j\|_{L_2(\Omega)}\Big) B^m$$

is an upper solution for problem (7.9). Hence, in view of Corollary 2.4, one gets the desired bound (7.7) for $m = M$. Applying a similar argument for any $M \ge 1$, we again deduce the desired bound $\forall m \ge 1$. ∎

Remark 7.5 (Ritz projection). The error bound (7.7) involves ρ, the error of the Ritz projection. For the latter, assuming that the spatial domain Ω is smooth or convex (or, more generally, such that $\|v\|_{W_2^2(\Omega)} \lesssim \|\mathcal{L}v\|_{L_2(\Omega)}$ for any sufficiently smooth v), one has

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

For $l=0$, see, e.g., [2, 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 $\dot u := \partial_t u$. Thus, under certain realistic assumptions on u (see, e.g., [17, Corollary 5.3 and Remark 5.4]), (7.7) yields $\|u(\cdot,t_m) - u_h^m\|_{L_2(\Omega)} \lesssim \mathcal{E}^m + h^{\ell+1}$.

Remark 7.6 (more general $\mathcal{L}$). Theorem 7.4 can be immediately extended to the case of more general $\mathcal{L} = \mathcal{L}(t)$ associated with a coercive bilinear form. The only modification required in the proof is to replace $\langle \nabla e_h^m, \nabla v_h\rangle$ in (7.8) by $\langle \mathcal{L}(t_m) e_h^m, v_h\rangle$. As $\langle \mathcal{L}(t_m) e_h^m, e_h^m\rangle \ge 0$, we again get (7.9), so the remainder of the proof works without any further changes. Note that the estimation of the error of the Ritz projection (such as discussed in Remark 7.5) will be more complicated in this case.

8. Generalizations.

8.1. A2 satisfied, but not A1. Suppose that f in (1.1) satisfies A2 but not A1 (as, e.g., in the Fisher equation with $f = u^2 - u$), and the initial condition is such that $\sigma_1 \le u_0 \le \sigma_2$. Also, let f be continuous in s and satisfy $f(\cdot,t,s) \in L_\infty(\Omega)\ \forall t > 0$ and $s \in [\sigma_1,\sigma_2]$.

Then one can replace f with a standard modification $\widetilde f = \widetilde f(\cdot,t,s)$ defined by $\widetilde f := f$ for $s \in [\sigma_1,\sigma_2]$, and $\widetilde f = f(\cdot,t,\sigma_1)$ for $s \le \sigma_1$, and $\widetilde f = f(\cdot,t,\sigma_2)$ for $s \ge \sigma_2$. Clearly $\widetilde f$ satisfies both A1 and A2, as well as A1*, so all our results on existence, uniqueness, and convergence properties of the discrete solutions will apply. Furthermore, with

the exception of Theorem 7.4, the computed solutions will lie between σ_1 and σ_2; hence they will also be (not necessarily unique) solutions of the corresponding discrete problems with the original f. Note also that the nonlinear discrete problems with $\widetilde{f}$ may be computationally more stable.

8.2. Nonhomogeneous Dirichlet boundary condition. Suppose that $u = \varphi$ on $\partial\Omega \times (0,T]$ in (1.1), where $\varphi(\cdot,t) \in H^1(\Omega) \cap L_\infty(\Omega)\ \forall\, t \in (0,T]$, while $\sigma_1 \le \varphi \le \sigma_2$ on $\partial\Omega \times (0,T]$ in A2. Then, with the obvious modifications $U^m = \varphi(\cdot,t_m)$ on $\partial\Omega$ in (2.1) and $\bar{U}^m \ge \varphi(\cdot,t_m)$ on $\partial\Omega$ in (2.2), and a similar change in (2.5), all results of section 2 remain valid. In particular, in the proof of Lemma 2.1(i), the existence of a unique solution of (2.4) such that $U^m - \varphi(\cdot,t_m) \in H_0^1(\Omega) \cap L_\infty(\Omega)$ can be shown imitating the argument used in the final paragraph of [9, section 2.1]. Furthermore, all error bounds of sections 5 and 6 remain valid for this case. Similarly, the results of section 7 can also be extended for this case with the obvious changes $u_h^m - \varphi_h(\cdot,t_m) \in S_h$ in (7.1) and $\mathcal{R}_h u - \varphi_h \in S_h$ in (7.2), where φ_h is a standard Lagrange interpolant of φ, except the bounds on the Ritz projection in Remarks 7.3 and 7.5 should now take into account the error induced by $\varphi - \varphi_h$.

