ESAIM: M2AN 57 (2023) 3091–3111
https://doi.org/10.1051/m2an/2023076

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

FULLY DISCRETE POINTWISE SMOOTHING ERROR ESTIMATES FOR MEASURE VALUED INITIAL DATA

Dmitriy Leykekhman[1], Boris Vexler[2] and Jakob Wagner[3]

Abstract. In this paper we analyze a homogeneous parabolic problem with initial data in the space of regular Borel measures. The problem is discretized in time with a discontinuous Galerkin scheme of arbitrary degree and in space with continuous finite elements of orders one or two. We show parabolic smoothing results for the continuous, semidiscrete and fully discrete problems. Our main results are interior L^∞ error estimates for the evaluation at the endtime, in cases where the initial data is supported in a subdomain. In order to obtain these, we additionally show interior L^∞ error estimates for L^2 initial data and quadratic finite elements, which extends the corresponding result previously established by the authors for linear finite elements.

Mathematics Subject Classification. 65N30, 65N15.

Received April 3, 2023. Accepted September 4, 2023.

1. Introduction

In this work we discuss smoothing properties of the fully discrete approximation of the homogeneous parabolic problem

$$\begin{aligned} \partial_t v - \Delta v &= 0 \quad \text{in} \quad I \times \Omega, \\ v &= 0 \quad \text{on} \quad I \times \partial\Omega, \\ v(0) &= v_0 \quad \text{in} \quad \Omega, \end{aligned} \tag{1}$$

where $\Omega \subset \mathbb{R}^N$, $N = 2, 3$ is a bounded, convex, polygonal/polyhedral domain, and $I = (0, T]$ a bounded time interval. In particular we are interested in pointwise error estimates in the case when the initial condition v_0 is a regular Borel measure $v_0 \in \mathcal{M}(\Omega)$ supported in some subdomain Ω_0 such that $\overline{\Omega}_0 \subset \Omega$, for example a linear combination of Dirac delta functions, $v_0 = \sum_j \beta_j \delta_{x_j}$. Our main result of this paper establishes the fully discrete error estimate of the form

$$\|(v - v_{kh})(T)\|_{L^\infty(\Omega_0)} \le C(\Omega_0, T) \left(k^{2r+1} + \ell_{kh} h^{s+1}\right) \|v_0\|_{\mathcal{M}(\Omega)}. \tag{2}$$

Keywords and phrases. Optimal control, sparse control, initial data identification, smoothing estimates, parabolic problems, finite elements, discontinuous Galerkin, error estimates, pointwise error estimates.

[1] Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA.

[2] Chair of Optimal Control, Technical University of Munich, School of Computation Information and Technology, Department of Mathematics, Boltzmannstraße 3, 85748 Garching b. Munich, Germany.

[3] Chair of Optimal Control, Technical University of Munich, School of Computation Information and Technology, Department of Mathematics, Boltzmannstraße 3, 85748 Garching b. Munich, Germany.

*Corresponding author: vexler@ma.tum.de

Here $r \geq 0$ is the order of time discretization, $s = 1, 2$ is the order of the space discretization, and ℓ_{kh} is a logarithmic term that depends on the mesh size h and the maximum time step k. In order to simplify the presentation, we assume that v_0 is supported in the same subdomain Ω_0, in which the L^∞ error is estimated, wheareas in general, two different subdomains could be chosen. We would like to point out that the piecewise linear case $s = 1$ does not require any additional smoothness assumptions beyond regularity results available on convex domains Ω. The higher order convergence of $s = 2$ requires some additional smoothness assumptions, which are available for example on rectangular domains (*cf.* Sect. 6). In this case, the logarithmic term ℓ_{kh} only depends on k.

The above problem is classical and many important results are available in the literature. The L^2 theory for a uniform time partition is well presented in the classical textbook of Thomée [20]. Extensions to variable time steps are available in Eriksson *et al.* [7]. The $L^1 \to L^\infty$ stability results, are technically more difficult and one of the first papers in this direction was the work of Schatz *et al.* [18], where such results were established in two space dimensions for piecewise linear elements and strongly A-stable single step methods with uniform time steps. The sharpest result in the case of smooth domains and uniform time steps was obtained by A. Hansbo in [9].

In our previous paper [15], for piecewise linear space discretizations on a convex polygonal/polyhedral domain Ω and $v_0 \in L^2(\Omega)$, we have obtained

$$\|(v - v_{kh})(T)\|_{L^\infty(\Omega_0)} \leq C(T)\left(k^{2r+1} + \ell_{kh} h^2\right)\|v_0\|_{L^2(\Omega)},$$

with explicit form of the constant $C(T)$. Such results were required for obtaining sharp results in initial data estimation of the parabolic problems from final time observation [15]. However, in order to extend the results to the situation when the final time observation is taken at a finite number of points [14], which is more relevant in applications, we require the results of the form (2). This yields an error estimate for the adjoint state, which satisfies a backwards-in-time problem, with a final time condition given by a measure supported in the observation points. Since these points are fixed, this support is contained in a proper subdomain, and hence the assumptions for (2) are satisfied. In summary, the main contribution of our paper is the establishment of fully discrete error estimates (2) for Galerkin methods on potentially highly varying time partitions and quasi-uniform meshes on convex polygonal/polyhedral domains, without any additional smoothness assumptions in the case of piecewise linear case and with additional smoothness assumptions in the case of quadratic elements.

The rest of the paper is organized as follows. In Section 2, we review the notion of very weak solutions for parabolic homogeneous problems with initial data given in the space of regular Borel measures. In Section 3, we discuss space-time discretization schemes and introduce the semidiscrete and fully discrete Galerkin solutions to (1). In Section 4, we review and show continuous and discrete smoothing estimates for the continuous, semidiscrete and fully discrete solutions. Our main result will be the pointwise fully discrete error estimate for initial data in $\mathcal{M}(\Omega)$, see Theorem 5.3, which we establish in Section 5. Finally, in Section 6 we extend our main result to a higher order space discretization.

2. Very weak solutions and regularity

We begin by introducing the proper setup for the existence and regularity of the solution with measure valued initial data. Throughout this work, we use standard notations $L^p(\Omega)$, $W^{k,p}(\Omega)$, $W_0^{k,p}(\Omega)$ for the Lebesque and Sobolev spaces and abbreviate them by $H^k(\Omega)$, $H_0^k(\Omega)$, in case $p = 2$. The $L^2(\Omega)$ inner product will be denoted by $(\cdot,\cdot)_\Omega$. We denote the Bochner spaces of $W^{k,p}(\Omega)$ valued, $q-$integrable functions over the time interval I by $L^q(I; W^{k,p}(\Omega))$, and denote by $(\cdot,\cdot)_{I\times\Omega}$ the inner product on $L^2(I; L^2(\Omega)) \cong L^2(I \times \Omega)$. The space $\mathcal{M}(\Omega)$ of regular Borel measures can be identified with the dual space of $C_0(\Omega) := \{v \in C(\bar{\Omega}) : v|_{\partial\Omega} = 0\}$, *i.e.* it holds $\mathcal{M}(\Omega) \cong (C_0(\Omega))^*$. The norm on $\mathcal{M}(\Omega)$ is then given as $\|\mu\|_{\mathcal{M}(\Omega)} := \sup_{0\neq v\in C_0(\Omega)} \frac{\langle v,\mu\rangle}{\|v\|_{C_0(\Omega)}}$. Note that this norm is equivalent to the total variation norm $|\mu|(\Omega) = \mu^+(\Omega) + \mu^-(\Omega)$, where $\mu = \mu^+ - \mu^-$ is the Jordan decomposition of $\mu \in \mathcal{M}(\Omega)$. With theses notations fixed, we can state the very weak formulation of (1).

Definition 2.1. Let $v_0 \in \mathcal{M}(\Omega)$ be given. A function $v \in L^1(I \times \Omega)$ is called a very weak solution to the heat equation (1), if

$$-\int_{I\times\Omega} \left(\frac{\partial\varphi}{\partial t} + \Delta\varphi\right) v \, dx \, dt = \int_\Omega \varphi(0) \, dv_0 \qquad \text{for all } \varphi \in \Phi_T, \tag{3}$$

where the space Φ_T of all test functions is given by

$$\Phi_T = \{\varphi \in \Phi : \; \varphi(T) = 0 \text{ in } \Omega\} \quad \text{with} \quad \Phi = \left\{\varphi \in L^2(I; H_0^1(\Omega)) : \; \partial_t\varphi + \Delta\varphi \in L^\infty(I \times \Omega) \wedge \varphi(T) \in L^2(\Omega)\right\}.$$

With this definition, we have the following result (see [5], Lem. 2.2):

Theorem 2.2. *For a given $v_0 \in \mathcal{M}(\Omega)$, there exists a unique solution v in the sense of (3). The solution v lies in the space $L^r(I; W_0^{1,p}(\Omega))$ for any $p, r \in [1, 2)$ with $\frac{2}{r} + \frac{N}{p} > N + 1$, with the estimate*

$$\|v\|_{L^r(I;W_0^{1,p}(\Omega))} \le C_{r,p} \|v_0\|_{\mathcal{M}(\Omega)}.$$

Moreover, $v \in C([0,T]; W^{-1,p}(\Omega))$, making the evaluation $v(t)$ well defined at any $t \in [0,T]$, and in addition $v(T) \in L^2(\Omega)$ with

$$\|v(T)\|_{L^2(\Omega)} \le C_T \|v_0\|_{\mathcal{M}(\Omega)}.$$

For any $\varphi \in \Phi$ there holds

$$\int_\Omega \varphi(T)v(T) - \int_{(0,T)\times\Omega} (\partial_t\varphi + \Delta\varphi) \, v \, dx \, dt = \int_\Omega \varphi(0) \, dv_0.$$

In the second estimate of the above theorem, the constant C_T depends on T. We shall make this dependence more explicit in Lemma 4.2. For the error analysis below, we will require the following result.

Lemma 2.3. *Let $\tau \in (0,T)$ and let v_1 be the very weak solution of the heat equation on the subinterval $(0,\tau)$ in the sense of (3). Let v_2 be the weak solution of the heat equation on the subinterval (τ, T) with initial data $v_1(\tau) \in L^2(\Omega)$. Then v defined as*

$$v(t) = \begin{cases} v_1(t) & t \in (0,\tau] \\ v_2(t) & t \in (\tau, T) \end{cases}$$

is the very weak solution in the sense of (3).

Proof. The proof is straightforward. □

3. Discretization

In this section we describe the semidiscrete and fully discrete finite element discretizations of the homogeneous equation (1) and present smoothing type error estimates. To discretize the problem, we use continuous Lagrange finite elements of order $s \ge 1$ in space and discontinuous Galerkin methods of order $r \ge 0$ in time.

3.1. Time discretization

To be more precise, we partition $I = (0,T]$ into subintervals $I_m = (t_{m-1}, t_m]$ of length $k_m = t_m - t_{m-1}$, where $0 = t_0 < t_1 < \cdots < t_{M-1} < t_M = T$. The maximal and minimal time steps are denoted by $k = \max_m k_m$ and $k_{\min} = \min_m k_m$, respectively. We impose the following conditions on the time mesh (as in [12] or [16]):

(i) There are constants $c, \beta > 0$ independent on k such that

$$k_{\min} \ge ck^\beta.$$

(ii) There is a constant $\kappa > 0$ independent on k such that for all $m = 1, 2, \dots, M-1$

$$\kappa^{-1} \le \frac{k_m}{k_{m+1}} \le \kappa.$$

(iii) It holds $k \le \frac{T}{2r+2}$.

