INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 3, Number 4, Pages 504–524

ERROR ESTIMATES UNDER MINIMAL REGULARITY FOR SINGLE STEP FINITE ELEMENT APPROXIMATIONS OF PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

L. S. HOU AND W. ZHU

Abstract. This paper studies error estimations for a fully discrete, single step finite element scheme for linear parabolic partial differential equations. Convergence in the norm of the solution space is shown and various error estimates in this norm are derived. In contrast to like results in the extant literature, the error estimates are derived in a stronger norm and under minimal regularity assumptions.

Key Words. fully discrete approximation, parabolic equations, error estimate, finite element methods, backward Euler method.

1. Introduction

This paper is devoted to the study of error estimations for a fully-discrete, single step finite element approximation of linear parabolic equations of the form

$$\frac{\partial u}{\partial t} - \operatorname{div}[A(\mathbf{x})\nabla u] = f(t,\mathbf{x}) \qquad \text{in } (0,T)\times\Omega \tag{1.1}$$

with the boundary and initial conditions

$$u = 0 \quad \text{on } (0,T)\times\partial\Omega, \qquad u(0,\mathbf{x}) = u_0(\mathbf{x}) \quad \text{in } \Omega\,, \tag{1.2}$$

where f is a given function and A is a matrix-valued, uniformly positive definite function. The fully discrete approximation scheme studied in this paper is a simple modification of the standard backward Euler method and it involves a temporal integral of the forcing term. This fully discrete scheme is well defined under minimal regularity assumptions on the forcing term and the initial condition; in particular, the forcing term f can be nondifferentiable in time, e.g., a temporal step function of the form $f = \sum_{i=1}^{J} f_i(t)\chi_{(t_i,t_{i+1})}(t)\Theta_i(\mathbf{x})$ where each (t_i, t_{i+1}) is a time interval in $[0,T]$ and $\chi_{(t_i,t_{i+1})}$ is the characteristic function for the interval (t_i, t_{i+1}) (such a choice of f corresponds to a setting in which different force patterns are applied on different time intervals.) The achievements of this paper include:

- fractional order error estimates in the norm of the solution space (this solution space will be made precise in Section 2.1) are derived under fractional order, uni-directional regularity assumptions on the forcing term;
- a first order δ error estimate (again in the norm of the solution space) is derived under standard assumptions that ensure solution regularity;
- convergence under minimal regularity.

Received by the editors November 18, 2005.
2000 *Mathematics Subject Classification.* 65M12, 65M15, 65M60.

Compared to the convergence results and error estimates in the extant literature for linear parabolic PDEs (see, e.g., [2, 3, 4, 5, 9, 10, 17, 19, 20]), the convergence and error estimates in this paper are derived in a stronger norm and under weaker regularity assumptions.

The fractional order error estimates established in this paper are new and they allow us to prove the convergence under weaker regularity hypotheses on the forcing term and the initial condition than those assumed in standard convergence results in the literature. Other types of fractional order error estimates can be found in the literature, e.g., [14]. The fractional order error estimates of [14] were measured in the $H^{p,p/2}_{\alpha}((0,T)\times\Omega)$ norm (see [14] for this notation) with the forcing term belonging to $H^{2p-1,p-1/2}_{-1}((0,T)\times\Omega)$, whereas those of this paper are measured in the $H^{1,-1}((0,T)\times\Omega)$ norm and requires only uni-directional regularity.

The results of this paper can be used in conjunction with the Brezzi, Rappaz, Raviart theory (see, e.g., [6, 12]) to study fully discrete approximations of semilinear parabolic PDEs. The set-up of the fully-discrete approximations for semilinear PDEs is more involved and is illustrated in [13]. The fractional order error estimates under uni-directional regularity assumptions play a crucial role in the derivation of fully discrete error estimates for semilinear parabolic PDEs; see [13].

Another significant application of the results of this paper is to prove the convergence of and error estimates for fully discrete approximations of optimal control problems constrained by parabolic PDEs; such applications will be discussed elsewhere.

The rest of the paper is organized as follows. In §2, we introduce continuous and discrete (finite element) function spaces and define a weak formulation for the problem (1.1)–(1.2). In §3, we define semidiscrete and fully discrete finite element approximations to that problem. In §4, we derive estimates for the difference between the semidiscrete and fully discrete approximate solutions and, in §5, we establish, under minimal regularity assumptions, convergence of and error estimates for the fully discrete approximation.

2. Function spaces, finite element spaces, and weak formulations

2.1. Function spaces. We use the standard notations (see, e.g. [1]) for Sobolev spaces $W^{s,p}(\Omega)$ for all real s and $p\geq 1$, with their norms denoted by $\|\cdot\|_{W^{s,p}(\Omega)}$. When $p=2$, we use the notation $H^s(\Omega)=W^{s,2}(\Omega)$ for all real s, with their norm simply denoted by $\|\cdot\|_s$. We let $H^1_0(\Omega)$ stand for the completion of $C^\infty_0(\Omega)$ with respect to the $H^1(\Omega)$ norm. Note that $H^0(\Omega)=L^2(\Omega)$ so that the $L^2(\Omega)$ norm is denoted by $\|\cdot\|_0$. The inner products on $L^2(\Omega)$ is denoted by $[\cdot,\cdot]$, i.e.,

$$[u,v]=\int_\Omega uv\,d\mathbf{x}\qquad\forall\,u,v\in L^2(\Omega)\,.$$

The duality paring between a Banach space B and its dual will be generically denoted by $\langle\cdot,\cdot\rangle$.

For a $p\in[1,\infty]$, an interval $(a,b)\subset\mathbb{R}$, and a Banach space B with norm $\|\cdot\|_B$, we denote by $L^p(a,b;B)$ the set of measurable functions $v:(a,b)\to B$ such that $\int_a^b\|v(t)\|_B^p\,dt<\infty$. The norm on $L^p(a,b;B)$ for $p\in[1,\infty)$ is defined by

$$\|v\|_{L^p(a,b;B)}=\Big(\int_a^b\|v(t)\|_B^p\,dt\Big)^{\frac{1}{p}}\qquad\forall\,v\in L^p(a,b;B)\,.$$

The norm on $L^\infty(a,b;B)$ is defined by

$$\|v\|_{L^\infty(a,b;B)} = \underset{(a,b)}{ess\,sup}\, \|v(t)\|_B \qquad \forall\, v \in L^\infty(a,b;B)\,.$$

We denote by $C([a,b];B)$ the set of all continuous functions $v : [a,b] \to B$ with the norm $\|v\|_{C([a,b];B)} = \max_{t\in[a,b]} \|v(t)\|_B$. We introduce

$$\mathcal{W}(a,b) = \Big\{ v \in L^2(a,b;H_0^1(\Omega)) : v' = \partial_t v \in L^2(a,b;H^{-1}(\Omega)) \Big\}$$

where $v' = \partial_t v$ is understood in the scalar distribution sense:

$$\int_a^b \langle v'(t), \phi(t) w\rangle \, dt = -\int_a^b \langle v(t), \phi'(t) w\rangle \, dt \qquad \forall\, \phi \in C_0^\infty(a,b),\ \forall\, w \in H_0^1(\Omega)$$

with $\langle\cdot,\cdot\rangle$ denoting the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$. The norm on $\mathcal{W}(a,b)$ is defined by

$$\|v\|_{\mathcal{W}(a,b)} = \Big(\|u\|^2_{L^2(a,b;H^1(\Omega))} + \|\partial_t v\|^2_{L^2(a,b;H^{-1}(\Omega))} \Big)^{1/2} \qquad \forall\, v \in \mathcal{W}(a,b).$$

The space $\mathcal{W}(a,b)$ is the standard solution space for the linear parabolic PDE [11]. For real numbers $s \geq 0$ and $p \geq 1$, the space $H^s(a,b;B)$ is defined as follows. First,

$$H^s(\mathbb{R};B) = \{ v \in L^2(\mathbb{R};B) : |\tau|^s \widehat{v} \in L^2(\mathbb{R};B) \}$$

endowed with the norm

$$\|v\|_{H^s(\mathbb{R};B)} = \Big(\int_{\mathbb{R}} \|v(t)\|_B^2 \, dt + \int_{\mathbb{R}} |\tau|^{2s} \|\widehat{v}(\tau)\|_B^2 \, d\tau \Big)^{1/2}$$

where $\widehat{v}$ is the temporal Fourier transform of v:

$$\widehat{v}(\tau) = \int_{\mathbb{R}} e^{-2i\pi t\tau} v(t)\, dt\,.$$

Then we set

$$H^s(a,b;B) = \Big\{ v = \widetilde{v}|_{[a,b]} : \widetilde{v} \in H^s(\mathbb{R};B) \Big\}$$

with the norm

$$\|v\|_{H^s(a,b;B)} = \inf_{\substack{\widetilde{v}\in H^s(\mathbb{R};B) \\ \widetilde{v}|_{[a,b]}=v}} \|\widetilde{v}\|_{H^s(\mathbb{R};B)} \qquad \forall\, v \in H^s(a,b;B)\,.$$

A function $v = v(t,\mathbf{x}) \in H^s(a,b;B)$ for some spatial function space B is often simply written as $v(t)$. Further discussions of Banach-space-valued Sobolev spaces $H^s(a,b;B)$ may be found in [15, 18].

Throughout, C denotes a generic constant that may depend on the domain Ω and time T; the value of C varies with context.

2.2. Weak formulation. We introduce the bilinear form

$$a[u,v] = \int_\Omega (A(\mathbf{x})\nabla u)\cdot \nabla v \, d\mathbf{x} \qquad \forall\, u, v \in H^1(\Omega)\,.$$

Given $u_0 \in L^2(\Omega)$ and $f \in L^2(0,T;H^{-1}(\Omega))$, the weak formulation for (1.1) and (1.2) is: seek a $u \in \mathcal{W}(0,T)$ such that

$$\text{(2.1)} \qquad \begin{cases} \big\langle \partial_t u(t), v\big\rangle + a[u(t),v] = \big\langle f(t), v\big\rangle, & \forall\, v \in H_0^1(\Omega),\ \text{a.e. } t \in [0,T]\,, \\ u(0) = u_0 \quad \text{in } L^2(\Omega). \end{cases}$$

The uniform positive definiteness of the matrix function A implies the coercivity for the bilinear form $a[\cdot,\cdot]$:

$$a[v,v] \geq C_a \|v\|_1^2 \qquad \forall\, v \in H_0^1(\Omega)\,. \tag{2.2}$$

It is well known (see, e.g., [11]) that there exists a unique weak solution for (2.1).