8.3. Periodic boundary conditions. As many of our arguments rely on the discrete maximum principle for the spatial operator $\mathcal{L}_h$, they can easily be extended to other types of boundary conditions. In particular, the results of section 6 for finite difference discretizations in $\Omega = (0,1)^d$, including Theorem 6.1, apply to the case of periodic boundary conditions (with standard modifications in (2.5) to reflect such boundary conditions). Note that a version of Lemma 2.5 from section 2.2 also holds true for this case assuming that the strict version of $\lambda\tau_j^\alpha \le \{\Gamma(2-\alpha)\}^{-1}$ is satisfied.

8.4. Neumann/Robin and mixed boundary conditions. Suppose that on a subset $\partial\Omega_R$ of the boundary $\partial\Omega$, the Dirichlet boundary condition in (1.1) is replaced by the homogeneous Neumann/Robin boundary condition of the form

$$\frac{\partial u}{\partial n} + \mu u = 0 \qquad \text{on } \ \partial\Omega_R \subseteq \partial\Omega, \qquad \text{where } \ \mu(x,t) \ge 0. \tag{8.1}$$

Then Lemma 2.5 from section 2.2 remains true provided that $\partial\Omega$ in (2.5) is replaced by $\partial\Omega \backslash \partial\Omega_R$, so Ω_h includes the nodes on $\partial\Omega_R$, and also the strict version of $\lambda\tau_j^\alpha \le \{\Gamma(2-\alpha)\}^{-1}$ is satisfied. (In fact, the latter is required only if $\partial\Omega_R = \partial\Omega$ and $\mu = 0$ on $\partial\Omega$.) Now, consider the treatment of (8.1) in finite difference and finite element approximations separately.

8.4.1. Finite difference discretizations. The material of section 6 also can be extended for (8.1). Using the standard finite difference discretization of the Robin boundary conditions (see, e.g., [32, section VII.1.9]), we modify the definition of $\mathcal{L}_h V(z)$ for $z \in \partial\Omega_R$ as follows. Whenever $z \in \partial\Omega_R$ and $z \pm h\mathbf{i}_k \notin \Omega$, we replace $V(z \pm h\mathbf{i}_k)$ in $\mathcal{L}_h V(z)$ by $V(z \mp h\mathbf{i}_k) + 2h\,\mu(z,t_m)\,V(z)$.

The same condition (6.1) ensures that $\mathcal{L}_h$ satisfies the discrete maximum principle also in this case. However, we need to modify the proof of Theorem 6.1, as the truncation error associated with the spatial discretization $r_h^m = (\mathcal{L}_h - \mathcal{L})u(\cdot,t_m)$ is only $O(h)$ on $\partial\Omega_R$ (while $|r_h^m| \lesssim h^2$ on $\Omega_h \backslash \partial\Omega_R$).

THEOREM 6.1*. *Let the coefficients $\{a_k\}$ in* (1.3) *be positive constants, and $\partial\Omega_R \subseteq \partial\Omega$. Then Theorem* 6.1 *holds true for the above finite difference discretization with $t_m^\alpha h^2$ in the right-hand side of the error bound* (6.2) *replaced by h^2.*

Proof. Imitating the proof of Theorem 6.1, we again get the following version of (5.2) in Ω_h (only now Ω_h includes the nodes on $\partial\Omega_R$):

$$\delta_t^\alpha e^m + \mathcal{L}_h e^m + [f(\cdot, t_m, U^m + e^m) - f(\cdot, t_m, U^m)] = r^m + r_h^m \qquad \forall\, m \ge 1. \tag{8.2}$$