The semidiscrete space X_k^r for the case $v_0 \in L^2(\Omega)$ is taken as

$$X_k^r = \left\{ \varphi_k \in L^2(I; H_0^1(\Omega)) \;\middle|\; \varphi_k|_{I_m} \in \mathbb{P}_r(I_m; H_0^1(\Omega)),\ m = 1, 2, \dots, M \right\},$$

where $\mathbb{P}_r(I_m; V)$ is the space of polynomial functions of degree r in time on I_m with values in a Banach space V. However, for the semidiscrete formulation of (1) with measure valued initial data $v_0 \in \mathcal{M}(\Omega)$ to be well defined, at initial time the test functions φ_k need to be in $C_0(\Omega)$. Since for $N \ge 2$ the space $H_0^1(\Omega)$ is not embedded in the space of continuous functions, we need to modify the spaces of trial and test functions, as

$$\widetilde{X}_k^r = \left\{ \varphi_k \in L^2(I \times \Omega) : \varphi_k|_{I_m} \in \mathbb{P}_r(I_m; W_0^{1,p}(\Omega)),\ m = 1, 2, \dots, M \right\}$$

and

$$\widehat{X}_k^r = \left\{ \varphi_k \in L^2(I \times \Omega) : \varphi_k|_{I_m} \in \mathbb{P}_r(I_m; W_0^{1,p'}(\Omega)),\ m = 1, 2, \dots, M \right\},$$

for some $\frac{2N}{N+2} < p < \frac{N}{N-1}$ and $\frac{2N}{N-2} > p' > N$ the dual exponent satisfying $\frac{1}{p} + \frac{1}{p'} = 1$. In this setting, the embedding $W_0^{1,p'}(\Omega) \hookrightarrow C_0(\Omega)$, yields that $\langle v_0, \varphi_{k,0}^+ \rangle$ is well defined for all test functions $\varphi_k \in \widehat{X}_k^r$, while every trial function $v_k \in \widetilde{X}_k^r$ satisfies $v_k(t) \in W_0^{1,p}(\Omega) \hookrightarrow L^2(\Omega)$ for every $t \in I$. With these spaces, the dG(r) semidiscrete (in time) solution v_k of (1) for $v_0 \in \mathcal{M}(\Omega)$ is given by $v_k \in \widetilde{X}_k^r$ that satisfies

$$B(v_k, \varphi_k) = \left\langle v_0, \varphi_{k,0}^+ \right\rangle \quad \text{for all } \ \varphi_k \in \widehat{X}_k^r. \tag{4}$$

Here the bilinear form $B(\cdot,\cdot)$ is defined by

$$B(w, \varphi) = \sum_{m=1}^{M} \langle \partial_t w, \varphi \rangle_{I_m \times \Omega} + (\nabla w, \nabla \varphi)_{I \times \Omega} + \sum_{m=2}^{M} ([w]_{m-1}, \varphi_{m-1}^+)_\Omega + (w_0^+, \varphi_0^+)_\Omega, \tag{5}$$

where $\langle \cdot,\cdot \rangle_{I_m \times \Omega}$ is the duality product between $L^2(I_m; W^{-1,p}(\Omega))$ and $L^2(I_m; W_0^{1,p'}(\Omega))$. In the above definition we use the usual notation for functions with possible discontinuities at the nodes t_m:

$$w_m^+ = \lim_{\varepsilon \to 0^+} w(t_m + \varepsilon), \quad w_m^- = \lim_{\varepsilon \to 0^+} w(t_m - \varepsilon), \quad [w]_m = w_m^+ - w_m^-.$$

Remark 3.1. Note that whenever $v_0 \in L^2(\Omega)$ the formulation (4) is equivalent to searching $v_k \in X_k^r$, satisfying

$$B(v_k, \varphi_k) = \left(v_0, \varphi_{k,0}^+ \right)_\Omega \quad \text{for all } \ \varphi_k \in X_k^r. \tag{6}$$

Remark 3.2. Since we are dealing with a homogeneous parabolic problem with constant coefficients, the discontinuous Galerkin method actually coincides with subdiagonal Padé approximations and one can use, for example, a rational representation of the semidiscrete solution. While it is more convenient for our analysis to use the definition based on the bilinear form $B(\cdot,\cdot)$, this rational expression allows us to show wellposedness of the semidiscrete problem.

Theorem 3.3. *Let $v_0 \in \mathcal{M}(\Omega)$. Then the semidiscrete problem (4) has a unique solution $v_k \in \widetilde{X}_k^r$.*

Proof. It is sufficient to show the claim on the first time interval. Let $\{\phi_j : j = 0,\dots,r\}$ denote a basis of $\mathbb{P}_r(I_1;\mathbb{R})$. It is well known, that in the setting of (6), there exist polynomials $P_j(z)$, $j = 0,\dots,r$ and $Q(z)$ of degrees r and $r+1$, respectively, such that the variational formulation is equivalent to the rational representation

$$v_k|_{I_1} = \sum_{j=0}^{r} \phi_j \frac{P_j(-k_1\Delta)}{Q(-k_1\Delta)} v_0,$$

see [11], Section 4.1 of [13]. Here $Q(z)$ corresponds to the denominator of the subdiagonal $(r, r+1)$ Padé approximation of e^{-z}. The polynomial $Q(z)$ posesses $r+1$ complex zeroes $\xi_n \in \mathbb{C}$, $n = 1,\dots,r+1$. These all satisfy $Re(\xi_n) < 0$, and thus $\xi_n \in \rho(-k_1\Delta)$, *i.e.*, they are contained in the resolvent set of $-k_1\Delta$, see Theorem 1.1 of [17]. This implies that the operators $(\xi_n + k_1\Delta)^{-1} : L^2(\Omega) \to H_0^1(\Omega)$, are well defined. By Theorem 8 of [21] and the fact that the $(r, r+1)$ Padé approximation is of order $2r+1$, we know that the zeroes ξ_n of $Q(z)$ are pairwise distinct. Hence for the coefficients of $v_k|_{I_1}$, there holds a partial fraction decomposition, and for some $c_{j,n} \in \mathbb{C}$, $j = 0,\dots,r$, $n = 1,\dots,r+1$ we have the representation

$$v_k|_{I_1} = \sum_{j=0}^{r} \phi_j \sum_{n=1}^{r+1} c_{j,n}(\xi_n + k_1\Delta)^{-1} v_0. \tag{7}$$

By the elliptic theory, we can show, that $(\xi_n + k_1\Delta)^{-1} : \mathcal{M}(\Omega) \to W_0^{1,p}(\Omega)$, are well defined, which implies that (7) holds also true for $v_0 \in \mathcal{M}(\Omega)$. To show wellposedness of the elliptic problems, we employ the following construction. For any $\mu \in \mathcal{M}(\Omega)$, due to Corollary 2.7 of [6], there exists a unique solution $u_\mu \in W_0^{1,p}(\Omega)$ with $\frac{2N}{N+2} < p < \frac{N}{N-1}$ to

$$k_1(\nabla u_\mu, \nabla v)_\Omega = \langle \mu, v\rangle \quad \text{for all } v \in W_0^{1,p'}(\Omega).$$

By the embedding $W_0^{1,p}(\Omega) \hookrightarrow L^2(\Omega)$, it holds $(\xi_n + k_1\Delta)^{-1}u_\mu \in H_0^1(\Omega)$. From this, we can construct $(\xi_n + k_1\Delta)^{-1}\mu$ *via* $\xi_n(\xi_n + k_1\Delta)^{-1}u_\mu - u_\mu \in W_0^{1,p}(\Omega)$, which concludes the proof. $\square$

Remark 3.4. Due to Remark 3.1, the semidiscrete problem for $v_0 \in \mathcal{M}(\Omega)$, can equivalently be formulated using $H_0^1(\Omega)$ for test and trial functions on the intervals I_m, $m = 2,\dots,M$, instead of using the spaces $\widetilde{X}_k^r, \widehat{X}_k^r$ defined above. By definition, it holds $v_{k,1}^- \in W_0^{1,p}(\Omega) \hookrightarrow L^2(\Omega)$, hence on subsequent intervals, the solution lies in X_k^r. This construction of spaces was pursued in [15].

Rearranging the terms in (5), we obtain an equivalent (dual) expression for B:

$$B(w,\varphi) = -\sum_{m=1}^{M} \langle w, \partial_t\varphi\rangle_{I_m\times\Omega} + (\nabla w, \nabla\varphi)_{I\times\Omega} - \sum_{m=1}^{M-1} (w_m^-, [\varphi]_m)_\Omega + (w_M^-, \varphi_M^-)_\Omega. \tag{8}$$

Notice that for the very weak solution v to (3) and the semidiscrete solution $v_k \in \widetilde{X}_k^r$ to (4) we have the following orthogonality property:

$$B(v - v_k, \varphi_k) = 0 \quad \text{for all } \ \varphi_k \in \widehat{X}_k^r,$$

which can be shown by splitting v at t_1 according to Lemma 2.3, using the weak formulation on $(t_1, T]$ and a density argument on I_1 to show, that the very weak solution can be tested with semidiscrete functions $\varphi_k \in \widehat{X}_k^r$. Next we define the fully discrete approximation scheme.

3.2. Space discretization

For some $h_0 > 0$ and $h \in (0, h_0]$ let $\mathcal{T}$ denote a quasi-uniform triangulation of Ω with mesh size h, *i.e.* $\mathcal{T} = \{\tau\}$ is a partition of Ω into cells (triangles or tetrahedrons) τ of diameter h_τ and measure $|\tau|$ such that for $h = \max_\tau h_\tau$,

$$h_\tau \le h \le C|\tau|^{\frac{1}{N}}, \quad \text{for all } \ \tau \in \mathcal{T},$$

hold. Let $V_h^s \subset H_0^1(\Omega)$ be the usual space of conforming piecewise polynomial finite elements of degree s. We define the following three operators to be used in the sequel: the discrete Laplacian $\Delta_h \colon V_h^s \to V_h^s$, defined by

$$(-\Delta_h v_h, w_h)_\Omega = (\nabla v_h, \nabla w_h)_\Omega \quad \text{for all} \;\; v_h, w_h \in V_h^s,$$

the L^2 projection $P_h \colon L^2(\Omega) \to V_h^s$, defined by

$$(P_h v, w_h)_\Omega = (v, w_h)_\Omega \quad \text{for all} \;\; w_h \in V_h^s,$$

and the Ritz projection $R_h \colon H_0^1(\Omega) \to V_h^s$, defined by

$$(\nabla R_h v, \nabla w_h)_\Omega = (\nabla v, \nabla w_h)_\Omega \quad \text{for all} \;\; w_h \in V_h^s.$$

To obtain the fully discrete approximation of (1) we consider the space-time finite element space

$$X_{k,h}^{r,s} = \{\, v_{kh} \in X_k^r \mid v_{kh}|_{I_m} \in \mathbb{P}_r(I_m; V_h^s), \; m = 1, 2, \ldots, M \,\}.$$

We define a fully discrete cG(s)dG(r) approximation $v_{kh} \in X_{k,h}^{r,s}$ of (1) by

$$B(v_{kh}, \varphi_{kh}) = \left\langle v_0, \varphi_{kh,0}^+ \right\rangle \quad \text{for all} \;\; \varphi_{kh} \in X_{k,h}^{r,s}. \tag{9}$$

Similarly to the time semidiscretization, we have the following orthogonality relation for the semidiscrete solution v_k to (4) and the fully discrete solution $v_{kh} \in X_{k,h}^{r,s}$ to (9):

$$B(v_k - v_{kh}, \varphi_{kh}) = 0 \quad \text{for all} \;\; \varphi_{kh} \in X_{k,h}^{r,s}. \tag{10}$$

Existence of a unique solution v_{kh} is shown, *e.g.*, in [20]. At the end of this section, we would like to introduce the following truncation argument, which we will use often in our proofs. For $w_k, \varphi_k \in X_k^r$, we let $\tilde{w}_k = \chi_{(t_{\tilde{m}},T]} w_k$ and $\tilde{\varphi}_k = \chi_{(t_{\tilde{m}},T]} \varphi_k$, where $\chi_{(t_{\tilde{m}},T]}$ is the characteristic function on the interval $(t_{\tilde{m}}, T]$, for some $1 \le \tilde{m} \le M$, *i.e.* $\tilde{w}_k = 0$ on $I_1 \cup \cdots \cup I_{\tilde{m}}$ for some $\tilde{m}$ and $\tilde{w}_k = w_k$ on the remaining time intervals. Then from (5), we have the identity

$$B(\tilde{w}_k, \varphi_k) = B(w_k, \tilde{\varphi}_k) + (w_{k,\tilde{m}}^-, \varphi_{k,\tilde{m}}^+)_\Omega. \tag{11}$$

The same identity holds of course for fully discrete functions $w_{kh}, \varphi_{kh} \in X_{k,h}^{r,s}$.