2.3. Finite element spaces. In the sequel we assume Ω is a two-dimensional polygon or a three dimensional polyhedron. Let V_h be a family of finite element subspaces of $H_0^1(\Omega)$ defined over a family of regular triangulations of Ω. The parameter h denotes the largest grid size for a given triangulation. For the convenience of stating approximation properties we introduce, as in [8], the spaces $\Phi_0^r(\Omega) = H_0^{\min(1,r)}(\Omega)$ for real r, i.e.,

$$\Phi_0^r(\Omega) = \begin{cases} H_0^1(\Omega) & \text{if } r \geq 1, \\ H_0^r(\Omega) & \text{if } 1/2 < r < 1, \\ H^r(\Omega) & \text{if } r \leq 1/2. \end{cases}$$

We assume that the finite element function space V_h satisfies the following approximation properties:

(i) For every $v \in \Phi_0^s(\Omega)$,

$$\inf_{v_h \in V_h} \|v - v_h\|_s \to 0 \quad \text{as } h \to 0 \qquad s = -1, 0, 1\,; \tag{2.1}$$

(ii) there exists a constant $C > 0$ such that for every $v \in H^{r+1}(\Omega) \cap \Phi_0^{r+1}(\Omega)$ and every $r \in [s-1, k]$,

$$\inf_{v_h \in V_h} \|v - v_h\|_s \leq C h^{r+1-s} \|v\|_{r+1}, \qquad s = -1, 0, 1\,, \tag{2.2}$$

where $k \geq 1$ is a positive integer that is usually determined by the order of the piecewise polynomials used to define V_h.

We also assume that finite element triangulations are uniformly regular so that the following inverse inequality hold:

$$\|v_h\|_1 \leq C h^{-1} \|v_h\|_0 \qquad \forall\, v_h \in V_h\,. \tag{2.3}$$

For detailed discussions of the properties (2.1)–(2.3) and constructions of the finite element spaces with these properties, see, e.g., [7].

We denote by P_h the $L^2(\Omega)$ projection from $L^2(\Omega)$ onto V_h, namely, for each $v \in L^2(\Omega)$,

$$(P_h v - v, w^h) = 0 \qquad \forall\, w^h \in V_h. \tag{2.4}$$

As a consequence of (2.3) we have

$$\|P_h v\|_r \leq C \|v\|_r \quad \forall\, v \in H^r(\Omega) \cap \Phi_0^r(\Omega),\; r \in [0,1]\,; \tag{2.5}$$

see [19] or [8].

3. Semidiscrete and fully discrete approximations of linear parabolic equations

We consider the linear parabolic problem (1.1)–(1.2), or more precisely, the corresponding weak formulation (2.1). We use the notations u, u_h, and $u_{\delta h}$ to respectively denote the exact, semi-discrete approximate, and fully discrete approximate solutions.

3.1. Semidiscrete finite element approximations. Let V_h be a family of finite element subspaces of $H_0^1(\Omega)$ introduced in §2.3. The semidiscrete finite element approximation of (2.1) is defined as follows: seek a $u_h \in H^1(0,T;V_h)$ such that

$$\begin{cases} \big\langle \partial_t u_h(t), v_h \big\rangle + a[u_h(t), v_h] = \big\langle f(t), v_h \big\rangle \qquad \forall\, v_h \in V_h, \text{ a.e. } t \in [0,T]\,, \\ u_h(0) = P_h u_0\,, \end{cases} \tag{3.1}$$

where P_h is the $L^2(\Omega)$-projection operator onto V_h defined in (2.4), i.e., $[u_h(0), v_h] = [u_0, v_h]$ for all $v_h \in V_h$.

We quote the following results of [8] concerning semidiscrete error estimates for linear parabolic problems:

Theorem 3.1. *Assume that $f \in L^2(0,T;H^{-1}(\Omega))$ and $u_0 \in L^2(\Omega)$. Let $u \in \mathcal{W}(0,T)$ be the solution of (2.1) and $u_h \in H^1(0,T;V_h)$ be the solution of (3.1). Then*

$$\|u - u_h\|_{\mathcal{W}(0,T)} \to 0 \quad \text{as } h \to 0. \tag{3.2}$$

If, in addition, $u \in L^2(0,T;H^{r+1}(\Omega)) \cap H^1(0,T;H^{r-1}(\Omega))$ for some $r \in [0,k]$, then

$$\|u - u_h\|_{\mathcal{W}(0,T)} \le Ch^r \Big(\|u\|_{L^2(0,T;H^{r+1}(\Omega))} + \|\partial_t u\|_{L^2(0,T;H^{r-1}(\Omega))} \Big). \tag{3.3}$$

3.2. Fully discrete finite element approximations. We partition $[0,T]$ into $0 = t_0 < t_1 < t_2 < \cdots < t_N = T$ with a uniform time steping δ. For $f \in L^2(0,T;H^{-1}(\Omega))$, we define f^m by

$$f^m = \frac{1}{\delta} \int_{t_{m-1}}^{t_m} f(t)\, dt, \qquad 1 \le m \le N\,. \tag{3.4}$$

The fully discrete approximate solution $u_{\delta h}$ is constructed as follows. We first solve for $U_h^m \in V_h$, $m = 0, 1, 2, \cdots, N$, from

$$\begin{aligned} & U_h^0 = P_h u_0\,, \\ & \Big\langle \frac{U_h^m - U_h^{m-1}}{\delta}, v_h \Big\rangle + a[U_h^m, v_h] = \langle f^m, v_h \rangle \qquad \forall\, v_h \in V_h,\ 1 \le m \le N\,; \end{aligned} \tag{3.5}$$

we then define $u_{\delta h} \in H^1([0,T];V_h)$ by

$$u_{\delta h}(t)\big|_{[t_{m-1},t_m]} = U_h^{m-1} + \Big(\frac{t - t_{m-1}}{\delta}\Big)(U_h^m - U_h^{m-1}), \qquad 1 \le m \le N. \tag{3.6}$$

Note that f^m (and hence the scheme) is well defined even when f has only the minimal regularity $L^2(0,T;H^{-1}(\Omega))$. The following two lemmas summarize some useful properties for $\{f^m\}$.

Lemma 3.2. *Assume $f \in L^2(0,T;H^{-1}(\Omega))$. Then the set $\{f^m\}_{m=1}^N$ defined by (3.4) satisfies*

$$\delta \sum_{m=1}^{N} \|f^m\|_{-1}^2 \le C\|f\|_{L^2(0,T;H^{-1}(\Omega))}^2 \tag{3.7}$$

and

$$\sum_{m=1}^{N} \int_{t_{m-1}}^{t_m} \|f^m - f(t)\|_{-1}^2\, dt \to 0 \quad \text{as } \delta \to 0\,. \tag{3.8}$$

If $f \in H^\gamma(0,T;H^{-1}(\Omega))$ *for some* $\gamma \in [0,1]$*, then*

$$\sum_{m=1}^{N} \int_{t_{m-1}}^{t_m} \|f^m - f(t)\|_{-1}^2 \, dt \le C\delta^{2\gamma} \|f\|^2_{H^\gamma(0,T;H^{-1}(\Omega))} \,. \tag{3.9}$$

Proof. Relations (3.7) and (3.8) follows from [18, p.221, Lemma 4.5] and [18, p.223, Lemma 4.9], respectively.

To prove (3.9) it suffices to examine the cases $\gamma = 0$ and $\gamma = 1$ thanks to interpolation theorems. If $f \in H^0(0,T;H^{-1}(\Omega)) = L^2(0,T;H^{-1}(\Omega))$, then

$$\begin{aligned} &\sum_{m=1}^{N} \int_{t_{m-1}}^{t_m} \|f^m - f(t)\|_{-1}^2 \, dt \\ &\quad \le 2\sum_{m=1}^{N} \int_{t_{m-1}}^{t_m} \|f^m\|_{-1}^2 \, dt + 2\sum_{m=1}^{N} \int_{t_{m-1}}^{t_m} \|f(t)\|_{-1}^2 \, dt \\ &\quad = 2\delta \sum_{m=1}^{N} \|f^m\|_{-1}^2 + 2\|f\|^2_{L^2(0,T;H^{-1}(\Omega))} \le C\|f\|^2_{L^2(0,T;H^{-1}(\Omega))} \,. \end{aligned} \tag{3.10}$$

If $f \in H^1(0,T;H^{-1}(\Omega))$, then for $1 \le m \le N$ and $t \in [t_{m-1}, t_m]$,

$$\begin{aligned} \|f^m - f(t)\|_{-1}^2 &= \Big\| \frac{1}{\delta} \int_{t_{m-1}}^{t_m} \{f(s) - f(t)\} \, ds \Big\|_{-1}^2 \\ &= \frac{1}{\delta^2} \Big\| \int_{t_{m-1}}^{t_m} \int_t^s \partial_t f(r) \, dr \, ds \Big\|_{-1}^2 \le \frac{1}{\delta^2} \Big| \int_{t_{m-1}}^{t_m} \int_{t_{m-1}}^{t_m} \|\partial_t f(r)\|_{-1} \, dr \, ds \Big|^2 \\ &\le \int_{t_{m-1}}^{t_m} \|\partial_t f(r)\|_{-1}^2 \, dr \int_{t_{m-1}}^{t_m} 1 \, dr = \delta \int_{t_{m-1}}^{t_m} \|\partial_t f(r)\|_{-1}^2 \, dr \end{aligned}$$

so that through an integration in t from t_{m-1} to t_m and summations in m we obtain:

$$\begin{aligned} \sum_{m=1}^{N} \int_{t_{m-1}}^{t_m} \|f^m - f(t)\|_{-1}^2 \, dt &\le \delta^2 \sum_{m=1}^{N} \int_{t_{m-1}}^{t_m} \|\partial_t f(r)\|_{-1}^2 \, dr \\ &\le \delta^2 \|f\|^2_{H^1(0,T;H^{-1}(\Omega))} \,. \end{aligned} \tag{3.11}$$

Interpolations between (3.10) and (3.11) yield (3.9). □

Similarly we have:

Lemma 3.3. *Assume* $f \in L^2(0,T;L^2(\Omega))$*. Then the set* $\{f^m\}_{m=1}^N$ *defined by (3.4) satisfies*

$$\delta \sum_{m=1}^{N} \|f^m\|_0^2 \le C\|f\|^2_{L^2(0,T;L^2(\Omega))} \tag{3.12}$$

and

$$\sum_{m=1}^{N} \int_{t_{m-1}}^{t_m} \|f^m - f(t)\|_0^2 \, dt \to 0 \quad \text{as } \delta \to 0 \,. \tag{3.13}$$

If $f \in H^\gamma(0,T;L^2(\Omega))$ *for some* $\gamma \in [0,1]$*, then*

$$\sum_{m=1}^{N} \int_{t_{m-1}}^{t_m} \|f^m - f(t)\|_0^2 \, dt \le C\delta^{2\gamma} \|f\|^2_{H^\gamma(0,T;L^2(\Omega))} \,. \tag{3.14}$$

Our task is to prove the convergence of the fully discrete solutions and ro derive error estimates for the fully discrete approximations defined in (3.5)–(3.6). Using the triangle inequality

$$\|u - u_{\delta h}\|_{\mathcal{W}(0,T)} \le \|u - u_h\|_{\mathcal{W}(0,T)} + \|u_h - u_{\delta h}\|_{\mathcal{W}(0,T)}$$

and recalling that semidiscrete error estimates are already known, we see that our task reduces to estimating the error between the semidiscrete solutions and the fully discrete solutions.

4. Estimation of the errors between semidiscrete and fully discrete approximate solutions

In this section we will estimate the errors between semidiscrete and fully discrete approximate solutions. Specifically, we will derive, respectively, $O(\delta)$ and $O(\delta^\gamma)$ ($\gamma \in [0,1]$) error estimates for $u_h(t) - u_{\delta h}(t)$ under various regularity assumptions.

The following lemma gives some useful estimates for generic solutions of the fully discrete schemes and will be invoked repeatedly in the sequel.