Next, similarly to obtaining (7.6) in the proof of Theorem 7.1, introduce $p^m \ge 0$ such that the above is rewritten in the form

$$(\delta_t^\alpha + \mathcal{L}_h + p^m - \lambda)e^m = r^m + r_h^m \qquad \forall\, m \ge 1. \tag{8.3}$$

Set $r_R^m := 0$ in Ω and $r_R^m := r_h^m = O(h)$ on $\partial\Omega_R$. As the above is a linear version of (8.2), we can separately estimate the components of the error that correspond to r_R^m and $r^m + (r_h^m - r_R^m)$. For the latter, exactly as in the proof of Theorem 6.1, we get a version of (6.3) with r_h^m replaced by $(r_h^m - r_R^m) = O(h^2)$, so the desired error bound of type (6.2) for this component of the error follows.

The remaining component of the error satisfies (8.3) with the right-hand-side r_R^m, and, in view of Lemma 2.5(ii), (iii), can be estimated using upper and lower solutions. To simplify the presentation, we shall assume that $\partial\Omega_R \subset \{x_1 = 1\}$ and b_1 is constant (as the other cases are similar). Let $B^0 = 0$ and $(\delta_t^\alpha - \lambda)B^m = 1$, so, in view of Theorem 3.1, $0 \le B^m \lesssim t_m^\alpha$. A calculation shows that $\mathcal{L}_h x_1 \ge -|b_1|$ in Ω (where we exploit that the coefficient a_1 is constant), while $\mathcal{L}_h x_1 \ge 2h^{-1}a_1$ on $\partial\Omega_R$. Noting that $(\delta_t^\alpha + \mathcal{L}_h + p^m - \lambda)x_1 \ge \mathcal{L}_h x_1 - \lambda$, one can check that the pair of discrete functions

$$\pm h(2a_1)^{-1}[x_1 + (\lambda + |b_1|)B^m] \max_{m=1,\dots,M} \|r_R^m\|_{L_\infty(\partial\Omega_R)}$$

gives an upper and a lower solution for the component of e^m that we are estimating. As $\|r_R^m\|_{L_\infty(\partial\Omega_R)} \lesssim h$, we conclude that this component of the error is $\lesssim h^2$. ∎

8.4.2. Lumped-mass linear finite elements. Next, consider an extension of the material of section 7.1 for (8.1). To simplify the presentation, let $\partial\Omega_R \ne \partial\Omega$ or $\mu > 0$ (to ensure that the Ritz projection is well-defined). In this case, with an obvious modification of S_h, the standard lumped-mass discretization (7.1) will include an additional term $\int_{\partial\Omega_R} \big(\mu(\cdot, t_m)\, u_h^m\, v_h\big)^I$ in the left-hand side. A similar modification applies to the definition of the Ritz projection (7.2), in which the left-hand side now includes an additional term $\int_{\partial\Omega_R} \big(\mu(\cdot, t_m)\, v_h\, \mathcal{R}_h u\big)^I$. Finally, in the definition of $\mathcal{L}_h$ in (7.3), the term $\langle \nabla u_h^m, \nabla \phi_z\rangle$ is now replaced by $\langle \nabla u_h^m, \nabla \phi_z\rangle + \int_{\partial\Omega_R} \big(\mu(\cdot, t_m) u_h^m \phi_z\big)^I$, while $\mathcal{N}$ denotes the set of nodes in $\Omega \cup \partial\Omega_R$. With these modifications, an inspection of the proof of Theorem 7.1 shows that this theorem remains true.

8.4.3. Finite elements without quadrature. Finally, we proceed to an extension of section 7.2. The treatment of the boundary condition (8.1) remains as in section 8.4.2, only all approximate integrals of type $\int_{\partial\Omega_R} (\cdots)^I$ are now replaced by their exact versions $\int_{\partial\Omega_R} (\cdots)$. Then an inspection of the proof of Theorem 7.4 shows that in (7.8) we need to add $\int_{\partial\Omega_R} \mu(\cdot, t_m)\, e_h^m\, v_h$ to $\langle \nabla e_h^m, \nabla v_h\rangle$, and afterward, when we set $v_h := e_h^m$ in (7.8), we now exploit the positivity of $\langle \nabla e_h^m, \nabla e_h^m\rangle$ and $\int_{\partial\Omega_R} \mu(\cdot, t_m)(e_h^m)^2$. Thus, we conclude that Theorem 7.4 remains valid for the considered finite element discretization.