4. Parabolic smoothing

In this section we review and establish smoothing properties of the continuous and discrete solutions, which are essential for the establishment of our main results.

4.1. Smoothing estimates for the continuous problem

It is well known that homogeneous parabolic problems have a strong smoothing effect. In particular, for $v_0 \in L^2(\Omega)$, the solution v to the problem (1) has the following smoothing property, see Chapter 1, Equation (1.14) of [3]

$$\|\partial_t^l v(t)\|_{L^2(\Omega)} + \|(-\Delta)^l v(t)\|_{L^2(\Omega)} \le \frac{C}{t^l} \|v_0\|_{L^2(\Omega)} \quad t > 0, \quad l = 0, 1, \ldots. \tag{12}$$

Remark 4.1. In many situations it is sufficient to have smoothing type estimates in L^2 norms and the corresponding smoothing results, for example in L^p norms, can obtained by the Gagliardo-Nirenberg inequality

$$\|g\|_{L^p(B)} \le C \, \|g\|_{H^2(B)}^{\alpha} \, \|g\|_{L^2(B)}^{1-\alpha}, \quad 2 \le p \le \infty, \quad \text{for } \alpha = \frac{N}{4} - \frac{N}{2p},$$

which holds for any subdomain $B \subset \Omega$ fulfilling the cone condition (in particular for $B = \Omega$) and for all $g \in H^2(B)$, see Theorem 3 of [1]. In particular, for $p = \infty$ on convex domains, we have

$$\|g\|_{L^\infty(\Omega)} \le C \|\Delta g\|_{L^2(\Omega)}^{\frac{N}{4}} \|g\|_{L^2(\Omega)}^{1-\frac{N}{4}}, \quad \text{for } g \in H^2(\Omega) \cap H_0^1(\Omega). \tag{13}$$

Thus using (13) and the smoothing estimates (12) for $v_0 \in L^2(\Omega)$, we immediately obtain

$$\|v(t)\|_{L^\infty(\Omega)} \le \frac{C}{t^{\frac{N}{4}}} \|v_0\|_{L^2(\Omega)} \quad t > 0. \tag{14}$$

Using a duality argument and the smoothing estimates, this result can be easily extended to the case $v_0 \in \mathcal{M}(\Omega)$. First we derive the explicit time dependence of the constant occuring in the estimate of Theorem 2.2.

Lemma 4.2. *Let $v_0 \in \mathcal{M}(\Omega)$ and v be the very weak solution of (1). Then*

$$\|v(t)\|_{L^2(\Omega)} \le \frac{C}{t^{\frac{N}{4}}} \|v_0\|_{\mathcal{M}(\Omega)} \quad t > 0.$$

Proof. We will establish the result for $t = T$. Define y to be the solution to the dual problem

$$\begin{cases} -\partial_t y - \Delta y = 0 & \text{in } I \times \Omega \\ y = 0 & \text{on } I \times \partial\Omega \\ y(T) = v(T) & \text{in } \Omega. \end{cases}$$

Then (14) applied to y yields $y(0) \in C_0(\Omega)$ and we have the estimate

$$\|v(T)\|_{L^2(\Omega)}^2 = (v(T), y(T))_\Omega = \langle v_0, y(0)\rangle \le C \|v_0\|_{\mathcal{M}(\Omega)} \|y(0)\|_{L^\infty(\Omega)} \le CT^{-\frac{N}{4}} \|v_0\|_{\mathcal{M}(\Omega)} \|v(T)\|_{L^2(\Omega)}.$$

Canceling $\|v(T)\|_{L^2(\Omega)}$ gives the result. □

Corollary 4.3. *Let $v_0 \in \mathcal{M}(\Omega)$ and v be the very weak solution of (1). Then*

$$\|\partial_t^l v(t)\|_{L^2(\Omega)} + \|(-\Delta)^l v(t)\|_{L^2(\Omega)} \le \frac{C}{t^{l+\frac{N}{4}}} \|v_0\|_{\mathcal{M}(\Omega)} \quad t > 0, \quad l = 0, 1, \dots .$$

Proof. This is a direct consequence of Theorem 2.2, Lemma 2.3 and the above smoothing result. The time dependency of the constant can be seen, by fixing $t \in (0, T)$ and setting $\tau = \frac{t}{2}$ in Lemma 2.3. Then by Lemma 4.2, we have $\left\|v\left(\frac{t}{2}\right)\right\|_{L^2(\Omega)} \le C\left(\frac{2}{t}\right)^{\frac{N}{4}} \|v_0\|_{\mathcal{M}(\Omega)}$. By (12), it also holds

$$\|\partial_t^l v(t)\|_{L^2(\Omega)} + \|(-\Delta)^l v(t)\|_{L^2(\Omega)} \le \frac{C}{(t - \frac{t}{2})^l} \left\|v\left(\frac{t}{2}\right)\right\|_{L^2(\Omega)} \le \frac{C}{t^{l+\frac{N}{4}}} \|v_0\|_{\mathcal{M}(\Omega)}.$$

□

By applying the Gagliardo-Nirenberg inequality (13), we immediately obtain:

Corollary 4.4. *Let $v_0 \in \mathcal{M}(\Omega)$ and v be the very weak solution of (1). Then*

$$\|v(t)\|_{L^\infty(\Omega)} \le \frac{C}{t^{\frac{N}{2}}} \|v_0\|_{\mathcal{M}(\Omega)}, \quad t > 0.$$

4.2. Smoothing estimates for the discrete solutions

For the time discontinuous Galerkin solutions, both the semidiscrete and the fully discrete, similar smoothing type estimates also hold (see [13], Thms. 3,4,5,10 and [15], Lem. 3.2 for general L^p norms, *cf.* also [7], Thm. 5.1 for the case of the L^2 norm).

Lemma 4.5. *Let $v_k \in X_k^r$ and $v_{kh} \in X_{k,h}^{r,s}$ be the semidiscrete and fully discrete solutions of (4) and (9), respectively, with $v_0 \in L^2(\Omega)$. Then, there exists a constant C independent of k and h such that*

$$\|v_k\|_{L^\infty(I;L^2(\Omega))} \le C\|v_0\|_{L^2(\Omega)} \quad \text{and} \quad \|v_{kh}\|_{L^\infty(I;L^2(\Omega))} \le C\|v_0\|_{L^2(\Omega)}.$$

Lemma 4.6. *Let $v_k \in X_k^r$ and $v_{kh} \in X_{k,h}^{r,s}$ be the semidiscrete and fully discrete solutions of (4) and (9), respectively, with $v_0 \in L^2(\Omega)$. Then, there exists a constant C independent of k and h such that*

$$\sup_{t\in I_m}\|\partial_t v_k(t)\|_{L^2(\Omega)} + \sup_{t\in I_m}\|\Delta v_k(t)\|_{L^2(\Omega)} + k_m^{-1}\|[v_k]_{m-1}\|_{L^2(\Omega)} \le \frac{C}{t_m}\|v_0\|_{L^2(\Omega)},$$
$$\sup_{t\in I_m}\|\partial_t v_{kh}(t)\|_{L^2(\Omega)} + \sup_{t\in I_m}\|\Delta_h v_{kh}(t)\|_{L^2(\Omega)} + k_m^{-1}\|[v_{kh}]_{m-1}\|_{L^2(\Omega)} \le \frac{C}{t_m}\|v_0\|_{L^2(\Omega)},$$

for $m = 1, 2, \ldots, M$. For $m = 1$ the jump term is understood as $[v_k]_0 = v_{k,0}^+ - v_0$ and $[v_{kh}]_0 = v_{kh,0}^+ - P_h v_0$.

The above estimates immediately imply the following stability result.

Corollary 4.7. *Under the assumptions of Lemma 4.6, we have*

$$\sum_{m=1}^{M}\Big(\|\partial_t v_k\|_{L^1(I_m;L^2(\Omega))} + \|\Delta v_k\|_{L^1(I_m;L^2(\Omega))} + k_m\|\Delta v_{k,m}^+\|_{L^2(\Omega)} + \|[v_k]_{m-1}\|_{L^2(\Omega)}\Big) \le C\ln\frac{T}{k}\|v_0\|_{L^2(\Omega)}$$

and

$$\sum_{m=1}^{M}\Big(\|\partial_t v_{kh}\|_{L^1(I_m;L^2(\Omega))} + \|\Delta_h v_{kh}\|_{L^1(I_m;L^2(\Omega))} + k_m\|\Delta_h v_{kh,m}^+\|_{L^2(\Omega)} + \|[v_{kh}]_{m-1}\|_{L^p(\Omega)}\Big) \le C\ln\frac{T}{k}\|v_0\|_{L^2(\Omega)}.$$

For sufficiently many time steps, applying Lemma 4.6 iteratively, we have the following result.

Lemma 4.8. *Let $v_k \in X_k^r$ and $v_{kh} \in X_{k,h}^{r,s}$ be the semidiscrete and fully discrete solutions of (4) and (9), respectively, with $v_0 \in L^2(\Omega)$. Then, for any $m \in \{1, 2, \ldots M\}$, any $l \le m$, there hold*

$$\sup_{t\in I_m}\|\partial_t(-\Delta)^{l-1}v_k(t)\|_{L^2(\Omega)} + \sup_{t\in I_m}\|(-\Delta)^l v_k(t)\|_{L^2(\Omega)} + k_m^{-1}\|[(-\Delta)^{l-1}v_k]_{m-1}\|_{L^2(\Omega)} \le \frac{C}{t_m^l}\|v_0\|_{L^2(\Omega)}$$

and

$$\sup_{t\in I_m}\|\partial_t(-\Delta_h)^{l-1}v_{kh}(t)\|_{L^2(\Omega)} + \sup_{t\in I_m}\|(-\Delta_h)^l v_{kh}(t)\|_{L^2(\Omega)} + k_m^{-1}\|[(-\Delta_h)^{l-1}v_{kh}]_{m-1}\|_{L^2(\Omega)} \le \frac{C}{t_m^l}\|v_0\|_{L^2(\Omega)},$$

provided $k \le \frac{t_m}{l+1}$.

Using the continuous (13) and the discrete version of the Gagliardo-Nirenberg inequality, namely

$$\|\chi\|_{L^\infty(\Omega)} \le C\|\Delta_h\chi\|_{L^2(\Omega)}^{\frac{N}{4}}\|\chi\|_{L^2(\Omega)}^{1-\frac{N}{4}}, \quad \text{for all } \chi \in V_h^s, \tag{15}$$

which for example was established for smooth domains in Lemma 3.3 of [9], but the proof is valid for convex domains as well, we immediately obtain the following smoothing result.

Corollary 4.9. *Under the assumptions of Lemmas 4.5 and 4.6 for all $m = 1, 2, \ldots, M$, we have*

$$\sup_{t\in I_m} \|v_k(t)\|_{L^\infty(\Omega)} \leq \frac{C}{t_m^{N/4}}\|v_0\|_{L^2(\Omega)} \quad \text{and} \quad \sup_{t\in I_m} \|v_{kh}(t)\|_{L^\infty(\Omega)} \leq \frac{C}{t_m^{N/4}}\|v_0\|_{L^2(\Omega)}.$$

Similarly to the continuous case, using a duality argument, the above smoothing results can be extended to $v_0 \in \mathcal{M}(\Omega)$.