Lemma 4.1. *Assume $g \in L^2(0,T;H^{-1}(\Omega))$ and $W_h^0 \in V_h$. Let $\{W_h^m\}_{m=1}^N \subset V_h$ be defined by*

$$\Big\langle \frac{W_h^m - W_h^{m-1}}{\delta}, v_h \Big\rangle + a[W_h^m, v_h] = \langle g^m, v_h\rangle \qquad \forall\, v_h \in V_h,\ 1 \le m \le N \tag{4.1}$$

where $g^m \equiv \delta^{-1}\int_{t_{m-1}}^{t_m} g(s)\,ds$. Then

$$\sum_{m=1}^N \|W_h^m - W_h^{m-1}\|_1^2 \le \frac{C}{\delta}\|g\|^2_{L^2(0,T;H^{-1}(\Omega))} + C\|W_h^0\|_1^2 \tag{4.2}$$

and

$$\sum_{m=1}^N \|W_h^m\|_1^2 \le \frac{C}{\delta}\|g\|^2_{L^2(0,T;H^{-1}(\Omega))} + \frac{C}{\delta}\|W_h^0\|_0^2\,. \tag{4.3}$$

If, in addition, $g \in L^2(0,T;L^2(\Omega))$, then

$$\sum_{m=1}^N \|W_h^m - W_h^{m-1}\|_1^2 \le C\|g\|^2_{L^2(0,T;L^2(\Omega))} + C\|W_h^0\|_1^2\,. \tag{4.4}$$

Proof. If $g \in L^2(0,T;H^{-1}(\Omega))$, by setting $v_h = W_h^m - W_h^{m-1}$ in (4.1) we obtain

$$\begin{aligned}
&\frac{1}{\delta}\|W_h^m - W_h^{m-1}\|_0^2 + \frac{1}{2}a[W_h^m, W_h^m] \\
&\qquad -\frac{1}{2}a[W_h^{m-1}, W_h^{m-1}] + \frac{1}{2}a[W_h^m - W_h^{m-1}, W_h^m - W_h^{m-1}] \\
&= \langle g^m, W_h^m - W_h^{m-1}\rangle \le \|g^m\|_{-1}\|W_h^i - W_h^{m-1}\|_1 \\
&\le \frac{1}{C_a}\|g^m\|_{-1}^2 + \frac{C_a}{4}\|W_h^m - W_h^{m-1}\|_1^2\,,
\end{aligned}$$

where C_a is the positive coercivity constant in (2.2). The last relation may be simplified as

$$a[W_h^m, W_h^m] - a[W_h^{m-1}, W_h^{m-1}] + \frac{C_a}{2}\|W_h^m - W_h^{m-1}\|_1^2 \le \frac{2}{C_a}\|g^m\|_{-1}^2\,.$$

Summations in m for $m = 1, \cdots, N$ yield

$$a[W_h^N, W_h^N] - a[W_h^0, W_h^0] + \frac{C_a}{2} \sum_{m=1}^{N} \|W_h^m - W_h^{m-1}\|_1^2 \leq \frac{2}{C_a} \sum_{m=1}^{N} \|g^m\|_{-1}^2$$

so that, by dropping the first term and applying Lemma 3.2, we obtain

$$\sum_{m=J}^{K} \|W_h^m - W_h^{m-1}\|_1^2 \leq \frac{C}{\delta}\, \delta \sum_{m=J}^{K} \|g^m\|_{-1}^2 + C\|W_h^{J-1}\|_1^2$$
$$\leq \frac{C}{\delta} \|g\|^2_{L^2(t_{J-1},t_K;H^{-1}(\Omega))} + C\|W_h^{J-1}\|_1^2 \,,$$

which proves (4.2).

Likewise, if $g \in L^2(0,T;L^2(\Omega))$, by setting $v_h = W_h^m - W_h^{m-1}$ in (4.1) we obtain

$$\frac{1}{\delta}\|W_h^m - W_h^{m-1}\|_0^2 + \frac{1}{2}a[W_h^m, W_h^m] - \frac{1}{2}a[W_h^{m-1}, W_h^{m-1}]$$
$$+\frac{1}{2}a[W_h^m - W_h^{m-1}, W_h^m - W_h^{m-1}] = \langle g^m, W_h^m - W_h^{m-1}\rangle$$
$$\leq \|g^m\|_0 \|W_h^m - W_h^{m-1}\|_0 \leq \frac{\delta}{2}\|g^m\|_0^2 + \frac{1}{2\delta}\|W_h^m - W_h^{m-1}\|_0^2 \,.$$

Thus, we have

$$a[W_h^m, W_h^m] - a[W_h^{m-1}, W_h^{m-1}] + C_a\|W_h^m - W_h^{m-1}\|_1^2 \leq \delta\|g^m\|_0^2 \,.$$

Summations in m with an application of Lemma 3.3 yield (4.4).

To prove (4.3), we proceed as follows. Setting $v_h = W_h^m$ in (4.1) we obtain

$$\frac{1}{2\delta}\|W_h^m\|_0^2 - \frac{1}{2\delta}\|W_h^{m-1}\|_0^2 + \frac{1}{2\delta}\|W_h^m - W_h^{m-1}\|_0^2 + a[W_h^m, W_h^m]$$
$$= \langle g^m, W_h^m\rangle \leq \frac{1}{2C_a}\|g^m\|_{-1}^2 + \frac{C_a}{2}\|W_h^m\|_1^2 \,.$$

This relation may be simplified as

$$\frac{1}{2\delta}\|W_h^m\|_0^2 - \frac{1}{2\delta}\|W_h^{m-1}\|_0^2 + \frac{C_a}{2}\|W_h^m\|_1^2 \leq \frac{1}{2C_a}\|g^m\|_{-1}^2 \,.$$

Summations in m for $m = 1, \cdots, N$ yield

$$\frac{1}{2\delta}\|W_h^N\|_0^2 - \frac{1}{2\delta}\|W_h^0\|_0^2 + \frac{C_a}{2}\sum_{m=1}^{N}\|W_h^m\|_1^2 \leq \frac{1}{2\delta C_a}\,\delta \sum_{m=1}^{N}\|g^m\|_{-1}^2$$

so that, by dropping the first term and applying Lemma 3.2, we obtain (4.3). □

The next lemma gives estimates for the error between a generic fully discrete solution and a generic semidiscrete solution. In the subsequent application of this lemma in this section (in the proof of Theorem 4.7) we will choose the generic discrete solutions to be the semidiscrete and fully discrete solutions defined in Sections 3.1 and 3.2, respectively. By choosing the semidiscrete solution as the zero solution thereby obtaining an estimate for the generic fully discrete solution we may apply this lemma to estimate a term in the proof of Theorem 5.1.

Lemma 4.2. *Assume* $g, \overline{g} \in L^2(0,T;H^{-1}(\Omega))$ *and* $W_h^0, \overline{W}_h^0 \in V_h$. *Let* $\{W_h^m\}_{m=1}^N \subset V_h$ *be defined by* (4.1), $w_{\delta h} \in H^1(0,T;V_h)$ *be defined by*

$$(4.5)\quad w_{\delta h}(t) = W_h^{m-1} + \frac{t - t_{m-1}}{\delta}(W_h^m - W_h^{m-1}) \text{ on } [t_{m-1}, t_m],\ m = 1, 2, \cdots, N,$$

and $\overline{w}_h \in H^1(0,T;V_h)$ *be defined by*

$$\text{(4.6)}\qquad \begin{cases} \overline{w}_h(0) = \overline{W}_h^0\,, \\ \langle \partial_t \overline{w}_h(t), v_h\rangle + a[\overline{w}_h(t), v_h] = \langle \overline{g}(t), v_h\rangle \qquad \forall\, v_h \in V_h,\ a.e.\ t\,. \end{cases}$$

Then

$$\text{(4.7)}\qquad \begin{aligned} \|\overline{w}_h - w_{\delta h}\|^2_{L^2(0,T;H^1(\Omega))} &\le C \sum_{m=1}^N \int_{t_{m-1}}^{t_m} \|\overline{g}(t) - g^m\|^2_{-1}\,dt \\ &\quad + C\delta \sum_{m=1}^N \|W_h^m - W_h^{m-1}\|_1^2 + C\|\overline{W}_h^0 - W_h^0\|_0^2 \end{aligned}$$

and

$$\text{(4.8)}\qquad \begin{aligned} \|\partial_t\overline{w}_h - \partial_t w_{\delta h}\|^2_{L^2(0,T;H^{-1}(\Omega))} &\le C \sum_{m=1}^N \int_{t_{m-1}}^{t_m} \|\overline{g}(t) - g^m\|^2_{-1}\,dt \\ &\quad + C\delta \sum_{m=1}^N \|W_h^m - W_h^{m-1}\|_1^2 + C\|\overline{W}_h^0 - W_h^0\|_0^2. \end{aligned}$$

Proof. Subtracting (4.1) from (4.6) and noting that

$$\partial_t w_{\delta h} = \frac{W_h^m - W_h^{m-1}}{\delta} \quad \text{for } t \in [t_{m-1}, t_m]\,,$$

we obtain

$$\text{(4.9)}\qquad \begin{aligned} \langle \partial_t\overline{w}_h(t) - \partial_t w_{\delta h}(t), v_h\rangle + a[\overline{w}_h(t) - w_{\delta h}, v_h] + a[w_{\delta h} - W_h^m, v_h] \\ = \langle \overline{g}(t) - g^m, v_h\rangle \qquad \forall\, v_h \in V_h,\ \forall\, t \in [t_{m-1}, t_m]\,. \end{aligned}$$

Denoting $\xi_{\delta h} = \overline{w}_h - w_{\delta h}$ and using the relations

$$W_h^m = W_h^{m-1} + \frac{t_m - t_{m-1}}{\delta}(W_h^m - W_h^{m-1})$$

and

$$w_{\delta h} = W_h^{m-1} + \frac{t - t_{m-1}}{\delta}(W_h^m - W_h^{m-1}) \qquad \text{on } [t_{m-1}, t_m]$$

we may rewrite (4.9) as:

$$\text{(4.10)}\qquad \begin{aligned} \langle \partial_t \xi_{\delta h}(t), v_h\rangle + a[\xi_{\delta h}(t), v_h] = \langle \overline{g}(t) - g^m, v_h\rangle \\ + \frac{t_m - t}{\delta}\, a[W_h^m - W_h^{m-1}, v_h] \quad \forall\, v_h \in V_h,\ \forall\, t \in [t_{m-1}, t_m]. \end{aligned}$$

Setting $v_h = \xi_{\delta h}(t)$ we have, for $t \in [t_{m-1}, t_m]$,

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\|\xi_{\delta h}(t)\|_0^2 + C_a\|\xi_{\delta h}(t)\|_1^2 \\ &\quad \le \langle \partial_t \xi_{\delta h}(t), \xi_{\delta h}(t)\rangle + a[\xi_{\delta h}(t), \xi_{\delta h}(t)] = \langle \overline{g}(t) - g^m, \xi_{\delta h}(t)\rangle \\ &\quad \le \|\overline{g}(t) - g^m\|_{-1}\|\xi_{\delta h}(t)\|_1 + \frac{C(t_m - t)}{\delta}\|W_h^m - W_h^{m-1}\|_1\|\xi_{\delta h}(t)\|_1 \\ &\quad \le \frac{1}{C_a}\|\overline{g}(t) - g^m\|_{-1}^2 + \frac{C_a}{4}\|\xi_{\delta h}(t)\|_1^2 + C\|W_h^m - W_h^{m-1}\|_1^2 + \frac{C_a}{4}\|\xi_{\delta h}(t)\|_1^2 \end{aligned}$$

which may be simplified as

$$\frac{1}{2}\frac{d}{dt}\|\xi_{\delta h}(t)\|_0^2 + \frac{C_a}{2}\|\xi_{\delta h}(t)\|_1^2 \le \frac{1}{C_a}\|\overline{g}(t) - g^m\|_{-1}^2 + C\|W_h^m - W_h^{m-1}\|_1^2\,.$$