9. Numerical results. As a test problem, consider (1.1) with $\mathcal{L} = -(\partial_{x_1}^2 + \partial_{x_2}^2)$ and an Allen–Cahn type nonlinearity $f = (u^3 - u)/\alpha$, posed in the square spatial domain $\Omega = (0, \pi)^2$ for $t \in [0, 1]$, subject to the initial condition $u(0, t) = u_0 =$

$\frac{2}{5}(2y - x^2)\sin x \sin y$. We shall test the error bound (6.2) of Theorem 6.1(i) given for finite difference discretizations in space combined with the L1 scheme in time. The graded temporal mesh $\{t_j = (j/M)^r\}_{j=0}^M$ will be used in all experiments. The spatial mesh is 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 are computed using the two-mesh principle.

First, note that condition A2 is satisfied with $-\sigma_1 = \sigma_2 = 1$, while the initial condition is in $[\sigma_1, \sigma_2] = [-1, 1]$. In full agreement with Theorem 6.1(ii), we have observed that all our computed solutions were also in this range.

Next, we look into the more interesting case of convergence in positive time $t \gtrsim 1$ and give, in Table 9.1, the maximum nodal errors for the graded temporal meshes with $r = 1$, $r = (2-\alpha)/0.9$, and $r = (2-\alpha)/\alpha$. Recalling Remark 4.3, for $r = 1$ we expect convergence rates in time close to 1. The other two values satisfy $r > 2-\alpha$, for which our error bound (6.2) combined with Remark 4.3 predicts the optimal convergence rate of order $2-\alpha$ with respect to time. This clearly agrees with the computational convergence rates given in Table 9.1. The spatial convergence rates are close to 2, which is also consistent with our theoretical bound.

The global maximum nodal errors for $t \in [0, 1]$ were computed for the optimal grading parameter $r = (2-\alpha)/\alpha$ (see the upper part of Table 9.2), as well as for $r = 1$ and $r = 2-\alpha$ (see the lower part of the same table). In view of Remark 4.4, the theoretical error bound (6.2) predicts the global convergence rates in time close

TABLE 9.1

Maximum nodal errors at $t = 1$ (odd rows) and computational rates q in M^{-q} or N^{-q} (even rows) on the graded mesh with $r = 1$, $r = (2-\alpha)/.9$, and $r = (2-\alpha)/\alpha$.

	Errors and convergence rates in time $N = 2M$				Errors and convergence rates in space $M = N^2$			
	$M = 2^5$	$M = 2^6$	$M = 2^7$	$M = 2^8$	$N = 2^3$	$N = 2^4$	$N = 2^5$	$N = 2^6$
				$r = 1$				
$\alpha = 0.3$	1.88e-3	8.98e-4	4.37e-4	2.15e-4	1.23e-2	2.99e-3	7.49e-4	1.87e-4
	1.07	1.04	1.02		2.05	2.00	2.00	
$\alpha = 0.5$	7.41e-4	3.35e-4	1.58e-4	7.65e-5	8.09e-3	2.07e-3	5.13e-4	1.28e-4
	1.15	1.08	1.05		1.97	2.01	2.00	
$\alpha = 0.7$	1.06e-3	4.83e-4	2.27e-4	1.08e-4	5.87e-3	1.48e-3	3.67e-4	9.14e-5
	1.13	1.09	1.06		1.98	2.02	2.01	
				$r = \frac{2-\alpha}{.9}$				
$\alpha = 0.3$	5.87e-4	1.79e-4	5.49e-5	1.69e-5	1.15e-2	2.81e-3	7.04e-4	1.75e-4
	1.71	1.71	1.70		2.04	2.00	2.00	
$\alpha = 0.5$	3.30e-4	1.09e-4	3.70e-5	1.29e-5	7.88e-3	2.01e-3	4.98e-4	1.24e-4
	1.60	1.56	1.53		1.97	2.01	2.00	
$\alpha = 0.7$	7.14e-4	2.83e-4	1.15e-4	4.75e-5	5.66e-3	1.42e-3	3.49e-4	8.66e-5
	1.33	1.30	1.28		1.99	2.02	2.01	
				$r = \frac{2-\alpha}{\alpha}$				
$\alpha = 0.3$	1.26e-3	4.10e-4	1.32e-4	4.21e-5	1.18e-2	2.82e-3	7.06e-4	1.76e-4
	1.62	1.64	1.65		2.06	2.00	2.01	
$\alpha = 0.5$	3.26e-4	1.03e-4	3.32e-5	1.10e-5	7.87e-3	2.01e-3	4.98e-4	1.24e-4
	1.67	1.63	1.59		1.97	2.01	2.00	
$\alpha = 0.7$	6.77e-4	2.58e-4	1.01e-4	4.02e-5	5.64e-3	1.41e-3	3.48e-4	8.63e-5
	1.39	1.35	1.33		2.00	2.02	2.01	