Lemma 4.10. *Let $v_0 \in \mathcal{M}(\Omega)$, and let $v_k \in \widetilde{X}_k^r$ and $v_{kh} \in X_{k,h}^{r,s}$ be the semidiscrete and the fully discrete solutions of (4) and (9) respectively. For any $m \in \{1, 2, \ldots M\}$, there hold*

$$\|v_k(t_m)\|_{L^2(\Omega)} + \|v_{kh}(t_m)\|_{L^2(\Omega)} \leq \frac{C}{t_m^{\frac{N}{4}}}\|v_0\|_{\mathcal{M}(\Omega)}.$$

Proof. Let $m = 1, 2, \ldots, M$, and define $y_k \in \widehat{X}_k^r$ to be the semidiscrete solution of the backward problem

$$B(\psi_k, y_k) = (\psi_{k,m}^-, v_k(t_m))_\Omega, \quad \forall \psi_k \in \widetilde{X}_k^r,$$

where the right hand side is well defined, due to the assumptions on $\widetilde{X}_k^r$, yielding $v_k(t_m) \in L^2(\Omega)$. Since for this dual problem, the test functions are taken from the weaker space $\widetilde{X}_k^r$, choosing $\psi_k = v_k \in \widetilde{X}_k^r$, and using Corollary 4.9 for the backward problem, we have

$$\|v_k(t_m)\|_{L^2(\Omega)}^2 = B(v_k, y_k) = \left\langle v_0, y_{k,0}^+\right\rangle \leq \|v_0\|_{\mathcal{M}(\Omega)}\|y_k(0)\|_{L^\infty(\Omega)} \leq \frac{C}{t_m^{N/4}}\|v_0\|_{\mathcal{M}(\Omega)}\|v_k(t_m)\|_{L^2(\Omega)}.$$

Canceling, we obtain the result for the time semidiscrete solution v_k. The argument for the fully discrete solution v_{kh} is almost identical. ◻

From Lemma 4.8, we can obtain the following result

Lemma 4.11. *Let $v_k \in \widetilde{X}_k^r$ and $v_{kh} \in X_{k,h}^{r,s}$ be the semidiscrete and the fully discrete solutions of (4) and (9) respectively. Let $m \in \{1, 2, \ldots M\}$ large enough and $l \leq m$, such that $k \leq \min\{\frac{t_m}{4}, \frac{t_m}{2(l+1)}\}$, then there hold*

$$\sup_{t\in I_m} \|\partial_t(-\Delta)^{l-1}v_k(t)\|_{L^2(\Omega)} + \sup_{t\in I_m} \|(-\Delta)^l v_k(t)\|_{L^2(\Omega)} + \sup_{t\in I_m} k_m^{-1}\|[(-\Delta)^{l-1}v_k]_{m-1}\|_{L^2(\Omega)} \leq \frac{C}{t_m^{l+\frac{N}{4}}}\|v_0\|_{\mathcal{M}(\Omega)}$$

and

$$\sup_{t\in I_m} \|\partial_t(-\Delta_h)^{l-1}v_{kh}(t)\|_{L^2(\Omega)} + \sup_{t\in I_m} \|(-\Delta_h)^l v_{kh}(t)\|_{L^2(\Omega)} + \sup_{t\in I_m} k_m^{-1}\|[(-\Delta_h)^{l-1}v_{kh}]_{m-1}\|_{L^2(\Omega)} \leq \frac{C}{t_m^{l+\frac{N}{4}}}\|v_0\|_{\mathcal{M}(\Omega)}.$$

Proof. We will only establish semidiscrete smoothing estimates for measure valued initial data, the analysis for the fully discrete solution is similar. Combining Lemma 4.8 with Lemma 4.10, gives us for all $m > \tilde{m} + l$:

$$\begin{aligned}&\sup_{t\in I_m} \|\partial_t(-\Delta)^{l-1}v_k(t)\|_{L^2(\Omega)} + \sup_{t\in I_m} \|(-\Delta)^l v_k(t)\|_{L^2(\Omega)} + k_m^{-1}\|[(-\Delta)^{l-1}v_k]_{m-1}\|_{L^2(\Omega)} \\ \leq &\frac{C}{(t_m - t_{\tilde{m}})^l}\|v_k(t_{\tilde{m}})\|_{L^2\Omega)} \leq \frac{C}{(t_m - t_{\tilde{m}})^l \ t_{\tilde{m}}^{\frac{N}{4}}} \|v_0\|_{\mathcal{M}(\Omega)}.\end{aligned}$$

For fixed t_m with m large enough such that $k \leq \min\{\frac{t_m}{4}, \frac{t_m}{2(l+1)}\}$ we apply the above argument to $t_{\tilde{m}}$ such that $\frac{t_m}{2} \in (t_{\tilde{m}-1}, t_{\tilde{m}}]$. By the requirements on k we obtain on the one hand that there are at least l timesteps between

$t_{\tilde{m}}$ and t_m, which allows us to use the smoothing estimate of Lemma 4.8 for the specified l. On the other hand, we obtain

$$t_m - t_{\tilde{m}} \geq \frac{t_m}{2} - k \geq \frac{t_m}{2} - \frac{t_m}{4} = \frac{t_m}{4} \qquad \text{yielding} \qquad (t_m - t_{\tilde{m}})^{-l} \leq 4^l \frac{1}{t_m^l}.$$

The choice $\frac{t_m}{2} \leq t_{\tilde{m}}$ gives $t_{\tilde{m}}^{-\frac{N}{4}} \leq 2^{\frac{N}{4}} t_m^{-\frac{N}{4}}$, which allows us to eliminate $t_{\tilde{m}}$ in the final bound and obtain

$$\sup_{t\in I_m} \|\partial_t(-\Delta)^{l-1} v_k(t)\|_{L^2(\Omega)} + \sup_{t\in I_m} \|(-\Delta)^l v_k(t)\|_{L^2(\Omega)} + k_m^{-1}\|[(-\Delta)^{l-1}v_k]_{m-1}\|_{L^2(\Omega)} \leq C(l,N) t_m^{-l-\frac{N}{4}} \|v_0\|_{\mathcal{M}(\Omega)}.$$

□

5. Smoothing type error estimates

First we review smoothing results with the initial data in $L^2(\Omega)$ and then extend the corresponding results to $\mathcal{M}(\Omega)$.

5.1. Review of pointwise smoothing error estimates for $v_0 \in L^2(\Omega)$

In [15], we have established the following pointwise fully discrete error estimate.

Proposition 5.1. *Let $v_0 \in L^2(\Omega)$, let v and $v_{kh} \in X_{k,h}^{r,1}$ satisfy (1) and (9), respectively. Then for any subdomain Ω_0 with $\overline{\Omega}_0 \subset \Omega$ there holds*

$$\|(v - v_{kh})(T)\|_{L^\infty(\Omega_0)} \leq C(T,\Omega_0) \left(\ell_{kh} h^2 + k^{2r+1}\right) \|v_0\|_{L^2(\Omega)},$$

where $\ell_{kh} = \ln\frac{T}{k} + |\ln h|$ and $C(T,\Omega_0)$ is a constant that depends on T and Ω_0 and the explicit form can be traced from the proof.

The proof of the above result was based on the following splitting of the error

$$(v - v_{kh})(T) = (v - v_k)(T) + (R_h v_k - v_{kh})(T) + (v_k - R_h v_k)(T). \tag{16}$$

Then each term was treated separately. The first error term was estimated in Theorem 3.8 of [15] by

$$\|(v - v_k)(T)\|_{L^\infty(\Omega)} \leq C(T) k^{2r+1} \|v_0\|_{L^2(\Omega)}, \tag{17}$$

with $C(T) \sim T^{-(2r+1+\frac{N}{4})}$. The above estimate follows from (see [15], Lem. 7.2)

$$\left\|(-\Delta)^j (v - v_k)(T)\right\|_{L^2(\Omega)} \leq C_j(T) k^{2r+1} \|v_0\|_{L^2(\Omega)}, \quad j = 0,1,\ldots.$$

The second error term in (16) satisfies,

$$\|(R_h v_k - v_{kh})(T)\|_{L^\infty(\Omega)} \leq C(T) \ln\frac{T}{k} h^2 \|v_0\|_{L^2(\Omega)},$$

which followed from (see [15], Lem. 8.2–8.3)

$$\|(-\Delta_h)^j (R_h v_k - v_{kh})(T)\|_{L^2(\Omega)} \leq C_j(T) \ln\frac{T}{k} h^2 \|v_0\|_{L^2(\Omega)}, \quad j = 0,1,$$

and the discrete Gagliardo-Nirenberg inequality (15). Here, we point out that the treatment of the first and the second terms of (16) do not require the condition $\overline{\Omega}_0 \subset \Omega$, they are global in nature. Finally, the estimate of the last term in (16) follows from the interior elliptic error estimate (*cf.* [19])

$$\|(v_k - R_h v_k)(T)\|_{L^\infty(\Omega_0)} \leq C(T,\Omega_0) |\ln h| h^2 \|v_0\|_{L^2(\Omega)}.$$

5.2. Pointwise smoothing error estimates for $v_0 \in \mathcal{M}(\Omega)$

We now turn towards proving the pointwise error estimate for measure valued initial data. To this end, first recall that in Lemma 5.1 of [15] we have shown the following L^2 error estimate for parabolic problems with initial data in $\mathcal{M}(\Omega)$, where for the spatial estimate we impose a condition on the support of v_0.

Lemma 5.2. *Let $v_0 \in \mathcal{M}(\Omega)$ with $\operatorname{supp} v_0 \subset \Omega_0$ for some subdomain $\Omega_0 \subset \overline{\Omega}_0 \subset \Omega$ and let v, $v_k \in \widetilde{X}_k^r$ and $v_{kh} \in X_{k,h}^{r,1}$ the continuous, semidiscrete and fully discrete solutions to (1), (4) and (9) respectively. Then there hold the estimates*

$$\begin{aligned}\|(v - v_k)(T)\|_{L^2(\Omega)} &\le C(T)k^{2r+1}\|v_0\|_{\mathcal{M}(\Omega)}\\ \|(v_k - v_{kh})(T)\|_{L^2(\Omega)} &\le C(\Omega_0, T)\ell_{kh}h^2\|v_0\|_{\mathcal{M}(\Omega)},\end{aligned}$$

where $\ell_{kh} = \ln\frac{T}{k} + |\ln h|$ and $C(T,\Omega_0)$ is a constant that depends on T and Ω_0 and the explicit form can be traced from the proof.

Our main result can now be obtained directly by introducing an auxiliary solution and the smoothing results presented in Section 4. We first prove the error estimate for the spatial discretization. The proof of the error estimate for the time semidiscretization follows the same steps under milder assumptions, see Lemma 5.4 below.

Theorem 5.3. *Let $v_0 \in \mathcal{M}(\Omega)$, let $v_k \in \widetilde{X}_k^r$ and $v_{kh} \in X_{k,h}^{r,1}$ satisfy (4) and (9), respectively. Then for any subdomain Ω_0 with $\overline{\Omega}_0 \subset \Omega$ and $\operatorname{supp} v_0 \subset \Omega_0$ there holds*

$$\|(v_k - v_{kh})(T)\|_{L^\infty(\Omega_0)} \le C(T,\Omega_0)\ell_{kh}h^2\|v_0\|_{\mathcal{M}(\Omega)},$$

where $\ell_{kh} = \ln\frac{T}{k} + |\ln h|$ and $C(T,\Omega_0)$ is a constant that depends on T and Ω_0 and the explicit form can be traced from the proof.