An integration with respect to t over $[t_{m-1}, t_m]$ leads us to

$$\|\xi_{\delta h}(t_m)\|_0^2 - \|\xi_{\delta h}(t_{m-1})\|_0^2 + C_a \int_{t_{m-1}}^{t_m} \|\xi_{\delta h}(t)\|_1^2 \, dt$$
$$\leq C \int_{t_{m-1}}^{t_m} \|\overline{g}(t) - g^m\|_{-1}^2 \, dt + C\delta \|W_h^m - W_h^{m-1}\|_1^2 \, .$$

Through summations in m over $m = 1, 2, \cdots, N$ and using the relation $\xi_{\delta h}(0) = \overline{W}_h^0 - W_h^0$ we obtain

$$\|\xi_{\delta h}(T)\|_0^2 + C_a \int_0^T \|\xi_{\delta h}(t)\|_1^2 \, dt$$
$$\leq C \sum_{m=1}^N \int_{t_{m-1}}^{t_m} \|\overline{g}(t) - g^m\|_{-1}^2 \, dt + C\delta \sum_{m=1}^N \|W_h^m - W_h^{m-1}\|_1^2 + \|\overline{W}_h^0 - W_h^0\|_0^0 \, ,$$

which implies (4.7).

Next we prove estimate (4.8). Since $\xi_{\delta h}(t) \in V_h$ a.e. on $[0, T]$, we have

$$\text{(4.11)} \qquad \langle \partial_t \xi_{\delta h}(t), v \rangle = \langle \partial_t e_{\delta h}(t), P_h v \rangle \qquad \forall\, v \in H_0^1(\Omega) \, , \text{ a.e. } t \, .$$

Using (4.11), (4.10) and (2.5) we obtain, for a.e. $t \in [t_{m-1}, t_m]$,

$$\langle \partial_t \xi_{\delta h}(t), v \rangle = \langle \partial_t \xi_{\delta h}(t), P_h v \rangle$$
$$= -a[\xi_{\delta h}(t), P_h v] + \langle \overline{g}(t) - g^m, P_h v \rangle + \frac{t_m - t}{\delta} a[W_h^m - W_h^{m-1}, P_h v]$$
$$\leq C \Big(\|\xi_{\delta h}(t)\|_1 + \|\overline{g}(t) - g^m\|_{-1} + \|W_h^m - W_h^{m-1}\|_1 \Big) \|v\|_1 \quad \forall\, v \in H_0^1(\Omega)$$

so that

$$\|\partial_t \xi_{\delta h}(t)\|_{-1} \leq C \Big(\|\xi_{\delta h}(t)\|_1 + \|\overline{g}(t) - g^m\|_{-1} + \|W_h^m - W_h^{m-1}\|_1 \Big) \, .$$

Hence,

$$\int_{t_{m-1}}^{t_m} \|\partial_t \xi_{\delta h}(t)\|_{-1}^2 \, dt$$
$$\leq C \Big(\int_{t_{m-1}}^{t_m} \|\xi_{\delta h}(t)\|_1^2 \, dt + \int_{t_{m-1}}^{t_m} \|\overline{g}(t) - g^m\|_{-1}^2 dt + \delta \|W_h^m - W_h^{m-1}\|_1^2 \Big) .$$

Through summations in m over $m = 1, 2, \cdots, N$ and applying (4.7) we have:

$$\int_0^T \|\partial_t \xi_{\delta h}(t)\|_{-1}^2 \, dt$$
$$\leq C \Big(\int_0^T \|\xi_{\delta h}(t)\|_1^2 \, dt + \sum_{m=1}^N \int_{t_{m-1}}^{t_m} \|\overline{g}(t) - g^m\|_{-1}^2 \, dt + \delta \sum_{m=1}^N \|W_h^m - W_h^{m-1}\|_1^2 \Big)$$
$$\leq C \Big(\sum_{m=1}^N \int_{t_{m-1}}^{t_m} \|\overline{g}(t) - g^m\|_{-1}^2 \, dt + \delta \sum_{m=1}^N \|W_h^m - W_h^{m-1}\|_1^2 + \|\overline{W}_h^0 - W_h^0\|_0^2 \Big) \, .$$

This proves (4.8). □

We now derive an $O(\delta)$ error estimate under the regularity hypotheses which are precisely those that ensure the standard regularity for the solution of the linear parabolic PDE – see [11, p.360-361,Theorem 5]. Compared to $O(\delta)$ error estimates in the literature such as those in [16], our $O(\delta)$ error estimate is derived in the norm for the solution space $\mathcal{W}(0, T)$ (instead of the typical discrete $L^\infty(0, T; L^2(\Omega))$ norm) and an approach of proof is given that is different than those in the literature.

Theorem 4.3. *Assume that* $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$ *and* $f \in H^1(0,T;L^2(\Omega))$. *Let* $u_h \in H^1(0,T;V_h)$ *be the solution of* (3.1) *and let* $u_{\delta h}$ *be defined by* (3.5)–(3.6). *Then*

$$\|u_h - u_{\delta h}\|_{\mathcal{W}(0,T)} \le C\delta\Big(\|f\|_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2\Big). \tag{4.12}$$

Proof. Applying Lemma 4.2 with $g \equiv f$, $\overline{g} \equiv f$, $W_h^0 \equiv U_h^0$, $\overline{W}_h^0 \equiv U_h^0$, $\overline{w}_h \equiv u_h$, $W_h^m \equiv U_h^m$ and $w_{\delta h} \equiv u_{\delta h}$ we obtain:

$$\begin{aligned} &\|u_h - u_{\delta h}\|^2_{\mathcal{W}(0,T)} \\ &\quad \le C\sum_{m=1}^N \int_{t_{m-1}}^{t_m} \|f(t) - f^m\|_{-1}^2\,dt + C\delta\sum_{m=1}^N \|U_h^m - U_h^{m-1}\|_1^2 \,. \end{aligned} \tag{4.13}$$

From (4.13) and Lemma 3.2 we obtain

$$\begin{aligned} \|u_h - u_{\delta h}\|^2_{\mathcal{W}(0,T)} &\le C\sum_{m=1}^N \int_{t_{m-1}}^{t_m} \|f(t) - f^m\|_{-1}^2\,dt + C\delta\sum_{m=1}^N \|U_h^m - U_h^{m-1}\|_1^2 \\ &\le C\delta^2\|f\|^2_{H^1(0,T;H^{-1}(\Omega))} + C\delta\sum_{m=1}^N \|U_h^m - U_h^{m-1}\|_1^2 \,. \end{aligned}$$

Thus (4.12) is proved if we can justify that

$$\delta\sum_{m=1}^N \|U_h^m - U_h^{m-1}\|_1^2 \le C\delta^2\Big(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2^2\Big)\,. \tag{4.14}$$

Subtracting consecutive equations in (3.5) and denoting $Y_h^m = U_h^m - U_h^{m-1}$ for $m = 1, 2, \cdots, N$ we obtain

$$\begin{aligned} &\Big\langle \frac{Y_h^m - Y_h^{m-1}}{\delta}, v_h \Big\rangle + a[Y_h^m, v_h] \\ &\qquad\qquad = \langle f^m - f^{m-1}, v_h\rangle \qquad \forall\ v_h \in V_h,\ 2 \le m \le N\,. \end{aligned}$$

Setting $v_h = Y_h^m$ we have, for $m = 2, \cdots, N$,

$$\begin{aligned} &\frac{1}{2\delta}\Big(\|Y_h^m\|_0^2 - \|Y_h^{m-1}\|_0^2 + \|Y_h^m - Y_h^{m-1}\|_0^2\Big) + a[Y_h^m, Y_h^m] \\ &\qquad = \langle f^m - f^{m-1}, Y_h^m\rangle \le C\|f^m - f^{m-1}\|_{-1}^2 + \frac{C_a}{2}\|Y_h^m\|_1^2 \end{aligned}$$

so that

$$\frac{1}{2\delta}\Big(\|Y_h^m\|_0^2 - \|Y_h^{m-1}\|_0^2\Big) + \frac{C_a}{2}\|Y_h^m\|_1^2 \le C\|f^m - f^{m-1}\|_{-1}^2\,.$$

Summations for $m = 2, \cdots, N$ yield

$$\frac{1}{2\delta}\Big(\|Y_h^N\|_0^2 - \|Y_h^1\|_0^2\Big) + \frac{C_a}{2}\sum_{m=2}^N \|Y_h^m\|_1^2 \le C\sum_{m=2}^N \|f^m - f^{m-1}\|_{-1}^2$$

so that

$$\delta\sum_{m=1}^N \|Y_h^m\|_1^2 \le C\delta\sum_{m=2}^N \|f^m - f^{m-1}\|_{-1}^2 + \delta\|Y_h^1\|_1^2 + C\|Y_h^1\|_0^2\,. \tag{4.15}$$

We proceed to estimate the three terms on the right hand side of (4.15).

The term $\delta \sum_{m=2}^{N} \|f^m - f^{m-1}\|_{-1}^2$ in (4.15) is estimated as follows. For each $m = 2, \cdots, N$,

$$\begin{aligned}\|f^m - f^{m-1}\|_{-1}^2 &= \frac{1}{\delta^2}\Big\| \int_{t_{m-1}}^{t_m} f(t)\,dt - \int_{t_{m-2}}^{t_{m-1}} f(t)\,dt\Big\|_{-1}^2 \\ &= \frac{1}{\delta^2}\Big\| \int_{t_{m-2}}^{t_{m-1}} (f(t+\delta) - f(t))\,dt\Big\|_{-1}^2 = \frac{1}{\delta^2}\Big\| \int_{t_{m-2}}^{t_{m-1}} \int_{t}^{t+\delta} f'(s)\,ds\,dt\Big\|_{-1}^2 \\ &\le \frac{1}{\delta^2}\Big(\int_{t_{m-2}}^{t_{m-1}} \int_{t_{m-2}}^{t_m} \|f'(s)\|_{-1}\,ds\,dt\Big)^2 \le C\delta \int_{t_{m-2}}^{t_m} \|f'(s)\|_{-1}^2\,ds\,.\end{aligned}$$

Thus,

$$\begin{aligned}\delta \sum_{m=2}^{N} \|f^m - f^{m-1}\|_{-1}^2 &\le C\delta^2 \sum_{m=2}^{N} \int_{t_{m-2}}^{t_m} \|f'(s)\|_{-1}^2\,ds \\ &\le C\delta^2 \int_0^T \|f'(s)\|_{-1}^2\,ds\,.\end{aligned} \tag{4.16}$$

The estimation of the terms $\delta\|Y_h^1\|_1^2$ and $\|Y_h^1\|_0^2$ in (4.15) is carried out in the following steps: derivation of a priori estimates for $u_h'(t)$; estimation of local errors $\|u_h(\delta) - U_h^1\|_0^2 + \delta\|u_h(\delta) - U_h^1\|_0^2$; and completion of proof by triangle inequalities.