TABLE 9.2

Global maximum nodal errors for $t \in [0,1]$ (odd rows) and computational rates q in M^{-q} or N^{-q} (even rows) on the graded mesh with $r = (2-\alpha)/\alpha$, $r = 1$, $r = 2-\alpha$.

	Errors and convergence rates in time $r = \frac{2-\alpha}{\alpha}$, $N = \frac{1}{2}M$				Errors and convergence rates in space $r = \frac{2-\alpha}{\alpha}$, $M = N^2$			
	$M = 2^8$	$M = 2^9$	$M = 2^{10}$	$M = 2^{11}$	$N = 2^3$	$N = 2^4$	$N = 2^5$	$N = 2^6$
$\alpha = 0.3$	1.49e-4	4.79e-5	1.55e-5	4.97e-6	1.96e-2	4.82e-3	1.20e-3	3.01e-4
	1.64	1.63	1.64		2.02	2.00	2.00	
$\alpha = 0.5$	3.91e-4	1.43e-4	5.20e-5	1.88e-5	1.24e-2	3.18e-3	7.95e-4	1.98e-4
	1.45	1.46	1.47		1.97	2.00	2.01	
$\alpha = 0.7$	8.90e-4	3.83e-4	1.63e-4	6.83e-5	1.43e-2	3.63e-3	8.76e-4	2.12e-4
	1.22	1.24	1.25		1.98	2.05	2.05	
	Errors and convergence rates in time $r = 1$, $N = \frac{1}{128}M$				$r = 2-\alpha$, $N = \frac{1}{4}M$			
	$M = 2^{15}$	$M = 2^{16}$	$M = 2^{17}$	$M = 2^{18}$	$M = 2^{10}$	$M = 2^{11}$	$M = 2^{12}$	$M = 2^{13}$
$\alpha = 0.3$	1.30e-2	1.13e-2	9.77e-3	8.37e-3	9.77e-3	7.47e-3	5.59e-3	4.11e-3
	0.20	0.21	0.22		0.39	0.42	0.45	
$\alpha = 0.5$	2.73e-3	1.95e-3	1.39e-3	9.88e-4	2.73e-3	1.64e-3	9.88e-4	5.94e-4
	0.49	0.49	0.49		0.73	0.73	0.73	
$\alpha = 0.7$	3.15e-4	1.93e-4	1.19e-4	7.33e-5	9.84e-4	5.27e-4	2.82e-4	1.51e-4
	0.70	0.70	0.70		0.90	0.90	0.90	

to αr, which is also in good agreement with the computational convergence rates in Table 9.2.

Overall, we conclude that our numerical results are consistent with our theoretical findings. We also refer the reader to numerical results in [19], which illustrate (for the linear case) that our error bounds are remarkably sharp in the pointwise-in-time sense.

REFERENCES

[1] H. AMANN, *Supersolutions, monotone iterations, and stability*, J. Differential Equations, 21 (1976), pp. 363–377.

[2] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Springer-Verlag, New York, 2008.