Proof. As done in the proofs of the smoothing results, we begin by splitting the time interval. To this end let $\tilde{m}$ be such that $\frac{T}{2} \in I_{\tilde{m}}$. We introduce a fully discrete auxiliary state $\hat{v}_{kh} \in X_{k,h}^{r,1}$, defined by

$$B(\hat{v}_{kh}, \varphi_{kh}) = (v_{k,\tilde{m}-1}^-, \varphi_{kh,\tilde{m}-1}^+)_\Omega \quad \text{for all} \quad \varphi_{kh} \in X_{k,h}^{r,1}.$$

Note that by definition $\hat{v}_{kh} \equiv 0$ on $I_1 \cup \ldots \cup I_{\tilde{m}-1}$ and it satisfies a discrete problem on $I_{\tilde{m}} \cup \ldots \cup I_M$ with initial condition $v_{k,\tilde{m}-1}^-$ at time $t_{\tilde{m}-1}$. By the triangle inequality, we obtain

$$\|(v_k - v_{kh})(T)\|_{L^\infty(\Omega_0)} \le \|(v_k - \hat{v}_{kh})(T)\|_{L^\infty(\Omega_0)} + \|(\hat{v}_{kh} - v_{kh})(T)\|_{L^\infty(\Omega_0)},$$

where for the first term, we obtain with Proposition 5.1 and the semidiscrete parabolic smoothing result of Lemma 4.10

$$\begin{aligned}\|(v_k - \hat{v}_{kh})(T)\|_{L^\infty(\Omega_0)} &\le C(T - t_{\tilde{m}-1}, \Omega_0)\ell_{kh}h^2\|v_{k,\tilde{m}-1}^-\|_{L^2(\Omega)}\\ &\le C(T - t_{\tilde{m}-1}, \Omega_0)t_{\tilde{m}-1}^{-\frac{N}{4}}\ell_{kh}h^2\|v_0\|_{\mathcal{M}(\Omega)}.\end{aligned}$$

For the second error term, we observe that the difference $\hat{v}_{kh} - v_{kh}$ satisfies a fully discrete parabolic equation on the intervals $I_{\tilde{m}} \cup \ldots \cup I_M$ for the initial data $v_{k,\tilde{m}-1}^- - v_{kh,\tilde{m}-1}^-$. Hence, the discrete Gagliardo-Nirenberg inequality (15) and the fully discrete smoothing results of Lemmas 4.6 and 4.8 yield

$$\begin{aligned}\|(\hat{v}_{kh} - v_{kh})(T)\|_{L^\infty(\Omega_0)} &\le C\|(\hat{v}_{kh} - v_{kh})(T)\|_{L^2(\Omega)}^{\frac{1}{2}}\|\Delta_h(\hat{v}_{kh} - v_{kh})(T)\|_{L^2(\Omega)}^{\frac{1}{2}}\\ &\le C(T - t_{\tilde{m}-1})^{-\frac{1}{2}-\frac{N}{4}}\|v_{k,\tilde{m}-1}^- - v_{kh,\tilde{m}-1}^-\|_{L^2(\Omega)}\end{aligned}$$

We apply Lemma 5.2 in order to estimate the L^2-error of the full discretization at the intermediate point in time, which yields

$$\|(\hat{v}_{kh} - v_{kh})(T)\|_{L^\infty(\Omega_0)} \le C(t_{\tilde{m}-1}, \Omega_0)(T - t_{\tilde{m}-1})^{-\frac{1}{2}-\frac{N}{4}} \ell_{kh} h^2 \|v_0\|_{\mathcal{M}(\Omega)}.$$

Since the assumptions on k and $\tilde{m}$ yield $\frac{T}{4} \le t_{\tilde{m}-1} \le \frac{T}{2}$ and $\frac{T}{2} \le T - t_{\tilde{m}-1} \le \frac{3T}{4}$, as before we can replace all quantities involving $t_{\tilde{m}-1}$ by ones only dependent of T, which concludes the proof. □

Note that by exactly the same technique, we can also derive the corresponding error estimate for the semidiscrete problem, which is global in Ω and no constraint on $\operatorname{supp} v_0$ is required. This is due to the fact, that the semidiscrete results of (17) and Lemma 5.1 of [15] hold in this more general setting. There holds the following result.

Lemma 5.4. *Let $v_0 \in \mathcal{M}(\Omega)$ and v and $v_k \in \widetilde{X}_k^r$ be the very weak and semidiscrete solutions to* (3) *and* (4) *respectively. Then there holds*

$$\|(v - v_k)(T)\|_{L^\infty(\Omega)} \le C(T) k^{2r+1} \|v_0\|_{\mathcal{M}(\Omega)},$$

with $C(T) \sim T^{-(2r+1+\frac{N}{2})}$.

6. Higher order space discretizations

Our main result from the previous section, Theorem 5.3, was established for piecewise linear finite elements only and does not require any additional smoothness assumptions on the solutions beyond H^2 regularity that is provided by the convexity of the domain. If additional regularity is available, for example,

$$|v|_{H^3(\Omega)} \le C \|\Delta v\|_{H_0^1(\Omega)}$$

for any $v \in H_0^1(\Omega)$ with $\Delta v \in H_0^1(\Omega)$, then the results of Proposition 5.1 can be extended (with an improved rate) to the case of quadratic Lagrange finite elements which we will denote by V_h^2 in this section.

Remark 6.1. Since due to Remark 3.4 for each $t \in (t_1, T]$, the solution v_k to the semidiscrete problem (4) satisfies $v_k(t) \in H_0^1(\Omega)$, one can also show straightforwardly that

$$\Delta v_k(t),\ \partial_t \Delta v_k(t) \in H_0^1(\Omega) \ \text{ for all } t \in I_m,\ m \ge 2 \quad \text{and} \quad \Delta^2 v_k(t) \in H_0^1(\Omega) \ \text{ for all } t \in I_m,\ m \ge 3.$$

Additional regularity is available on special domains, for example on rectangles, right or equilateral triangles. We make the following assumption of the domain Ω.

Assumption 6.2. *For every $u \in H_0^1(\Omega)$ with $\Delta u \in H_0^1(\Omega)$ there holds $u \in H^3(\Omega)$. Moreover, there exists a constant C independent of u such that*

$$\|u\|_{H^3(\Omega)} \le C \|\nabla \Delta u\|_{L^2(\Omega)}.$$

Example 6.3. This assumption holds for example on a rectangle, see Lemma 2.4 of [10]. In this case the solution u to the elliptic equation

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

with $f \in H_0^1(\Omega)$ possesses the $H^3(\Omega)$ regularity and the estimate

$$\|u\|_{H^3(\Omega)} \le C \|\nabla f\|_{L^2(\Omega)}$$

holds. Thus, Assumption 6.2 is satisfied in this case.

Lemma 6.4. *Let Ω satisfy Assumption 6.2.*
(1) Let $u, \Delta u \in H_0^1(\Omega)$, and $\Delta^2 u \in L^2(\Omega)$. Then there holds

$$\|u\|_{H^3(\Omega)}^2 \le C\|\Delta u\|_{L^2(\Omega)}\|\Delta^2 u\|_{L^2(\Omega)}. \tag{18}$$

(2) Let Ω_0 be a subdomain with $\overline{\Omega}_0 \subset \Omega$, let $u, \Delta u, \Delta^2 u \in H_0^1(\Omega)$, and $\Delta^3 u \in L^2(\Omega)$. Then there holds

$$\|u\|_{H^5(\Omega_0)}^2 \le C\|\Delta^2 u\|_{L^2(\Omega)}\|\Delta^3 u\|_{L^2(\Omega)}.$$

Proof. (1) For $v \in H_0^1(\Omega)$ with $\Delta v \in L^2(\Omega)$ one directly obtains

$$\|\nabla v\|_{L^2(\Omega)}^2 \le \|v\|_{L^2(\Omega)}\|\Delta v\|_{L^2(\Omega)}.$$

Due to $\Delta u \in H_0^1(\Omega)$ and $\Delta^2 u \in L^2(\Omega)$ this inequality can be applied to $v = \Delta u$ leading to

$$\|\nabla \Delta u\|_{L^2(\Omega)}^2 \le \|\Delta u\|_{L^2(\Omega)}\|\Delta^2 u\|_{L^2(\Omega)}.$$

Thus, Assumption 6.2 implies the desired estimate.
(2) Using a higher interior regularity result, see Chapter 6.3, Theorem 2 of [8], we obtain

$$\|u\|_{H^5(\Omega_0)} \le C(\|\Delta u\|_{H^3(\Omega)} + \|u\|_{L^2(\Omega)}).$$

Since $\Delta^2 u \in H_0^1(\Omega)$ and $\Delta^3 u \in L^2(\Omega)$ we can apply (18) to Δu leading to

$$\|\Delta u\|_{H^3(\Omega)}^2 \le C\|\Delta^2 u\|_{L^2(\Omega)}\|\Delta^3 u\|_{L^2(\Omega)}.$$

This leads to

$$\|u\|_{H^5(\Omega_0)}^2 \le C\|\Delta^2 u\|_{L^2(\Omega)}\|\Delta^3 u\|_{L^2(\Omega)} + C\|\Delta u\|_{L^2(\Omega)}\|\Delta^2 u\|_{L^2(\Omega)},$$

which proves the desired result by $\|\Delta^j u\|_{L^2(\Omega)} \le C\|\Delta^{j+1} u\|_{L^2(\Omega)}$ for $j = 1, 2$. □

Complementing the standard error estimates for the Ritz projection in the L^2 and H^1 norms, under Assumption 6.2, we also have the following negative norm estimate. Note that even though no H^3 regularity of the solution u is used explicitly in the estimates, the duality argument used to prove the result, requires the assumption to hold true for any H_0^1 right hand side.

Lemma 6.5. *Let $u \in H_0^1(\Omega)$ and Assumption 6.2 hold true. Then it holds*

$$\|u - R_h u\|_{H^{-1}(\Omega)} \le Ch^2\|\nabla(u - R_h u)\|_{L^2(\Omega)}.$$

If further $u \in H^2(\Omega)$, then it holds

$$\|u - R_h u\|_{H^{-1}(\Omega)} \le Ch^3\|u\|_{H^2(\Omega)} \le Ch^3\|\Delta u\|_{L^2(\Omega)}.$$

Proof. The first estimate is proved by a duality argument in Theorem 5.8.3 of [4]. The second estimate then follows with the standard H^1 error estimate and H^2 regularity. □

Under Assumption 6.2 we can establish the main results of this section. We first consider again the case of L^2 initial data. The extension to $v_0 \in \mathcal{M}(\Omega)$ then follows analogously to the case of linear finite elements.

Theorem 6.6. *Let $v_0 \in L^2(\Omega)$, let v and $v_{kh} \in X_{k,h}^{r,2}$ satisfy (1) and (9), respectively. Then for any subdomain Ω_0 with $\overline{\Omega}_0 \subset \Omega$ there holds*

$$\|(v - v_{kh})(T)\|_{L^\infty(\Omega_0)} \le C(T,d)\left(\ell_k h^3 + k^{2r+1}\right)\|v_0\|_{L^2(\Omega)},$$

where $\ell_k = \ln \frac{T}{k}$, $d = \mathrm{dist}(\Omega_0, \partial\Omega)$ and $C(T,d)$ is a constant depending on T and d.