First, we derive a priori estimates for $u_h'(t)$ from (3.1). Note that the regularity assumptions on f ensures that $f \in C([0,T]; H^{-1}(\Omega))$ so that the forcing terms in (3.1) are continuous; then standard ODE theories imply that the solution $u_h(t)$ of (3.1) is C^1 and that Equation (3.1) holds pointwise in t. Setting $t = 0$ in (3.1) and then choosing $v_h = u_h'(0)$ we have

$$\begin{aligned}\|u_h'(0)\|_0^2 &= [f(0), u_h'(0)] - a[u_0, u_h'(0)] + a[u_0 - U_h^0, u_h'(0)] \\ &\le \|f(0)\|_0 \|u_h'(0)\|_0 + [\operatorname{div}(A(\mathbf{x})\nabla u_0), u_h'(0)] + C\|u_0 - U_h^0\|_1 \|u_h'(0)\|_1 \\ &\le C\|f(0)\|_0 \|u_h'(0)\|_0 + C\|u_0\|_2 \|u_h'(0)\|_0 + Ch\|u_0\|_2 \frac{C}{h} \|u_h'(0)\|_0\end{aligned}$$

so that

$$\|u_h'(0)\|_0 \le C\Big(\|f(0)\|_0 + \|u_0\|_2\Big). \tag{4.17}$$

Differentiating (3.1) with respect to t and then setting $v_h = u_h'(t)$ we obtain:

$$\frac{1}{2}\frac{d}{dt}\|u_h'(t)\|_0^2 + a[u_h'(t), u_h'(t)] = [f'(t), u_h'(t)] \le C\|f'(t)\|_{-1}^2 + \frac{C_a}{2}\|u_h'(t)\|_1^2$$

so that

$$\frac{1}{2}\frac{d}{dt}\|u_h'(t)\|_0^2 + \frac{C_a}{2}\|u_h'(t)\|_1^2 \le C\|f'(t)\|_{-1}^2\,.$$

The last estimate and (4.17) readily yield

$$\begin{aligned}\|u_h'(s)\|_0^2 + \int_0^\delta \|u_h'(t)\|_1^2\,dt &\le C\int_0^\delta \|f'(t)\|_{-1}^2\,dt + C\|u_h'(0)\|_0^2 \\ &\le C\int_0^\delta \|f'(t)\|_{-1}^2\,dt + C\Big(\|f(0)\|_0^2 + \|u_0\|_2^2\Big) \\ &\le C\Big(\|f\|_{H^1(0,T;L^2(\Omega))}^2 + \|u_0\|_2^2\Big) \qquad \forall\, s \in [0,\delta].\end{aligned} \tag{4.18}$$

Here we also used the continuous embedding $H^1(0,T;L^2(\Omega)) \hookrightarrow C([0,T];L^2(\Omega))$ to estimate $\|f(0)\|_0$.

Next, we estimate the local error $\|u_h(\delta) - U_h^1\|_0^2 + \delta\|u_h(\delta) - U_h^1\|_0^2$. Integrating (3.1) from $t_0 = 0$ to $t_1 = \delta$ we obtain

$$\langle u_h(\delta) - U_h^0, v_h\rangle + a[\textstyle\int_0^\delta u_h(t)\,dt, v_h] = \delta\,\langle f^1, v_h\rangle \qquad \forall\, v_h \in V_h.$$

From (3.5) with $m = 1$ we deduce

$$\langle U_h^1 - U_h^0, v_h\rangle + \delta\, a[U_h^1, v_h] = \delta\,\langle f^1, v_h\rangle \qquad \forall\, v_h \in V_h\,.$$

Subtracting the last two relations we obtain:

$$\begin{aligned}\langle u_h(\delta) - U_h^1, v_h\rangle &+ \delta a[u_h(\delta) - U_h^1, v_h]\\ &+ a[\textstyle\int_0^\delta u_h(t)\,dt - \delta u_h(\delta), v_h] = 0 \qquad \forall\, v_h \in V_h\,.\end{aligned}$$

Setting $v_h = u_h(\delta) - U_h^1$ we are led to:

$$\begin{aligned}\|u_h(\delta) - U_h^1\|_0^2 &+ \delta\, a[u_h(\delta) - U_h^1, u_h(\delta) - U_h^1]\\ &= -a[\textstyle\int_0^\delta u_h(t)\,dt - \delta u_h(\delta), u_h(\delta) - U_h^1]\\ &\le \frac{C}{\delta}\Big\|\int_0^\delta u_h(t)\,dt - \delta u_h(\delta)\Big\|_1^2 + \frac{C_a\delta}{2}\|u_h(\delta) - U_h^1\|_1^2\end{aligned}$$

so that

$$\begin{aligned}&\|u_h(\delta) - U_h^1\|_0^2 + \frac{C_a\delta}{2}\|u_h(\delta) - U_h^1\|_1^2\\ &\quad\le \frac{C}{\delta}\Big\|\int_0^\delta \big(u_h(t) - u_h(\delta)\big)\,dt\Big\|_1^2 = \frac{C}{\delta}\Big\|\int_0^\delta\int_\delta^t u_h'(s)\,ds\,dt\Big\|_1^2\\ &\quad\le \frac{C}{\delta}\Big(\int_0^\delta\int_0^\delta \|u_h'(s)\|_1\,ds\,dt\Big)^2 \le C\delta\Big(\int_0^\delta \|u_h'(s)\|_1\,ds\Big)^2 \le C\delta^2\int_0^\delta \|u_h'(s)\|_1^2\,ds\,.\end{aligned}$$

By virtue of (4.18) the last relation reduces to

$$\|u_h(\delta) - U_h^1\|_0^2 + \delta\,\|u_h(\delta) - U_h^1\|_1^2 \le C\delta^2\Big(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2^2\Big)\,. \tag{4.19}$$

Now, we use the triangle inequality and (4.18)–(4.19) to estimate $\delta\|Y_h^1\|_1^2$ and $\|Y_h^1\|_0^2$:

$$\begin{aligned}\|Y_h^1\|_1^2 = \|U_h^1 - U_h^0\|_1^2 &\le 2\|U_h^1 - u_h(\delta)\|_1^2 + 2\|u_h(\delta) - u_h(0)\|_1^2\\ &\le C\delta\Big(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2^2\Big) + C\Big\|\int_0^\delta u_h'(s)\,ds\Big\|_1^2\\ &\le C\delta\Big(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2^2\Big) + C\delta\int_0^\delta \|u_h'(s)\|_1^2\,ds\\ &\le C\delta\Big(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2^2\Big)\end{aligned}$$

so that

$$\delta\|Y_h^1\|_1^2 \le C\delta^2\Big(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2^2\Big)\,; \tag{4.20}$$

similarly,

$$
\begin{aligned}
\|Y_h^1\|_0^2 &= \|U_h^1 - U_h^0\|_0^2 \le 2\|U_h^1 - u_h(\delta)\|_0^2 + 2\|u_h(\delta) - u_h(0)\|_0^2 \\
&\le 2\|U_h^1 - u_h(\delta)\|_0^2 + 2\Big\| \int_0^\delta u_h'(s)\,ds \Big\|_0^2 \\
&\le C\delta^2\Big(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2^2\Big) + C\delta \int_0^\delta \|u_h'(s)\|_0^2\,ds \\
&\le C\delta^2\Big(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2^2\Big) + C\delta^2\Big(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2^2\Big).
\end{aligned}
\tag{4.21}
$$

Substituting estimates (4.16), (4.20) and (4.21) into (4.15) we obtain

$$
\begin{aligned}
\delta \sum_{m=1}^N \|Y_h^m\|_1^2 &\le C\delta^2 \|f'\|^2_{H^1(0,T;L^2(\Omega))} + C\delta^2\Big(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2^2\Big) \\
&\le C\delta^2\Big(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2^2\Big),
\end{aligned}
$$

which is precisely (4.14). ☐

Corollary 4.4. *Assume that* $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$ *and* $f \in L^2(0,T;H_0^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))$. *Let* $u_h \in H^1(0,T;V_h)$ *be the solution of* (3.1) *and let* $u_{\delta h}$ *be defined by* (3.5)–(3.6). *Then*

$$
\begin{aligned}
&\|u_h - u_{\delta h}\|_{\mathcal{W}(0,T)} \\
&\qquad \le C\delta\Big(\|f\|_{L^2(0,T;H^1(\Omega))} + \|f\|_{H^1(0,T;H^{-1}(\Omega))} + \|u_0\|_2\Big).
\end{aligned}
\tag{4.22}
$$

Proof. A careful review of the proof of Theorem 4.3 reveals that the $H^1(0,T;L^2(\Omega))$ assumption on f may be weakened to: $f \in H^1(0,T;H^{-1}(\Omega))$ and $f(0) \in L^2(\Omega)$ (see also [18, p.202, Theorem 3.5].) Since $f \in L^2(0,T;H_0^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega)) \hookrightarrow C([0,T];L^2(\Omega))$, we have $f(0) \in L^2(\Omega)$. Hence, the conclusions of Theorem 4.3 hold. ☐

Remark 4.5. In Corollary 4.4 we assumed that f has a homogeneous boundary condition. This is reasonable as the differential equation $\partial_t u - \operatorname{div}[A(\mathbf{x})\nabla u] = f$ holds in $H^1(\Omega)$ for a.e. t and u has a homogeneous boundary condition.

Remark 4.6. Recall that the basic regularity for f is $L^2(0,T;H^{-1}(\Omega))$. It seems from Theorem 4.3 and Corollary 4.4 that raising the regularity of f by one integer order in either t or $\mathbf{x}$ is not enough to guarantee $O(\delta)$ error estimate. Theorem 4.3 assumed one order regularity in t and one order regularity in $\mathbf{x}$, while Corollary 4.4 assumed two orders of regularity in $\mathbf{x}$.

The regularity assumptions on f in Theorem 4.3 is mixed in both t and $\mathbf{x}$ since $H^1(0,T;L^2(\Omega))$ is more regular in both t and $\mathbf{x}$ than $L^2(0,T;H^{-1}(\Omega))$ which is the basic regularity of f. Such mixed regularity assumptions are harder to verify than uni-directional regularity assumptions such as $H^\alpha(0,T;H^{-1}(\Omega))$ and $L^2(0,T;H^\sigma(\Omega))$ for some $\alpha > 0$ and $\sigma > -1$. In the next theorem we establish estimates for $\|u_h - u_{\delta h}\|_{\mathcal{W}(0,T)}$ under suitable uni-directional regularity assumptions. We will use the $\Phi_0^\gamma(\Omega)$ notation introduced in Section 2.3.