[3] A. A. KILBAS, H. M. SRIVASTAVA, AND J. J. TRUJILLO, *Theory and Applications of Fractional Differential Equations*, Elsevier Science, Amsterdam, 2006.

[4] H. BRÉZIS AND W. A. STRAUSS, *Semi-linear second-order elliptic equations in L^1*, J. Math. Soc. Japan, 25 (1973), pp. 565–590.

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

[6] H. BRUNNER, *Collocation Methods for Volterra Integral and Related Functional Differential Equations*, Cambridge University Press, Cambridge, UK, 2004.

[7] K. DIETHELM, *The Analysis of Fractional Differential Equations*, Lecture Notes in Math. 2004, Springer-Verlag, Berlin, 2010.

[8] Q. DU, J. YANG, AND Z. ZHOU, *Time-Fractional Allen-Cahn Equations: Analysis and Numerical Methods*, arXiv:1906.06584, 2019.

[9] A. DEMLOW AND N. KOPTEVA, *Maximum-norm a posteriori error estimates for singularly perturbed elliptic reaction-diffusion problems*, Numer. Math., 133 (2016), pp. 707–742.

[10] L. C. EVANS, *Partial Differential Equations*, AMS, Providence, RI, 1998.

[11] D. GILBARG AND N. S. TRUDINGER, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, 1998.

[12] 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.
[13] B. Ji, H.-L. Liao, and L. Zhang, *Simple Maximum-Principle Preserving Time-Stepping Methods for Time-Fractional Allen-Cahn Equation*, arXiv:1906.11693, 2019.
[14] 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.
[15] B. Jin, B. Li, and Z. Zhou, *Numerical analysis of nonlinear subdiffusion equations*, SIAM J. Numer. Anal., 56 (2018), pp. 1–23.
[16] N. Kopteva, *Maximum norm error analysis of a 2D singularly perturbed semilinear reaction-diffusion problem*, Math. Comp., 76 (2007), pp. 631–646.
[17] 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.
[18] N. Kopteva, *Error analysis of an L2-type method on graded meshes for a fractional-order parabolic problem*, Math. Comp., to appear.
[19] N. Kopteva and X. Meng, *Error analysis for a fractional-derivative parabolic problem on quasi-graded meshes using barrier functions*, SIAM J. Numer. Anal., 58 (2020), pp. 1217–1238.
[20] N. Kopteva and S. B. Savescu, *Pointwise error estimates for a singularly perturbed time-dependent semilinear reaction-diffusion problem*, IMA J. Numer. Anal., 31 (2011), pp. 616–639.
[21] 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.
[22] 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.
[23] H.-L. Liao, T. Tang, and T. Zhou, *A Second-Order and Nonuniform Time-Stepping Maximum-Principle Preserving Scheme for Time-Fractional Allen-Cahn Equations*, arXiv: 1909.10216, 2019.
[24] W. McLean, *Regularity of solutions to a time-fractional diffusion equation*, ANZIAM J., 52 (2010), pp. 123–138.
[25] W. McLean and K. Mustapha, *A second-order accurate numerical method for a fractional wave equation*, Numer. Math., 105 (2007), pp. 481–510.
[26] 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.
[27] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
[28] C. V. Pao, *Nonlinear Parabolic and Elliptic Equations*, Plenum Press, New York, 1992.
[29] C. V. Pao, *Accelerated monotone iterative methods for finite difference equations of reaction-diffusion*, Numer. Math., 79 (1998), pp. 261–281.
[30] C. V. Pao and X. Lu, *Block monotone iterative method for semilinear parabolic equations with nonlinear boundary conditions*, SIAM J. Numer. Anal., 47 (2010), pp. 4581–4606.
[31] 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.
[32] A. A. Samarski, *Theory of Difference Schemes*, Nauka, Moscow, 1989 (in Russian)
[33] M. Stynes, *Too much regularity may force too much uniqueness*, Fract. Calc. Appl. Anal., 19 (2016), pp. 1554–1562.
[34] 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.
[35] J. Xu and L. Zikatanov, *A monotone finite element scheme for convection-diffusion equations*, Math. Comp., 68 (1999), pp. 1429–1446.