6.1. Proof of Theorem 6.6

The exact dependence of the constant C on T and d is available in the proof of this result. The rest of this section is devoted to the establishment of the above theorem. The proof for the quadratic case is similar to the proof for the piecewise linear case, but requires some modifications. As it was done in [15], we split the error as

$$(v - v_{kh})(T) = (v - v_k)(T) + (v_k - R_h v_k)(T) + (R_h v_k - v_{kh})(T) =: T_1 + T_2 + T_3. \tag{19}$$

The first time semidiscrete term T_1 is already estimated in Theorem 3.8 of [15]. The second term T_2 can again be estimated by the interior pointwise error estimates of Theorem 5.1 of [19],

$$\|(v_k - R_h v_k)(T)\|_{L^\infty(\Omega_0)} \le C\|v_k(T) - \chi\|_{L^\infty(\Omega_d)} + Cd^{-N/2}\|(v_k - R_h v_k)(T)\|_{L^2(\Omega)}, \tag{20}$$

for any $\chi \in V_h^2$, where Ω_d is a subdomain satisfying $\overline{\Omega}_0 \subset \Omega_d \subset \overline{\Omega}_d \subset \Omega$ and $d = \mathrm{dist}(\Omega_0, \partial\Omega_d)$. We note that in contrast to the linear elements, in the above estimate the logarithmic term is not needed. By the approximation theory and the Sobolev embedding $H^5(\Omega_d) \hookrightarrow W^{3,\infty}(\Omega_d)$ (see *e.g.* [2], Thm. 4.12), Lemma 6.4 and the discrete parabolic smoothing result of Lemma 4.6, we obtain

$$\begin{aligned}\|v_k(T) - \chi\|_{L^\infty(\Omega_d)} &\le Ch^3\|v_k(T)\|_{W^{3,\infty}(\Omega_d)} \le Ch^3\|v_k(T)\|_{H^5(\Omega_d)}\\ &\le Ch^3\|\Delta^2 v_k(T)\|^{\frac12}_{L^2(\Omega)}\|\Delta^3 v_k(T)\|^{\frac12}_{L^2(\Omega)} \le \frac{Ch^3}{T^{\frac52}}\|v_0\|_{L^2(\Omega)}.\end{aligned}$$

The pollution term $\|(v_k - R_h v_k)(T)\|_{L^2(\Omega)}$ from (20), can be estimated using global elliptic estimates in L^2 norm, Lemmas 6.4 and 4.11 as

$$\|(v_k - R_h v_k)(T)\|_{L^2(\Omega)} \le Ch^3\|v_k(T)\|_{H^3(\Omega)} \le Ch^3\|\Delta v_k(T)\|^{\frac12}_{L^2(\Omega)}\|\Delta^2 v_k(T)\|^{\frac12}_{L^2(\Omega)} \le \frac{Ch^3}{T^{\frac32}}\|v_0\|_{L^2(\Omega)}.$$

Thus,

$$\|(v_k - R_h v_k)(T)\|_{L^\infty(\Omega_0)} \le C(T, \Omega_0)h^3\|v_0\|_{L^2(\Omega)},$$

and it remains to estimate the last term T_3 of (19). As done in Lemmas 8.2–8.3 of [15], this will be achieved by estimating

$$\|(-\Delta_h)^j (R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}, \quad j = 0, 1, \tag{21}$$

and the discrete Gagliardo-Nirenberg inequality (15). The proof of the above estimates was facilitated by the following technical lemma, see Lemma 8.1 of [15].

Lemma 6.7. *Let $v_0 \in L^2(\Omega)$, let $v_k \in \tilde{X}_k^r$ and $v_{kh} \in X_{kh}^{r,1}$ satisfy (4) and (9), respectively. There exists a constant C independent of k, h, and T such that*

$$\|\Delta_h^{-1}(P_h v_k - v_{kh})(T)\|_{L^2(\Omega)} \le Ch^2 \ln\frac{T}{k}\|v_0\|_{L^2(\Omega)}.$$

In order to prove Theorem 6.6 we thus first extend Lemma 6.7 to quadratic finite elements in space, in order to estimate the terms of (21).

Lemma 6.8. *Let $v_k \in X_k^r$ and $v_{kh} \in X_{k,h}^{r,2}$ be the semidiscrete and fully discrete solutions of (4) and (9), respectively for $v_0 \in L^2(\Omega)$. Then there exists a constant C independent of h,k and T such that*

$$\|\Delta_h^{-2}\left(P_h v_k - v_{kh}\right)(T)\|_{L^2(\Omega)} \le Ch^3 \ln\frac{T}{k}\|v_0\|_{L^2(\Omega)}.$$

Proof. Let $z_{kh} \in X^{r,2}_{k,h}$ be the solution to a dual problem with $z_{kh}(T) = \Delta_h^{-2}(P_h v_k - v_{kh})(T)$, *i.e.*

$$B(\chi_{kh}, z_{kh}) = \big(\chi_{kh}(T), \Delta_h^{-2}(P_h v_k - v_{kh})(T)\big) \quad \text{for all} \quad \chi_{kh} \in X^{r,2}_{k,h}.$$

Choosing $\chi_{kh} = \Delta_h^{-2}(P_h v_k - v_{kh})$, and using the Galerkin orthogonality (10) of v_k and v_{kh}, we obtain

$$\mathcal{Z} := \|\Delta_h^{-2}(P_h v_k - v_{kh})(T)\|^2_{L^2(\Omega)} = B(\Delta_h^{-2}(P_h v_k - v_{kh}), z_{kh}) = B(P_h v_k - v_{kh}, \Delta_h^{-2} z_{kh}) = B(P_h v_k - v_k, \Delta_h^{-2} z_{kh}).$$

Note, that since z_{kh} is piecewise polynomial in time, with values in V_h^2, for every $t \in I$, it holds $\Delta_h^{-2} z_{kh}(t) \in V_h^2$ and for every $t \in I \setminus \{t_0, t_1, \ldots, t_M\}$ it holds $\partial_t \Delta_h^{-2} z_{kh}(t) \in V_h^2$. Using the dual representation of B, given in (8), and the definition P_h, all $L^2(\Omega)$ inner products vanish, and it holds

$$\mathcal{Z} = (\nabla(P_h v_k - v_k), \nabla(\Delta_h^{-2} z_{kh}))_{I\times\Omega}.$$

In this inner product, we can replace v_k by its Ritz projection $R_h v_k$ and obtain after applying the definitions of Δ_h, P_h and the duality pairing

$$\begin{aligned}\mathcal{Z} = (\nabla(P_h v_k - R_h v_k), \nabla(\Delta_h^{-2} z_{kh}))_{I\times\Omega} &= -(P_h v_k - R_h v_k, \Delta_h^{-1} z_{kh})_{I\times\Omega} = -(v_k - R_h v_k, \Delta_h^{-1} z_{kh})_{I\times\Omega}\\ &\le \int_I \|(v_k - R_h v_k)(t)\|_{H^{-1}(\Omega)} \|\Delta_h^{-1} z_{kh}(t)\|_{H_0^1(\Omega)}\, dt.\end{aligned}$$

The second term in the integral for each fixed t, can be estimated as follows,

$$\begin{aligned}\|\Delta_h^{-1} z_{kh}(t)\|^2_{H_0^1(\Omega)} &\le C\left(\nabla \Delta_h^{-1} z_{kh}(t), \nabla \Delta_h^{-1} z_{kh}(t)\right)_\Omega\\ &= -C\left(z_{kh}(t), \Delta_h^{-1} z_{kh}(t)\right)_\Omega\\ &\le C\|z_{kh}(t)\|_{L^2(\Omega)} \|\Delta_h^{-1} z_{kh}(t)\|_{L^2(\Omega)}\\ &\le C\|z_{kh}(t)\|_{L^2(\Omega)} \|\Delta_h^{-1} z_{kh}(t)\|_{H_0^1(\Omega)},\end{aligned}$$

yielding $\|\Delta_h^{-1} z_{kh}(t)\|_{H_0^1(\Omega)} \le C\|z_{kh}(t)\|_{L^2(\Omega)}$ for almost all t. Using this estimate together with Lemma 6.5, we get

$$\mathcal{Z} \le Ch^3 \int_I \|\Delta v_k(t)\|_{L^2(\Omega)} \|z_{kh}(t)\|_{L^2(\Omega)}\, dt \le Ch^3 \|\Delta v_k\|_{L^1(I;L^2(\Omega))} \|z_{kh}\|_{L^\infty(I;L^2(\Omega))}.$$

Using Corollary 4.7, we finally obtain

$$\mathcal{Z} \le Ch^3 \ln\frac{T}{k} \|v_0\|_{L^2(\Omega)} \|z_{kh}(T)\|_{L^2(\Omega)} \le Ch^3 \ln\frac{T}{k} \|v_0\|_{L^2(\Omega)} \|\Delta_h^{-2}(P_h v_k - v_{kh})(T)\|_{L^2(\Omega)}.$$

Canceling $\|\Delta_h^{-2}(P_h v_k - v_{kh})(T)\|_{L^2(\Omega)}$ gives the result. ☐

Using this auxiliary result, we can prove the next lemmas estimating $R_h v_k - v_{kh}$.

Lemma 6.9. *Let $v_k \in X_k^r$ and $v_{kh} \in X^{r,2}_{k,h}$ be the semidiscrete and fully discrete solutions of (4) and (9), respectively for $v_0 \in L^2(\Omega)$. There exists a constant C independent of k,h, and T such that*

$$\|\left(R_h v_k - v_{kh}\right)(T)\|_{L^2(\Omega)} \le Ch^3 \left(\frac{1}{T^2} + \frac{1}{T^{\frac{3}{2}}}\right) \ln\frac{T}{k} \|v_0\|_{L^2(\Omega)}.$$

Proof. Let $y_{kh} \in X^{r,2}_{k,h}$ be the solution to a dual problem with $y_{kh}(T) = \left(R_h v_k - v_{kh}\right)(T)$, i.e $y_{kh} \in X^{r,2}_{k,h}$ satisfies

$$B(\varphi_{kh}, y_{kh}) = (\varphi_{kh}(T), (R_h v_k - v_{kh})(T)) \quad \text{for all} \quad \varphi_{kh} \in X^{r,2}_{k,h}.$$

To simplify notation, we define $\psi_{kh} := R_h v_k - v_{kh} \in X^{r,2}_{k,h}$. We introduce $\tilde{\psi}_{kh} \in X^{r,2}_{k,h}$ to be zero on $I_1 \cup \ldots \cup I_{\tilde{m}}$ and $\tilde{\psi}_{kh} = \psi_{kh}$ on $I_{\tilde{m}+1} \cup \ldots I_M$ for $\tilde{m}$ chosen such that $\frac{T}{2} \in I_{\tilde{m}}$. Analogously we define $\tilde{y}_{kh}$. Choosing $\tilde{\psi}_{kh}$ as test function in the definition of y_{kh} and transfering the cutoff from one argument to the other, by (11), we get

$$\begin{aligned}
\|\left(R_h v_k - v_{kh}\right)(T)\|^2_{L^2(\Omega)} &= B(\tilde{\psi}_{kh}, y_{kh}) \\
&= B(\psi_{kh}, \tilde{y}_{kh}) + \left(\psi^-_{kh,\tilde{m}}, y^+_{kh,\tilde{m}}\right) \\
&= B(R_h v_k - v_{kh}, \tilde{y}_{kh}) + \left((R_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, y^+_{kh,\tilde{m}}\right) \\
&= B(R_h v_k - v_k, \tilde{y}_{kh}) + \left((R_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, y^+_{kh,\tilde{m}}\right) \\
&= J_1 + J_2.
\end{aligned}$$

Here we also have used the Galerkin orthogonality (10) with respect to the bilinear form B. By the definition of the Ritz projection the terms $(\nabla(R_h v_k - v_k), \nabla \tilde{y}_{kh})_{I_m \times \Omega}$ vanish from the form B, such that the remaining terms in J_1 are

$$\begin{aligned}
J_1 &= -\sum_{m=\tilde{m}+1}^{M} (R_h v_k - v_k, \partial_t y_{kh})_{I_m \times \Omega} - \sum_{m=\tilde{m}+1}^{M} (R_h v^-_{k,m} - v^-_{k,m}, [y_{kh}]_m) - \left(R_h v^-_{k,\tilde{m}} - v^-_{k,\tilde{m}}, y^+_{kh,\tilde{m}}\right) \\
&\le \|R_h v_k - v_k\|_{L^\infty((t_{\tilde{m}-1},T);L^2(\Omega))} \left(\|\partial_t y_{kh}\|_{L^1(I;L^2(\Omega))} + \sum_{m=1}^{M} \|[y_{kh}]_m\|_{L^2(\Omega)} + \|y^+_{kh,\tilde{m}}\|_{L^2(\Omega)} \right),
\end{aligned}$$

where we used the dual form of $B(\cdot,\cdot)$ and Hölders inequality in space and time to estimate the terms. Applying Corollary 4.7, gives

$$\|\partial_t y_{kh}\|_{L^1(I;L^2(\Omega))} + \sum_{m=1}^{M} \|[y_{kh}]_m\|_{L^2(\Omega)} + \|y^+_{kh,\tilde{m}}\|_{L^2(\Omega)} \le C \ln \frac{T}{k} \|(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}.$$