Theorem 4.7. *Let* $u_h \in H^1(0,T;V_h)$ *be the solution of* (3.1) *and let* $u_{\delta h}$ *be defined by* (3.5).

i) *If for some* $\gamma \in [0,1]$,

$$u_0 \in H^{1+\gamma}(\Omega) \cap H_0^1(\Omega), \quad f \in L^2(0,T;H^{-1+2\gamma}(\Omega) \cap \Phi_0^{-1+2\gamma}(\Omega)) \text{ and } f \in H^\gamma(0,T;H^{-1}(\Omega)), \tag{4.23}$$

then

$$\|u_h - u_{\delta h}\|_{\mathcal{W}(0,T)} \leq C\delta^\gamma \Big(\|f\|_{L^2(0,T;H^{-1+2\gamma}(\Omega))} + \|f\|_{H^\gamma(0,T;H^{-1}(\Omega))} + \|u_0\|_{1+\gamma} \Big). \tag{4.24}$$

ii) *If* $\delta \leq Ch$, *and for some* $\gamma \in [0,1]$,

$$u_0 \in H^{(1+3\gamma)/2}(\Omega) \cap \Phi_0^{(1+3\gamma)/2}(\Omega), \quad f \in L^2(0,T;H^{-1+2\gamma}(\Omega) \cap \Phi_0^{-1+2\gamma}(\Omega)) \text{ and } f \in H^\gamma(0,T;H^{-1}(\Omega)), \tag{4.25}$$

then

$$\|u_h - u_{\delta h}\|_{\mathcal{W}(0,T)} \leq C\delta^\gamma \Big(\|f\|_{L^2(0,T;H^{-1+2\gamma}(\Omega))} + \|f\|_{H^\gamma(0,T;H^{-1}(\Omega))} + \|u_0\|_{(1+3\gamma)/2} \Big). \tag{4.26}$$

iii) *If* $\delta \leq Ch^2$, *and for some* $\gamma \in [0,1]$,

$$u_0 \in H^{2\gamma}(\Omega) \cap \Phi_0^{2\gamma}(\Omega), \quad f \in L^2(0,T;H^{-1+2\gamma}(\Omega) \cap \Phi_0^{-1+2\gamma}(\Omega)) \text{ and } f \in H^\gamma(0,T;H^{-1}(\Omega)), \tag{4.27}$$

then

$$\|u_h - u_{\delta h}\|_{\mathcal{W}(0,T)} \leq C\delta^\gamma \Big(\|f\|_{L^2(0,T;H^{-1+2\gamma}(\Omega))} + \|f\|_{H^\gamma(0,T;H^{-1}(\Omega))} + \|u_0\|_{2\gamma} \Big). \tag{4.28}$$

Proof. We will furnish a complete proof for case i). For cases ii) and iii) we will merely indicate the changes to be made to the proof of case i).

Case i) Thanks to interpolation theorems (see [15],) it suffices to prove (4.24) for the cases $\gamma = 0$ and $\gamma = 1$.

For $\gamma = 0$, regularity assumption (4.23) reduces to $u_0 \in H_0^1(\Omega)$ and $f \in L^2(0,T;H^{-1}(\Omega))$. Applying Lemma 4.2 with $g \equiv f$, $\overline{g} \equiv f$, $W_h^0 \equiv U_h^0$, $\overline{W}_h^0 \equiv U_h^0$, $\overline{w}_h \equiv u_h$, $W_h^m \equiv U_h^m$ and $w_{\delta h} \equiv u_{\delta h}$ we obtain:

$$\|u_h - u_{\delta h}\|^2_{\mathcal{W}(0,T)} \leq C \sum_{m=1}^N \int_{t_{m-1}}^{t_m} \|f(t) - f^m\|^2_{-1}\, dt + C\delta \sum_{m=1}^N \|U_h^m - U_h^{m-1}\|_1^2 . \tag{4.29}$$

Applying Lemma 4.1 with $W_h^m \equiv U_h^m$, $W_h^0 \equiv U_h^0$ and $g \equiv f$, also noting that $\|U_h^0\|_1 = \|P_h u_0\|_1 \leq C\|u_0\|_1$ we obtain:

$$\begin{aligned} \sum_{m=1}^N \|U_h^m - U_h^{m-1}\|_1^2 &\leq \frac{C}{\delta} \|f\|^2_{L^2(0,T;H^{-1}(\Omega))} + C\|U_h^0\|_1^2 \\ &\leq \frac{C}{\delta} \|f\|^2_{L^2(0,T;H^{-1}(\Omega))} + C\|u_0\|_1^2 \qquad \text{if } f \in L^2(0,T;H^{-1}(\Omega)) . \end{aligned} \tag{4.30}$$

Combining (4.29), (4.30), and (3.9) with $\gamma = 0$ we deduce that

$$\|u_h - u_{\delta h}\|^2_{\mathcal{W}(0,T)} \leq C\|f\|^2_{L^2(0,T;H^{-1}(\Omega))} + C\delta\|u_0\|_1^2 . \tag{4.31}$$

For $\gamma = 1$, regularity assumption (4.23) is equivalent to

$$u_0 \in H^2(\Omega) \cap H_0^1(\Omega) \text{ and } f \in L^2(0,T;H_0^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega)) . \tag{4.32}$$

Thus by Corollary 4.4, estimate (4.22) holds.

Interpolations between (4.31) and (4.22) yield (4.24).

Case ii) For $\gamma = 0$, regularity assumption (4.25) reduces to $u_0 \in H^{1/2}(\Omega)$ and $f \in L^2(0,T;H^{-1}(\Omega))$. Using the relation $\delta \le Ch$ and the inverse inequality $\|U_h^0\|_1 \le Ch^{-1/2}\|U_h^0\|_{1/2}$ in (4.30) we obtain from (4.29) and (4.30) that

$$\|u_h - u_{\delta h}\|^2_{\mathcal{W}(0,T)} \le C\Big(\|f\|^2_{L^2(0,T;H^{-1}(\Omega))} + \|u_0\|^2_{1/2}\Big)\,. \tag{4.33}$$

The rest of the proof is similar to that of Case i).

Case iii) For $\gamma = 0$, regularity assumption (4.27) reduces to $u_0 \in L^2(\Omega)$ and $f \in L^2(0,T;H^{-1}(\Omega))$. Using the relation $\delta \le Ch^2$ and the inverse inequality $\|U_h^0\|_1 \le Ch^{-1}\|U_h^0\|_0$ in (4.30) we obtain from (4.29) and (4.30) that

$$\|u_h - u_{\delta h}\|^2_{\mathcal{W}(0,T)} \le C\Big(\|f\|^2_{L^2(0,T;H^{-1}(\Omega))} + \|u_0\|^2_0\Big)\,. \tag{4.34}$$

The rest of the proof is similar to that of Case i). □

Remark 4.8. Using the definition of $\Phi_0^r(\Omega)$ we see that for suitable values of γ, Theorem 4.7 may be stated without the need of the $\Phi_0^r(\Omega)$ notation. For instance, case i) can be stated as follows: if for some $\gamma \in [0, 3/4]$,

$$u_0 \in H^{1+\gamma}(\Omega) \cap H_0^1(\Omega),$$
$$f \in L^2(0,T;H^{-1+2\gamma}(\Omega)) \text{ and } f \in H^\gamma(0,T;H^{-1}(\Omega))\,,$$

then

$$\|u_h - u_{\delta h}\|_{\mathcal{W}(0,T)} \le C\delta^\gamma\Big(\|f\|_{L^2(0,T;H^{-1+2\gamma}(\Omega))} + \|f\|_{H^\gamma(0,T;H^{-1}(\Omega))} + \|u_0\|_{1+\gamma}\Big).$$

5. Convergence of and error estimates for fully discrete approximations

In this section we first prove that $\|u_h - u_{\delta h}\|_{\mathcal{W}(0,T)} \to 0$ as $h, \delta \to 0$; the proof will be based on the denseness of smooth functions in $L^2(0,T;H^{-1}(\Omega))$ and the error estimates of Theorem 4.7. We then prove $\|u - u_{\delta h}\|_{\mathcal{W}(0,T)} \to 0$ as $h, \delta \to 0$ and derive fully discrete error estimates.

Theorem 5.1. *Assume that $f \in L^2(0,T;H^{-1}(\Omega))$ and $u_0 \in L^2(\Omega)$. Let $u_h \in H^1(0,T;V_h)$ be the solution of* (3.1) *and $u_{\delta h}$ be defined by* (3.5)–(3.6). *If* i) $u_0 \in H_0^1(\Omega)$, *or* ii) $u_0 \in H^{1/2}(\Omega)$ *and $\delta \le Ch$, or* iii) $u_0 \in L^2(\Omega)$ *and $\delta \le Ch^2$, then*

$$\|u_h - u_{\delta h}\|_{\mathcal{W}(0,T)} \to 0 \quad \text{as } h, \delta \to 0\,. \tag{5.1}$$

Proof. Using the denseness of $C_0^\infty([0,T]\times\overline{\Omega})$ in $L^2(0,T;L^2(\Omega)) \subset L^2(0,T;H^{-1}(\Omega))$ we may choose a family of functions $\{f^\epsilon\} \subset C_0^\infty([0,T]\times\overline{\Omega}) \subset L^2(0,T;L^2(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))$ such that $\|f^\epsilon - f\|_{L^2(0,T;H^{-1}(\Omega))} \to 0$ as $\epsilon \to 0$. Likewise, we may choose a family of functions $u_0^\epsilon \subset H_0^1(\Omega)$ such that $\|u_0^\epsilon - u_0\|_1 \to 0$ as $\epsilon \to 0$ in Case i), or $\|u_0^\epsilon - u_0\|_{1/2} \to 0$ as $\epsilon \to 0$ in Case ii), or $\|u_0^\epsilon - u_0\|_0 \to 0$ as $\epsilon \to 0$ in Case iii).

We use the $\widetilde{\ }$ notation to denote the ϵ dependence, e.g., $\widetilde{f}$ denotes f^ϵ, $\widetilde{u}_0$ denotes u_0^ϵ, and $\widetilde{u}_h$ Denotes $\widetilde{u}_h^\epsilon$. Let $\widetilde{u}_h \in H^1(0,T;V_h)$ be the solution of

$$\begin{cases} \big\langle \partial_t \widetilde{u}_h(t), v\big\rangle + a[\widetilde{u}_h(t), v_h] = \big\langle \widetilde{f}(t), v_h\big\rangle \qquad \forall\, v_h \in V_h,\ \text{a.e. } t \in [0,T]\,, \\ \widetilde{u}_h(0) = P_h \widetilde{u}\,. \end{cases} \tag{5.2}$$

We define $\widetilde{U}_h^m \in V_h$ and $\widetilde{u}_{\delta h} \in H^1(0,T;V_h)$ as follows:

$$(5.3)\qquad \begin{cases} \widetilde{U}_h^0 = P_h \widetilde{u}_0; \\ \Big\langle \dfrac{\widetilde{U}_h^m - \widetilde{U}_h^{m-1}}{\delta}, v_h \Big\rangle + a[\widetilde{U}_h^m, v_h] = \langle \widetilde{f}^m, v_h\rangle \qquad \forall\, v_h \in V_h,\ 1 \le m \le N \end{cases}$$

and

$$(5.4)\qquad \widetilde{u}_{\delta h}(t)\big|_{[t_{m-1},t_m]} = \widetilde{U}_h^{m-1} + \frac{t-t_{m-1}}{\delta}(\widetilde{U}_h^m - \widetilde{U}_h^{m-1}), \qquad 1 \le m \le N\,,$$

where in (5.3), $\widetilde{f}^m = \delta^{-1}\int_{t_{m-1}}^{t_m} \widetilde{f}(s)\,ds$.

Using triangle inequalities and denoting $\theta_h = u_h - \widetilde{u}_h$ and $\theta_{\delta h} = u_{\delta h} - \widetilde{u}_{\delta h}$ we deduce

$$(5.5)\qquad \|u_h - u_{\delta h}\|_{\mathcal{W}(0,T)} \le \|\theta_h\|_{\mathcal{W}(0,T)} + \|\widetilde{u}_h - \widetilde{u}_{\delta h}\|_{\mathcal{W}(0,T)} + \|\theta_{\delta h}\|_{\mathcal{W}(0,T)}\,.$$

We need to estimate the three terms on the right hand side of (5.5).

We first estimate the term $\|\widetilde{u}_h - \widetilde{u}_{\delta h}\|_{\mathcal{W}(0,T)}$. Applying Theorem 4.7 Case i) with $\gamma = 1/2$ we have

$$(5.6)\qquad \begin{aligned} &\|\widetilde{u}_h - \widetilde{u}_{\delta h}\|_{\mathcal{W}(0,T)} \\ &\quad \le C\delta^{1/2}\Big(\|\widetilde{f}\|_{L^2(0,T;L^2(\Omega))} + \|\widetilde{f}\|_{H^{1/2}(0,T;H^{-1}(\Omega))} + \|u_0\|_1\Big)\,. \end{aligned}$$

Next, we estimate $\|\theta_h\|_{\mathcal{W}(0,T)} \equiv \|u_h - \widetilde{u}_h\|_{\mathcal{W}(0,T)}$. Subtracting (5.2) from (3.1) we obtain

$$(5.7)\qquad \begin{cases} \big\langle \partial_t\theta_h(t), v_h\big\rangle + a[\theta_h(t), v_h] \\ \qquad\qquad = \big\langle f(t) - \widetilde{f}(t), v_h\big\rangle \quad \forall\, v_h \in V_h,\ \text{a.e. } t \in [0,T], \\ \theta_h(0) = u_h(0) - U_h^0 = P_h u_0 - P_h\widetilde{u}_0\,. \end{cases}$$

Setting $v_h = \theta_h(t)$ we have that

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|\theta_h(t)\|_0^2 + a[\theta_h(t),\theta_h(t)] &= \big\langle f(t) - \widetilde{f}(t), \theta_h(t)\big\rangle \\ &\le C\|f(t) - \widetilde{f}(t)\|_{-1}^2 + \frac{C_a}{2}\|\theta_h(t)\|_1^2 \end{aligned}$$

so that

$$\frac{1}{2}\frac{d}{dt}\|\theta_h(t)\|_0^2 + \frac{C_a}{2}\|\theta_h(t)\|_1^2 \le C\|f(t) - \widetilde{f}(t)\|_{-1}^2\,.$$

An integration in t together with the relation $\theta_h(0) = P_h u_0 - P_h\widetilde{u}_h$ yields

$$(5.8)\qquad \begin{aligned} \|\theta_h\|_{L^2(0,T;H^1(\Omega))}^2 &\le C\Big(\|f - \widetilde{f}\|_{L^2(0,T;H^{-1}(\Omega))}^2 + \|P_h(u_0 - \widetilde{u}_0)\|_0^2\Big) \\ &\le C\Big(\|f - \widetilde{f}\|_{L^2(0,T;H^{-1}(\Omega))}^2 + \|u_0 - \widetilde{u}_0\|_0^2\Big)\,. \end{aligned}$$

We estimate $\|\partial_t\theta_h\|_{L^2(0,T;H^1(\Omega))}^2$ as follows. Let $v \in H_0^1(\Omega)$ be arbitrarily given. Since $\theta_h(t) \in V_h$ for almost every t, the definition of the projection $P_h : L^2(\Omega) \to V_h$ implies $\big\langle \partial_t\theta_h(t), v\big\rangle = \big\langle \partial_t\theta_h(t), P_h v\big\rangle$ a.e. t; thus, using (5.7) we deduce

$$\begin{aligned} &\big\langle \partial_t\theta_h(t), v\big\rangle = \big\langle \partial_t\theta_h(t), P_h v\big\rangle = -a[\theta_h(t), P_h v] + \big\langle f(t) - \widetilde{f}(t), P_h v\big\rangle \\ &\le C\Big(\|\theta_h(t)\|_1 + \|f(t) - \widetilde{f}(t)\|_{-1}\Big)\|P_h v\|_1 \le C\Big(\|\theta_h(t)\|_1 + \|f(t) - \widetilde{f}(t)\|_{-1}\Big)\|v\|_1\,. \end{aligned}$$

By taking the supremum over $v \in H_0^1(\Omega)$ with $\|v\|_1 \le 1$ we are led to

$$\|\partial_t\theta_h(t)\|_{-1}^2 \le C\Big(\|\theta_h(t)\|_1^2 + \|f(t) - \widetilde{f}(t)\|_{-1}^2\Big)\,.$$

Integrating in t and utilizing (5.8) we arrive at

$$\|\partial_t \theta_h\|^2_{L^2(0,T;H^{-1}(\Omega))} \le C\Big(\|f-\widetilde{f}\|^2_{L^2(0,T;H^{-1}(\Omega))} + \|u_0 - \widetilde{u}_0\|_0^2\Big)\,. \tag{5.9}$$

The term $\|\theta_{\delta h}\|_{\mathcal{W}(0,T)} \equiv \|u_{\delta h} - \widetilde{u}_{\delta h}\|_{\mathcal{W}(0,T)}$ may be estimated as follows. Subtracting (5.3)–(5.4) from the corresponding member of (3.5)–(3.6) and denoting $\Theta_h^m = U_h^m - \widetilde{U}_h^m$ we obtain

$$\begin{cases} \Theta_h^0 = U_h^0 - \widetilde{U}_h^0 = P_h(u_0 - \widetilde{u}_0); \\ \Big\langle \dfrac{\Theta_h^m - \Theta_h^{m-1}}{\delta}, v_h\Big\rangle + a[\Theta_h^m, v_h] \\ \qquad\qquad = \langle f^m - \widetilde{f}^m, v_h\rangle \quad \forall\, v_h \in V_h,\ 1 \le m \le N \end{cases} \tag{5.10}$$

and

$$\theta_{\delta h}(t)\big|_{[t_{m-1},t_m]} = \Theta_h^{m-1} + \frac{t - t_{m-1}}{\delta}(\Theta_h^m - \Theta_h^{m-1}), \qquad 1 \le m \le N. \tag{5.11}$$

Applying Lemma 4.1 to (5.10)–(5.11) we have

$$\begin{aligned} &\delta \sum_{m=1}^N \|\Theta_h^m\|_1^2 + \delta \sum_{m=1}^N \|\Theta_h^m - \Theta_h^{m-1}\|_1^2 \\ &\qquad \le C\|f - \widetilde{f}\|^2_{L^2(0,T;H^{-1}(\Omega))} + C\delta\|P_h(u_0 - \widetilde{u}_0)\|_1^2 + C\|P_h(u_0-\widetilde{u}_0)\|_0^2\,. \end{aligned} \tag{5.12}$$

Thus,

$$\begin{aligned} \|\theta_{\delta h}\|^2_{L^(0,T;H^1(\Omega))} &= \sum_{m=1}^N \int_{t_{m-1}}^{t_m} \|\theta_{\delta h}(t)\|_1^2\,dt \\ &\le 2\sum_{m=1}^N \int_{t_{m-1}}^{t_m} \Big(\|\Theta_h^{m-1}\|_1^2 + \|\Theta_h^m - \Theta_h^{m-1}\|_1^2\Big)\,dt \\ &= 2\delta \sum_{m=1}^N \|\Theta_h^{m-1}\|_1^2 + 2\delta\sum_{m=1}^N \|\Theta_h^m - \Theta_h^{m-1}\|_1^2 \\ &\le C\|f-\widetilde{f}\|^2_{L^2(0,T;H^{-1}(\Omega))} + C\delta\|P_h(u_0-\widetilde{u}_0)\|_1^2 + C\|P_h(u_0-\widetilde{u}_0)\|_0^2\,. \end{aligned}$$

To estimate $\|\partial_t\theta_{\delta h}\|_{L^2(0,T;H^{-1}(\Omega))}$ we note that using (5.11), the projection property of P_h and (5.10) we obtain, for a.e. $t \in [t_{m-1}, t_m]$,

$$\begin{aligned} \langle \partial_t \theta_{\delta h}(t), v\rangle &= \frac{1}{\delta}\langle \Theta_h^m - \Theta_h^{m-1}, v\rangle = \frac{1}{\delta}\langle \Theta_h^m - \Theta_h^{m-1}, P_h v\rangle \\ &= -a[\Theta_h^m, P_h v] + \langle f^m - \widetilde{f}^m, P_h v\rangle \\ &\le C\Big(\|\Theta_h^m\|_1 + \|f^m - \widetilde{f}^m\|_{-1}\Big)\|v\|_1 \qquad \forall\, v \in H_0^1(\Omega) \end{aligned}$$

so that

$$\|\partial_t\theta_{\delta h}(t)\|_{-1} \le C\Big(\|\Theta_h^m\|_1 + \|f^m - \widetilde{f}^m\|_{-1}\Big)\,.$$

Hence,

$$\begin{aligned} \|\partial_t\theta_{\delta h}\|^2_{L^2(0,T;H^{-1}(\Omega))} &= \sum_{m=1}^N \int_{t_{m-1}}^{t_m} \|\partial_t \theta_{\delta h}(t)\|_{-1}^2\,dt \\ &\le C\sum_{m=1}^N \Big(\int_{t_{m-1}}^{t_m} \|\Theta_h^m\|_1^2\,dt + \int_{t_{m-1}}^{t_m} \|f^m - \widetilde{f}^m\|_{-1}^2\,dt\Big) \end{aligned}$$

Adding up the above estimates for $\theta_{\delta h}$ and $\partial_t\theta_{\delta h}$ and utilizing (5.12) we have

$$\|\theta_{\delta h}\|^2_{\mathcal{W}(0,T)} \le C\|f-\widetilde{f}\|^2_{L^2(0,T;H^{-1}(\Omega))} + C\delta\|P_h(u_0-\widetilde{u}_0)\|_1^2 + C\|P_h(u_0-\widetilde{u}_0)\|_0^2\,. \tag{5.13}$$

Combining (5.5) with (5.6), (5.8), (5.9) and (5.13) we conclude

$$\|u_h-u_{\delta h}\|_{\mathcal{W}(0,T)} \le C\|f-\widetilde{f}\|_{L^2(0,T;H^{-1}(\Omega))} + C\delta\|P_h(u_0-\widetilde{u}_0)\|_1^2 + C\|u_0-\widetilde{u}_0\|_0^2 + C\delta^{1/2}\Big(\|\widetilde{f}\|_{L^2(0,T;L^2(\Omega))} + \|\widetilde{f}\|_{H^{1/2}(0,T;H^{-1}(\Omega))} + \|\widetilde{u}_0\|_1\Big)\,.$$