Note that y_{kh} is a solution to a dual problem and we use $y_{kh}(T)$ as bound. Using the L^2 error estimate for the Ritz projection, together with the estimate (18) of Lemma 6.4, we obtain for any $t \in (t_{\tilde{m}-1}, T]$

$$\|(R_h v_k - v_k)(t)\|_{L^2(\Omega)} \le C h^3 \|v_k(t)\|_{H^3(\Omega)} \le C h^3 \|\Delta v_k(t)\|^{\frac{1}{2}}_{L^2(\Omega)} \|\Delta^2 v_k(t)\|^{\frac{1}{2}}_{L^2(\Omega)}.$$

Using the smoothing results of Lemma 4.8, we obtain

$$\begin{aligned}
\sup_{t \in (t_{\tilde{m}-1},T]} \|(R_h v_k - v_k)(t)\|_{L^2(\Omega))} &\le C h^3 \sup_{t \in (t_{\tilde{m}-1},T]} \|\Delta v_k(t)\|^{\frac{1}{2}}_{L^2(\Omega)} \|\Delta^2 v_k(t)\|^{\frac{1}{2}}_{L^2(\Omega)} \\
&\le C \frac{h^3}{t^{\frac{3}{2}}_{\tilde{m}}} \|v_0\|_{L^2(\Omega)} \le C \frac{h^3}{T^{\frac{3}{2}}} \|v_0\|_{L^2(\Omega)}.
\end{aligned} \tag{22}$$

In the last step, we used the estimate $\frac{1}{t_{\tilde{m}}} \le \frac{2}{T}$ which holds true, since $t_{\tilde{m}}$ was chosen such that $\frac{T}{2} \in I_{\tilde{m}}$, and thus, $\frac{T}{2} \le t_{\tilde{m}}$. Combining these results gives the proposed estimate for J_1:

$$J_1 \le C \frac{h^3}{T^{\frac{3}{2}}} \ln \frac{T}{k} \|v_0\|_{L^2(\Omega)} \|(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}.$$

To estimate J_2 we insert an artificial zero by adding and subtracting $v^-_{k,\tilde{m}}$,

$$J_2 = ((R_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, y^+_{kh,\tilde{m}}) = ((R_h v_{k,\tilde{m}} - v_{k,\tilde{m}})^-, y^+_{kh,\tilde{m}}) + ((v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, y^+_{kh,\tilde{m}}) := J_{21} + J_{22}.$$

The term J_{21} can be estimated with (22), the discrete smoothing result of Lemma 4.8 applied to $\|y^+_{kh,\tilde{m}}\|_{L^2(\Omega)}$ and the special choice of $\tilde{m}$:

$$\begin{aligned} J_{21} &\le \sup_{t\in(t_{\tilde{m}-1},T)} \|(R_h v_k - v_k)(t)\|_{L^2(\Omega))} \|y^+_{kh,\tilde{m}}\|_{L^2(\Omega)} \\ &\le C\frac{h^3}{t_{\tilde{m}}^{\frac{3}{2}}}\|v_0\|_{L^2(\Omega)}\|y_{kh}(T)\|_{L^2(\Omega)} \le C\frac{h^3}{T^{\frac{3}{2}}}\|v_0\|_{L^2(\Omega)}\|(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}. \end{aligned}$$

To estimate J_{22} we use Lemma 6.8 by using $t_{\tilde{m}}$ as artificial endtime. Here it is of importance, that the derived constant does not depend on the endtime, since we need to replace $t_{\tilde{m}}$ by T later. This can only be done, when the explicit dependence of the result of Lemma 6.8 on the endtime is known. We thus get after inserting the L^2 projection operator:

$$\begin{aligned} J_{22} &= ((P_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, y^+_{kh,\tilde{m}}) \\ &= (\Delta_h^{-2}(P_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, \Delta_h^2 y^+_{kh,\tilde{m}}) \\ &\le \|\Delta_h^{-2}(P_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-\|_{L^2(\Omega)}\|\Delta_h^2 y^+_{kh,\tilde{m}}\|_{L^2(\Omega)} \\ &\le C\ln\left(\frac{t_{\tilde{m}}}{k}\right)\|v_0\|_{L^2(\Omega)}\frac{h^3}{(T-t_{\tilde{m}})^2}\|(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}. \end{aligned}$$

In the last step, we have used Lemma 6.8 for $\|\Delta_h^{-2}(P_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-\|_{L^2(\Omega)}$ and the discrete smoothing result of Lemma 4.8 for $\|\Delta_h^2 y^+_{kh,\tilde{m}}\|_{L^2(\Omega)}$. Since y_{kh} is a solution to a backwards problem, we use $\frac{1}{T-t_{\tilde{m}}}$, instead of $\frac{1}{t_{\tilde{m}}}$. We now replace all occurrences of $t_{\tilde{m}}$ by T. As before, we use $\frac{T}{2} \in I_{\tilde{m}}$ yielding $t_{\tilde{m}} \le \frac{T}{2} + k$. For fine enough time discretizations (*i.e.* $\frac{T}{4} > k$) we have

$$T - t_{\tilde{m}} \ge T - \frac{T}{2} - k \ge T - \frac{T}{2} - \frac{T}{4} = \frac{T}{4},$$

thus giving $\frac{1}{T-t_{\tilde{m}}} \le C\frac{1}{T}$. To estimate the logarithmic term, we use the following consideration: Let $x \in \mathbb{R}$ such that $x > 2$. Then it holds $x + 1 \le x^2$. With the monotonicity of the logarithm, we obtain $\ln(x+1) \le \ln(x^2) = 2\ln(x)$. Applying this to the logarithmic term, while using $t_{\tilde{m}} \le \frac{T}{2} + k$ and $\frac{T}{2k} > 2$, yields

$$\ln\left(\frac{t_{\tilde{m}}}{k}\right) \le \ln\left(\frac{\frac{T}{2}+k}{k}\right) = \ln\left(\frac{T}{2k}+1\right) \le 2\ln\left(\frac{T}{2k}\right) \le 2\ln\left(\frac{T}{k}\right).$$

This gives a bound for J_{22}, depending on the final time T,

$$J_{22} \le C\frac{h^3}{T^2}\ln\frac{T}{k}\|v_0\|_{L^2(\Omega)}\|(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}.$$

Dividing all considered terms by $\|(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}$ gives the proposed estimate. ☐

We now show a similar result for the discrete Laplacian.

Lemma 6.10. *Let $v_k \in X_k^r$ and $v_{kh} \in X_{k,h}^{r,2}$ be the semidiscrete and fully discrete solutions of (4) and (9), respectively. Then there exists a constant independent of k, h, and T such that*

$$\|\Delta_h(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)} \le Ch^3\left(\frac{1}{T^3} + \frac{1}{T^{\frac{5}{2}}}\right)\ln\frac{T}{k}\|v_0\|_{L^2(\Omega)}.$$

Proof. Let $y_{kh} \in X^{r,2}_{k,h}$ be the solution to a dual problem with $y_{kh}(T) = \Delta_h\left(R_h v_k - v_{kh}\right)(T)$, i.e $y_{kh} \in X^{r,2}_{k,h}$ satisfies

$$B(\varphi_{kh}, y_{kh}) = (\varphi_{kh}(T), \Delta_h(R_h v_k - v_{kh})(T)) \quad \text{for all} \quad \varphi_{kh} \in X^{r,2}_{k,h}.$$

As in the previous lemma, in order to simplify notation, we define $\psi_{kh} := R_h v_k - v_{kh} \in X^{r,2}_{k,h}$. We introduce $\tilde{\psi}_{kh} \in X^{r,2}_{k,h}$ to be zero on $I_1 \cup \ldots \cup I_{\tilde{m}}$ and $\tilde{\psi}_{kh} = \psi_{kh}$ on $I_{\tilde{m}+1} \cup \ldots I_M$ for $\tilde{m}$ chosen such that $\frac{T}{2} \in I_{\tilde{m}}$. We define $\tilde{y}_{kh}$ analogously. Choosing $\Delta_h \tilde{\psi}_{kh}$ as test function in the definition of y_{kh}, and transfering the cutoff from the first argument of B to the second, by applying (11), we get

$$\begin{aligned}
\|\Delta_h\left(R_h v_k - v_{kh}\right)(T)\|^2_{L^2(\Omega)} &= B(\Delta_h \tilde{\psi}_{kh}, y_{kh}) \\
&= B(\tilde{\psi}_{kh}, \Delta_h y_{kh}) \\
&= B(\psi_{kh}, \Delta_h \tilde{y}_{kh}) + \left(\psi^-_{kh,\tilde{m}}, \Delta_h y^+_{kh,\tilde{m}}\right) \\
&= B(R_h v_k - v_{kh}, \Delta_h \tilde{y}_{kh}) + \left((R_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, \Delta_h y^+_{kh,\tilde{m}}\right) \\
&= B(R_h v_k - v_k, \Delta_h \tilde{y}_{kh}) + \left((R_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, \Delta_h y^+_{kh,\tilde{m}}\right) \\
&= J_1 + J_2.
\end{aligned}$$

Here we have used the Galerkin orthogonality (10) with respect to the bilinear form B. By the definition of the Ritz projection the terms $(\nabla(R_h v_k - v_k), \nabla\Delta_h \tilde{y}_{kh})_{I_m \times \Omega}$ vanish from the form B, such that the remaining terms of J_1 are

$$J_1 = \sum_{m=\tilde{m}+1}^{M} (\partial_t(R_h v_k - v_k), \Delta_h y_{kh})_{I_m \times \Omega} + \sum_{m=\tilde{m}+1}^{M} ([R_h v_k - v_k]_m, \Delta_h y^+_{kh,m}).$$

Applying Hölder's inequality in space and time gives

$$J_1 \leq \|\partial_t(R_h v_k - v_k)\|_{L^\infty((t_{\tilde{m}},T);L^2(\Omega))} \|\Delta_h y_{kh}\|_{L^1((t_{\tilde{m}},T);L^2(\Omega))} + \sum_{m=\tilde{m}+1}^{M} \|[R_h v_k - v_k]_m\|_{L^2(\Omega)} \|\Delta_h y^+_{kh,m}\|_{L^2(\Omega)}.$$

Introducing an artificial factor 1 as $k_m \cdot k_m^{-1}$ in the sum allows us to extract the term $\max_{\tilde{m}\leq m\leq M}\left\{k_m^{-1}\|[R_h v_k - v_k]_m\|_{L^2(\Omega)}\right\}$ out of the sum. This gives

$$\begin{aligned}
J_1 \leq &\max_{\tilde{m}\leq m\leq M} \|\partial_t(R_h v_k - v_k)\|_{L^\infty(I_m;L^2(\Omega))} \|\Delta_h y_{kh}\|_{L^1((t_{\tilde{m}},T);L^2(\Omega))} \\
&+ \max_{\tilde{m}\leq m\leq M}\left\{k_m^{-1}\|[R_h v_k - v_k]_m\|_{L^2(\Omega)}\right\} \left(\sum_{m=\tilde{m}+1}^{M} k_m \|\Delta_h y^+_{kh,m}\|_{L^2(\Omega)}\right).
\end{aligned}$$