From the facts that $\widetilde{f}=f^\epsilon\to f$ in $L^2(0,T;H^{-1}(\Omega))$ and $\widetilde{u}_0=u_0^\epsilon\to u_0$ in $L^2(\Omega)$ as $\epsilon\to 0$ it is evident that all terms except $C\delta\|P_h(u_0-\widetilde{u}_0)\|_1^2$ on the right hand side of the last estimate tend to zero as $\epsilon\to0$ and $h,\delta\to0$. In case i), i.e., the case where $u_0\in H_0^1(\Omega)$, we may simply choose $u_0^\epsilon=u_0$ so that $\delta\|P_h(u_0-\widetilde{u}_0)\|_1^2\to0$ as $h,\delta\to0$. In case ii), i.e., the case where $u_0\in H^{1/2}(\Omega)$ and $\delta\le Ch$, we have

$$\delta\|P_h(u_0-\widetilde{u}_0)\|_1^2 \le \frac{C\delta}{h}\|P_h(u_0-\widetilde{u}_0)\|_{1/2}^2 \le C\|u_0-\widetilde{u}_0\|_{1/2}^2$$

so that $\delta\|P_h(u_0-\widetilde{u}_0)\|_1^2\to0$ as $h,\delta\to0$. In case iii), i.e., the case where $u_0\in L^2(\Omega)$ and $\delta\le Ch^2$, we have

$$\delta\|P_h(u_0-\widetilde{u}_0)\|_1^2 \le \frac{C\delta}{h^2}\|P_h(u_0-\widetilde{u}_0)\|_0^2 \le C\|u_0-\widetilde{u}_0\|_0^2$$

so that $\delta\|P_h(u_0-\widetilde{u}_0)\|_1^2\to0$ as $h,\delta\to0$. Hence, in all three cases we have proved that $\|u_h-u_{\delta h}\|_{\mathcal{W}(0,T)}\to0$ as $h,\delta\to0$. □

Remark 5.2. An alternative proof for the estimate of the norm $\|\theta_{\delta h}\|_{\mathcal{W}(0,T)}$ can be obtained through an application of Lemma 4.2 to (5.10) with $\overline{g}\equiv0, \overline{W}_h^0\equiv0$, $\overline{w}_h\equiv0$, $g\equiv f-\widetilde{f}$, $W_h^m\equiv\Theta_h^m$ and $w_{\delta h}\equiv\theta_{\delta h}$, and an application of Lemma 4.1.

Finally, using the triangle inequality

$$\|u-u_{\delta h}\|_{\mathcal{W}(0,T)} \le \|u-u_h\|_{\mathcal{W}(0,T)} + \|u_h-u_{\delta h}\|_{\mathcal{W}(0,T)}$$

and combining the results of Theorems 3.1, 4.7, 4.3 and 5.1 we arrive at the following results concerning fully discrete approximations of the linear parabolic problem:

Theorem 5.3. *Assume that $f\in L^2(0,T;H^{-1}(\Omega))$ and $u_0\in L^2(\Omega)$. Let $u\in\mathcal{W}(0,T)$ be the solution of (2.1) and $u_{\delta h}$ be defined by (3.5)–(3.6). If $u_0\in H_0^1(\Omega)$, or $u_0\in H^{1/2}(\Omega)$ and $\delta\le Ch$, Or $u_0\in H_0^1(\Omega)$ and $\delta\le Ch^2$, then*

$$\|u-u_{\delta h}\|_{\mathcal{W}(0,T)}\to0 \quad \text{as } \delta,h\to0\,.$$

Also, the following error estimates hold:

i): *If $u_0\in H^{1+\gamma}(\Omega)\cap H_0^1(\Omega)$ and*

$$f\in L^2(0,T;H^{-1+2\gamma}(\Omega)\cap\Phi_0^{-1+2\gamma}(\Omega))\cap H^\gamma(0,T;H^{-1}(\Omega))$$

for a $\gamma\in[0,1]$, then

$$\|u-u_{\delta h}\|_{\mathcal{W}(0,T)} \le C(\delta^\gamma+h^{2\gamma})\Big(\|f\|_{L^2(0,T;H^{-1+2\gamma}(\Omega))} + \|f\|_{H^\gamma(0,T;H^{-1}(\Omega))} + \|u_0\|_{1+\gamma}\Big).$$

ii): *If* $\delta \le Ch$, $u_0 \in H^{(1+3\gamma)/2}(\Omega) \cap \Phi_0^{(1+3\gamma)/2}(\Omega)$ *and*

$$f \in L^2(0,T;H^{-1+2\gamma}(\Omega) \cap \Phi_0^{-1+2\gamma}(\Omega)) \cap H^\gamma(0,T;H^{-1}(\Omega))$$

for a $\gamma \in [0,1]$, *then*

$$\|u - u_{\delta h}\|_{\mathcal{W}(0,T)} \le C(\delta^\gamma + h^{2\gamma})\Big(\|f\|_{L^2(0,T;H^{-1+2\gamma}(\Omega))} + \|f\|_{H^\gamma(0,T;H^{-1}(\Omega))} + \|u_0\|_{(1+3\gamma)/2}\Big).$$

iii): *If* $\delta \le Ch^2$, $u_0 \in H^{(1+3\gamma)/2}(\Omega) \cap \Phi_0^{(1+3\gamma)/2}(\Omega)$ *and*

$$f \in L^2(0,T;H^{-1+2\gamma}(\Omega) \cap \Phi_0^{-1+2\gamma}(\Omega)) \cap H^\gamma(0,T;H^{-1}(\Omega))$$

for a $\gamma \in [0,1]$, *then*

$$\|u - u_{\delta h}\|_{\mathcal{W}(0,T)} \le C(\delta^\gamma + h^{2\gamma})\Big(\|f\|_{L^2(0,T;H^{-1+2\gamma}(\Omega))} + \|f\|_{H^\gamma(0,T;H^{-1}(\Omega))} + \|u_0\|_{2\gamma}\Big).$$

iv): *If* $u_0 \in H^2(\Omega)$ *and* $f \in H^1(0,T;L^2(\Omega))$, *then*

$$\|u - u_{\delta h}\|_{\mathcal{W}(0,T)} \le C(\delta + h)\Big(\|f\|_{H^1(0,T;L^2(\Omega))} + \|u_0\|_2\Big).$$

Furthermore, in each of the cases i)–iv), *if*

$$u \in L^2(0,T;H^{r+1}(\Omega)) \cap H^1(0,T;H^{r-1}(\Omega))$$

for some $r \in [1,k]$, *then the* $O(h^{2\gamma})$ *term in* i)–iii) *or the* $O(h)$ *term in* iv) *on the right hand side of the error estimate can be replaced by*

$$Ch^r\Big(\|u\|_{L^2(0,T;H^{r+1}(\Omega))} + \|\partial_t u\|_{L^2(0,T;H^{r-1}(\Omega))}\Big).$$

Remark 5.4. Some remarks are in order for the last theorem. a) The conclusion in Case iv) follows from Theorem 4.3 and the implied regularity for the solution u (see [11]): $u \in L^\infty(0,T;H^2(\Omega))$ and $\partial_t u \in L^\infty(0,T;L^2(\Omega))$. The significance of this case is that no spatial differentiability is required of f. b) The last statement in the theorem about the $O(h^r)$ estimates follows from the known error estimates for semidiscrete finite element approximations.

We conclude this work by reiterating some features of the fully discrete scheme and its error estimates. i) The scheme is well defined for $f \in L^2(0,T;H^{-1}(\Omega))$ and $u_0 \in L^2(\Omega)$. ii) The $\mathcal{W}(0,T)$ norm we use to measure the error is stronger than the more commonly used $L^\infty(0,T;L^2(\Omega))$ norm. iii) The error estimates and convergence are obtained under minimal regularity; in particular, f is allowed have only a fractional order temporal derivative – in contrast, standard fully discrete error estimates in the literature such as those of [16] require $\partial_t f \in L^2(0,T;L^2(\Omega))$ (which rules out f being a timewise step function with values in $H^{-1}(\Omega)$). iv) Fractional order error estimates are derived under uni-directional regularity assumptions that are easier to verify than the $(t,\mathbf{x})$-mixed regularity assumptions; such results facilitate the derivation of error estimates for fully discrete approximations of semilinear parabolic PDEs and for fully discrete approximations of control control problems constrained by parabolic PDEs.

References

[1] R. Adams, *Sobolev Spaces*, Academic, Boston, 1975.
[2] C. Baiocchi and F. Brezzi, *Optimal error estimates for linear parabolic problems under minimal regularity assmptions*, Calcolo, 20 (1983) pp.143–176.
[3] G.A. Baker, J.H. Bramble and V. Thomee, *Single step Galerkin approximations for parabolic problems*, Math. Comp., 31 (1977), pp.818–847.
[4] J.H. Bramble and V. Thomee, *Discrete time Galerkin methods for a parabolic boundary value problem*, Mat. Pura Appl, 101 (1974), pp.115–152.
[5] J.H. Bramble and P. Sammon, *Efficient higher order single step methods for parabolic problems: Part 1*, Math. Comp., 35 (1980), pp.655–677.
[6] F. Brezzi, J. Rappaz, and P.-A. Raviart, *Finite dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions*, Numer. Math., 36 (1980), pp.1–25.
[7] P. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
[8] K. Chrysafinos and L.S. Hou, *Error estimates for semidiscrete finite element approximations of linear and semilinear parabolic equations under minimal regularity assumptions*, SIAM. J. Numer. Anal., 40 (2002), pp.282–306.
[9] J. Douglas Jr. and T.F. Dupont, *Galerkin methods for parabolic equations*, SIAM J. Numer. Anal., 7 (1970), pp.575-626.
[10] K. Eriksson, C. Johnson and V. Thomee, *Time discretization of parabolic problems by the discontinuous Galerkin method*, RAIRO Model. Math. Anal. Numer., 19 (1985), pp.611–643.
[11] L.C. Evans, *Partial Differential Equations*, AMS, Providence, RI, 1998.
[12] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer, Berlin, 1986.
[13] M.D. Gunzburger, L.S. Hou and W. Zhu, *Fully discrete finite element approximations of the forced Fisher equation*, J. Math. Anal. Appl., 313 (2006), pp.419-440.
[14] W. Hackbusch, *Optimal $H^{p.p/2}$ error estimates for a parabolic Galerkin method*, SIAM J. Numer. Anal., **18** (1981), pp.681-692.
[15] J.L. Lions and E. Magenes, *Non-homogeneous Boundary Value Problems and Applications*, I, II, Springer-Verlag, New York-Heidelberg-Berlin, 1972.
[16] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer-Verlag, Berlin-Heidelberg, 1994.
[17] P. Sammon, *Fully discrete approximation methods for parabolic problems with nonsmooth initial data*, SIAM J. Numer. Anal., 20 (1983), pp.437–470.
[18] R. Temam, *Navier-Stokes Equations*, AMS Chelsea Publishing, Providence, Rhode Island, 2001.
[19] V. Thomee, *Galerkin Finite Element Methods for Parabolic Equations*, Springer, Berlin, 1997.
[20] M. Zlamal, *Finite element multistep discretizations of parabolic boundary value problems*, Math. Comp., 29 (1975), pp.350–359.

Department of Mathematics, Iowa State University, Ames, IA 50011-2064, USA
E-mail: stevehou@iastate.edu

Department of Mathematics, Penn State University, University Park, PA 16802, USA
E-mail: zhu_w@math.psu.edu