Using Corollary 4.7 and the L^2 error estimate for the Ritz projection for the other terms, gives the following estimate,

$$J_1 \leq Ch^3 \left(\max_{\tilde{m}\leq m\leq M} \|\partial_t v_k\|_{L^\infty(I_m;H^3(\Omega))} + \max_{\tilde{m}\leq m\leq M}\left\{k_m^{-1}\|[v_k]_m\|_{H^3(\Omega))}\right\}\right) \ln\frac{T}{k} \|\Delta_h(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}.$$

Similar to the previous lemma, by the estimate (18) of Lemma 6.4 and Remark 6.1, yielding $\partial_t \Delta v_k(t) \in H^1_0(\Omega)$ for $t \in (t_{\tilde{m}-1}, T]$, we obtain

$$\|\partial_t v_k(t)\|_{H^3(\Omega)} \leq C\|\partial_t \Delta v_k(t)\|^{\frac{1}{2}}_{L^2(\Omega)} \|\partial_t \Delta^2 v_k(t)\|^{\frac{1}{2}}_{L^2(\Omega)}.$$

Taking the supremum over $(t_{\tilde{m}-1}, T]$ and using Lemma 4.8 then yields

$$\begin{aligned}\max_{\tilde{m}\le m\le M} \|\partial_t v_k(t)\|_{L^\infty(I_m;H^3(\Omega))} &\le C \max_{\tilde{m}\le m\le M} \|\partial_t \Delta v_k(t)\|^{\frac{1}{2}}_{L^\infty(I_m;L^2(\Omega))} \left\|\partial_t \Delta^2 v_k(t)\right\|^{\frac{1}{2}}_{L^\infty(I_m;L^2(\Omega))}\\ &\le C\frac{1}{t_{\tilde{m}}^{\frac{5}{2}}}\|v_0\|_{L^2(\Omega)} \le C\frac{1}{T^{\frac{5}{2}}}\|v_0\|_{L^2(\Omega)}.\end{aligned}$$

Applying the same arguments, using $\Delta v_k(t) \in H_0^1(\Omega)$ for $t \in (t_{\tilde{m}-1}, T)$, and Lemma 4.8 gives

$$\max_{\tilde{m}\le m\le M} \left\{k_m^{-1}\|[v_k]_m\|_{H^3(\Omega)}\right\} \le \max_{\tilde{m}\le m\le M}\left\{k_m^{-1}\left(\|\Delta[v_k]_m\|_{L^2(\Omega)}\|\Delta^2[v_k]_m\|_{L^2(\Omega)}\right)^{\frac{1}{2}}\right\} \le C\frac{1}{t_{\tilde{m}}^{\frac{5}{2}}}\|v_0\|_{L^2(\Omega)} \le C\frac{1}{T^{\frac{5}{2}}}\|v_0\|_{L^2(\Omega)}.$$

Summarizing all above results yields the final bound for J_1:

$$J_1 \le Ch^3\frac{1}{T^{\frac{5}{2}}}\ln\frac{T}{k}\|v_0\|_{L^2(\Omega)}\|\Delta_h(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}.$$

To estimate J_2 we insert an artificial zero like before by adding and subtracting $v^-_{k,\tilde{m}}$:

$$J_2 = ((R_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, \Delta_h y^+_{kh,\tilde{m}}) = ((R_h v_{k,\tilde{m}} - v_{k,\tilde{m}})^-, \Delta_h y^+_{kh,\tilde{m}}) + ((v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, \Delta_h y^+_{kh,\tilde{m}}) := J_{21} + J_{22}.$$

The term J_{21} can be estimated similarly to the previous lemma, applying (22), the discrete smoothing result of Lemma 4.8 for $\|\Delta_h y^+_{kh,\tilde{m}}\|_{L^2(\Omega)}$ and using the special choice of $\tilde{m}$:

$$\begin{aligned}J_{21} &\le \sup_{t\in(t_{\tilde{m}-1},T)} \|(R_h v_k - v_k)(t)\|_{L^2(\Omega)}\|\Delta_h y^+_{kh,\tilde{m}}\|_{L^2(\Omega)}\\ &\le C\frac{h^3}{t_{\tilde{m}}^{\frac{3}{2}}}\|v_0\|_{L^2(\Omega)}\frac{1}{T - t_{\tilde{m}}}\|y_{kh}(T)\|_{L^2(\Omega)}\\ &\le C\frac{h^3}{T^{\frac{5}{2}}}\|v_0\|_{L^2(\Omega)}\|\Delta_h(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}.\end{aligned}$$

We estimate J_{22} by replacing $v_{k,\tilde{m}}$ with its L^2-projection:

$$\begin{aligned}J_{22} &= ((P_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, \Delta_h y^+_{kh,\tilde{m}})\\ &= (\Delta_h^{-2}(P_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-, \Delta_h^3 y^+_{kh,\tilde{m}})\\ &\le \|\Delta_h^{-2}(P_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-\|_{L^2(\Omega)}\|\Delta_h^3 y^+_{kh,\tilde{m}}\|_{L^2(\Omega)}\\ &\le C\frac{h^3}{(T - t_{\tilde{m}})^3}\ln\left(\frac{t_{\tilde{m}}}{k}\right)\|v_0\|_{L^2(\Omega)}\|\Delta_h(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}.\end{aligned}$$

In the last step, we have used Lemma 6.8 for $\|\Delta_h^{-2}(P_h v_{k,\tilde{m}} - v_{kh,\tilde{m}})^-\|_{L^2(\Omega)}$ and the discrete smoothing result of Lemma 4.8 for $\|\Delta_h^3 y^+_{kh,\tilde{m}}\|_{L^2(\Omega)}$. Since y_{kh} is the solution to a problem backward in time, we use $\frac{1}{T-t_{\tilde{m}}}$ instead of $\frac{1}{t_{\tilde{m}}}$ in the application of this result. Analogously to the previous lemma, we can replace the terms involving $t_{\tilde{m}}$ by ones dependent only of T because of the special choice of $t_{\tilde{m}}$, thus giving the final bound for J_{22},

$$J_{22} \le C\frac{h^3}{T^3}\ln\frac{T}{k}\|v_0\|_{L^2(\Omega)}\|\Delta_h(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}.$$

Dividing all considered terms by $\|\Delta_h(R_h v_k - v_{kh})(T)\|_{L^2(\Omega)}$ gives the proposed estimate. □

Combining Lemmas 6.9 and 6.10 with the discrete Gagliardo-Nirenberg inequality (15) gives the following result:

Corollary 6.11. *Let $v_k \in X_k^r$ and $v_{kh} \in X_{k,h}^{r,2}$ be the semidiscrete and fully discrete solutions of of* (1) *with $v_0 \in L^2(\Omega)$. Then there exists a constant independent of k and h such that*

$$\|(R_h v_k - v_{kh})(T)\|_{L^\infty(\Omega)} \leq C h^3 \ln \frac{T}{k} \left(\frac{1}{T^3} + \frac{1}{T^{\frac{5}{2}}} \right)^{\frac{N}{4}} \left(\frac{1}{T^2} + \frac{1}{T^{\frac{3}{2}}} \right)^{\left(1 - \frac{N}{4}\right)} \|v_0\|_{L^2(\Omega)}.$$

This result now allows us to estimate the final term T_3 of (19) and thus proves Theorem 6.6.

6.2. Estimates for $(v - v_{kh})(T)$ with $v_0 \in \mathcal{M}(\Omega)$

Now that we have established Theorem 6.6 for $v_0 \in L^2(\Omega)$, following exactly the proof of Theorem 5.3, and using the Assumption 6.2, we can establish

Theorem 6.12. *Let Ω_0 be a subdomain with $\overline{\Omega}_0 \subset \Omega$, $v_0 \in \mathcal{M}(\Omega)$ with $\operatorname{supp} v_0 \subset \Omega_0$. Let v and $v_{kh} \in X_{k,h}^{r,2}$ satisfy* (1) *and* (9)*, respectively. In addition, let Ω be such that Assumption 6.2 holds.*

$$\|(v - v_{kh})(T)\|_{L^\infty(\Omega_0)} \leq C(T, \Omega_0) \left(\ell_k h^3 + k^{2r+1} \right) \|v_0\|_{\mathcal{M}(\Omega)},$$

where $\ell_k = \ln \frac{T}{k}$, $d = \operatorname{dist}(\Omega_0, \partial\Omega)$ and $C(T, d)$ is a constant depending on T and d.

References

[1] R.A. Adams and J. Fournier, Cone conditions and properties of Sobolev spaces. *J. Math. Anal. Appl.* **61** (1977) 713–734.

[2] R.A. Adams and J. Fournier, Sobolev spaces, 2nd edition. In Vol. 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam (2003).

[3] A. Ashyralyev and P.E. Sobolevskii, Well-Posedness of Parabolic Difference Equations. Birkhäuser Basel (1994).

[4] S.C. Brenner and L.R. Scott, The mathematical theory of finite element methods. 3rd edition. In Vol. 15 of *Texts in Applied Mathematics*. Springer, New York (2008).

[5] E. Casas, B. Vexler and E. Zuazua, Sparse initial data identification for parabolic PDE and its finite element approximations. *Math. Control Relat. Fields* **5** (2015) 377–399.

[6] I. Drelichman, R.G. Durán and I. Ojea, A weighted setting for the numerical approximation of the poisson problem with singular sources. *SIAM J. Numer. Anal.* **58** (2020) 590–606.

[7] K. Eriksson, C. Johnson and S. Larsson, Adaptive finite element methods for parabolic problems. VI. Analytic semigroups. *SIAM J. Numer. Anal.* **35** (1998) 1315–1325.

[8] L.C. Evans, Partial differential equations, 2nd edition. In no. 19 *Graduate Studies in mathematics, American Mathematical Society* (2010).

[9] A. Hansbo, Strong stability and non-smooth data error estimates for discretizations of linear parabolic problems. *BIT Numer. Math.* **42** (2002) 351–379.

[10] T. Hell, A. Ostermann and M. Sandbichler, Modification of dimension-splitting methods—overcoming the order reduction due to corner singularities. *IMA J. Numer. Anal.* **35** (2015) 1078–1091.

[11] P. Lesaint and P. Raviart, On a finite element method for solving the neutron transport equation. In *Mathematical Aspects of Finite Elements in Partial Differential Equations*. Academic Press (1974) 89–123.

[12] D. Leykekhman and B. Vexler, Pointwise best approximation results for Galerkin finite element solutions of parabolic problems. *SIAM J. Numer. Anal.* **54** (2016) 1365–1384.

[13] D. Leykekhman and B. Vexler, Discrete maximal parabolic regularity for Galerkin finite element methods. *Numer. Math.* **135** (2017) 923–952.

[14] D. Leykekhman, B. Vexler and J. Wagner, Numerical analysis of sparse initial data identification for parabolic problems with pointwise final time observations, in preparation, (2023).

[15] D. Leykekhman, B. Vexler and D. Walter, Numerical analysis of sparse initial data identification for parabolic problems. *ESAIM Math. Model. Numer. Anal.* **54** (2020) 1139–1180.

[16] D. Meidner, R. Rannacher and B. Vexler, A priori error estimates for finite element discretizations of parabolic optimization problems with pointwise state constraints in time. *SIAM J. Control Optim.* **49** (2011) 1961–1997.

[17] E.B. Saff and R.S. Varga, On the zeros and poles of Padé approximants to eˆz. *Numer. Math.* **25** (1975) 1–14.

[18] A.H. Schatz, V.C. Thomée and L.B. Wahlbin, Maximum norm stability and error estimates in parabolic finite element equations. *Comm. Pure Appl. Math.* **33** (1980) 265–304.

[19] A.H. Schatz and L.B. Wahlbin, Interior maximum norm estimates for finite element methods. *Math. Comput.* **31** (1977) 414–442.

[20] V. Thomée, Galerkin finite element methods for parabolic problems, 2nd edition. In Vol. 25 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin (2006).

[21] G. Wanner, E. Hairer and S.P. Nørsett, Order stars and stability theorems. *BIT Numer. Math.* **18** (1978) 475–489.