ESAIM: M2AN 59 (2025) 449–485
https://doi.org/10.1051/m2an/2024073
ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

ERROR ESTIMATES FOR DISCONTINUOUS GALERKIN TIME-STEPPING SCHEMES FOR THE PARABOLIC p-LAPLACIAN: A QUASI-NORM APPROACH

Konstantinos Chrysafinos[1,2,*] and Panagiotis Paraschis[1]

Abstract. Error estimates for arbitrary order fully-discrete schemes for the parabolic p-Laplacian are considered. The schemes combine the discontinuous Galerkin time-stepping approach for the temporal discretization with classical conforming finite elements in space. In particular, a symmetric – Céa Lemma type – error estimate is established for a suitable quasi-norm, under minimal regularity assumptions on the data. The above estimate leads to error bounds of arbitrary order in space and time provided that the necessary regularity is present, without imposing any restrictions between the temporal and spatial discretization parameters. The symmetric structure of the estimate also leads to various error estimates at partition points as well as for the natural energy $L^p(I; W^{1,p}(\Omega))$ norm. Furthermore, an unconditional $L^\infty(I; L^2(\Omega))$ stability and error estimate is proved under minimal regularity assumptions, as well as an optimal $L^\infty(I; L^2(\Omega))$ error estimate under a suitable restriction between the temporal and spatial discretization parameters and additional regularity of the solution.

Mathematics Subject Classification. 65M15, 65M60, 35K92.

Received October 26, 2023. Accepted September 24, 2024.

1. Introduction

We consider a quasilinear partial differential equation with p-Laplacian structure, *i.e.*,

$$\begin{cases} \partial_t u - \operatorname{div}(|\nabla u|^{p-2}\nabla u) = f & \text{in } \Omega \times I, \\ u = 0 & \text{on } \partial\Omega \times I, \\ u(\cdot,0) = u_0 & \text{in } \Omega. \end{cases} \tag{1.1}$$

Here f and u_0 denote the forcing term and the initial data respectively, while the parameter $p \in (1,\infty)$, defines the diffusion properties of the underlying partial differential equation (PDE) and it is the critical parameter involved. In the above context, we consider $\Omega \subset \mathbb{R}^d$ (with $d = 1,2,3$), an open bounded set with suitably piecewise smooth boundary $\partial\Omega$ and a time interval $I = (0,T]$, for a fixed $T > 0$. Equations with p-Laplacian structure serve as prototype models in various nonlinear diffusion problems arising in mathematical biology,

Keywords and phrases. Error estimates, discontinuous-in-time Galerkin schemes, Quasi-norm.

[1] Department of Mathematics, School of Mathematical and Physical Sciences, National Technical University of Athens, Zografou 15780, Greece.

[2] IACM, FORTH, 70013 Heraklion, Crete, Greece.

*Corresponding author: chrysafinos@math.ntua.gr

chemical kinetics, as well as non-Newtonian fluids mechanics [2, 13, 22, 28, 29, 31]. It is worth noting that the value of p critically determines the behaviour of the solution of the underlying PDE. In particular, one should distinguish two cases with respect to p, $p \in (2,\infty)$, and $p \in (1,2)$. The case $p = 2$ is omitted, since it corresponds to the linear heat equation.

1.1. A brief description of the main results

The main aim of this work is the development of error estimates for a general class of fully-discrete schemes that maintain an "almost" symmetric structure, resembling the classical Céa's Lemma estimate of the standard finite element discretization of linear elliptic PDEs. To achieve this, we consider fully-discrete schemes (of any order in time and space) based on discontinuous-in-time Galerkin (dG) scheme approach for the temporal discretization, combined with standard conforming (in space) finite elements for spatial discretization. Such schemes are known to inherit stability properties of the underlying nonlinear PDE, without imposing additional regularity assumptions on the solution (see, *e.g.*, [10]), and provide the natural framework for time adaptivity since they allow the use of different finite element subspaces every other or every few time steps. Indeed, for a time partition $0 = t_0 < t_1 < \ldots < t_N = T$, of I, we consider a family of conforming finite element spaces $\{V_h^n\}_{n=0}^N$ and we denote by

$$u_h \in \mathcal{U}_h := \Big\{ w_h \in L^p(I;W_0^{1,p}(\Omega)) : \ w_h|_{I_n} \in \mathcal{P}_k(I_n;V_h^n) \ , \, n = 1,\ldots,N \Big\}$$

the fully-discrete solution computed by the discontinuous-in-time Galerkin (dG(k)) scheme (the precise statement is given in Sect. 3). Here, we denote by $\mathcal{P}_k(I_n;V_h^n)$ polynomials of degree $k \geq 0$, in-time taking values in V_h^n, in each time interval $I_n = (t_{n-1},t_n]$. We observe that only $L^p(I,W_0^{1,p}(\Omega))$ regularity is assumed, and that $w_h \in \mathcal{U}_h$ is discontinuous-in-time at partition points. For the above class of schemes, first, in Section 3, following the techniques of [9, 10], it is proved that these schemes are unconditionally stable in $L^\infty(I;L^2(\Omega)) \cap L^p(I;W^{1,p}(\Omega))$, without imposing any restriction between the size of the associated temporal and spatial discretization parameters τ and h, respectively. The main results, regarding error estimates, are presented in Section 4. The key ingredient of our error analysis is the proof of an almost best-approximation error estimate, posed in a suitable quasi-norm setting. To this end, we denote by $\|w\|^2_{L^2(I;p,\psi)} := \int_I \|w\|^2_{(p,\psi)}\,\mathrm{d}t = \int_I\int_\Omega (|\nabla\psi| + |\nabla w|)^{p-2}|\nabla w|^2\,\mathrm{d}x\,\mathrm{d}t$, the associated quasi-norm, defined for any weight $\psi \in L^p(I;W^{1,p}(\Omega))$. The error is decomposed as, $e := u_h - u := e_h + \hat{e}$, where $\hat{e} = P_h^{\mathrm{loc}}u - u$, with P_h^{loc} denoting the natural fully-discrete projector $P_h^{\mathrm{loc}} : C(\bar{I};L^2(\Omega)) \to \mathcal{U}_h$ associated to the discontinuous-in-time Galerkin scheme that exhibits best approximation properties (see, *e.g.*, [33]). Our main result is an "almost symmetric" estimate (a precise statement is given in Sect. 4), of the form:

$$\max_{1\le n\le N} \|e(t_n)\|^2_{L^2(\Omega)} + \lambda_p \|e\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)} \le C\Bigg(\sum_{n=0}^{N-1} \|(\mathcal{I}-P_h^n)u(t_n)\|^2_{L^2(\Omega)} + \max_{1\le n\le N} \|(\mathcal{I}-P_h^n)u(t_n)\|^2_{L^2(\Omega)} + \|\hat{e}\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)}\Bigg). \tag{1.2}$$

In the above P_h^n stands for the standard L^2-projection on V_h^n, and λ_p denotes the coercivity constant, depending on p, associated with the p-Laplacian operator, and $\mathcal{I}$ denotes the identity operator. Observe that the first term of the left-hand side measures the error due to the possible change of the finite element spaces and it is measured using the standard L^2 norms, involving L^2-type projections. This term can be omitted when $V_h^n = V_h^0$, for $n = 1,\cdots,N$, to recover a Céa's type approximation result. The last term of (1.2) measures the approximation error on the quasi-norm and it is related to best-approximation type errors in $L^p(0,T;W^{1,p}(\Omega))$ norm. We note that the estimate is derived under minimal regularity assumptions without imposing any restriction between τ and h.

The use of quasi-norm approaches for the backward Euler scheme has been developed before in the pioneering works [3, 12]. It should be noted that our choice of quasi-norm differs from the above mentioned works, in the sense that it uses the natural discontinuous-in-time Galerkin projection, as an additional weight instead of u. It turns out that this is crucial since it allows to use the space-time Galerkin orthogonality, and leads to the almost symmetric estimate (1.2), which can be viewed as an analogue of the classical Céa's Lemma estimate. In addition, we emphasize that our quasi-norm based on $P_h^{\text{loc}}u$ weight, recovers the same convergence rates compared to the classical quasi-norm $\|\cdot\|_{L^2(I;p,u)}$ based on weight u. The estimate (1.2) leads to various estimates on the natural energy norm $\|e\|_{L^p(I;W^{1,p}(\Omega))}$ and quite surprising it is possible to derive estimates in $L^\infty(I;L^2(\Omega))$ by performing a "boot-strap" argument based on the technique of [9, 10]. Our analysis is also applicable when high order schemes are being used for the temporal and spatial discretizations. Even though higher regularity properties for the solution of (1.1) are not always observed, we may exploit the structure of high-order discontinuous-in-time Galerkin schemes in conjunction with classical piecewise linear finite elements discretization in space, to compute approximate solutions based on coarse time-stepping approaches. Coarse time-stepping approaches require the solution of fewer "nonlinear" systems associated to the nonlinear structure of the operator, which in certain scenarios is proven to be cost effective. Our computational examples, given in Section 6, verify the applicability of such "coarse" time-stepping approach.

To summarize, the discontinuous-in-time Galerkin framework, combined with this particular choice of quasi-norm structure, apart from being the key asset for developing estimate (1.2), distinguishes our work from previous approaches. To our best knowledge the estimate (1.2) is the first estimate of such type for the underlying PDE and it is applicable in a variety of situation depending on the regularity of the solution.

1.2. Related results

The numerical analysis of the parabolic p-Laplacian is considered in the pioneering works by Barrett and Liu [3], where error estimates for the discretization with the backward Euler method in time and linear finite elements in space are established, by Diening *et al.* [12] and by Berselli and Růžička [4], where the discontinuous piecewise constant-in-time scheme together with linear finite elements in space has been applied to general parabolic systems with p-structure; including the p-Laplacian. In the two aforementioned papers, optimal quasi-norm error estimates have been obtained, under restrictions between the time and space discretization parameters. Those restrictions have been removed by Breit *et al.* [7] for the parabolic p-Laplacian, discretized with a backward Euler scheme in time and conforming finite elements in space. In the latter paper, quasi-norm error estimates have been derived with suitable fractional differentiability assumptions for the exact solution and its gradient, in the context of Nikolskiĭ spaces. Several numerical experiments have been presented, for different examples with low regularity solutions. Moreover, Berselli *et al.* [5] have obtained optimal quasi-norm error estimates for an implicit-explicit Euler time discretization along with conforming finite elements in space, for the incompressible p-Navier–Stokes equations, *i.e.*, the Navier–Stokes equations with the p-Laplacian instead of the Laplacian. The corresponding p-Stokes equations have been also analyzed by Eckstein and Růžička [15]. In the latter paper, quasi-norm error estimates have been obtained for a fully-discrete scheme, consisting of the backward Euler method in time and conforming finite elements in space.

A high order fully-discrete scheme for the parabolic p-Laplacian is analyzed by Hansen [21], where algebraically stable Runge–Kutta methods in time with standard conforming finite elements in space have been applied to nonilnear parabolic problems, that include the parabolic p-Laplacian with $p \in (2,\infty)$. Error estimates with suboptimal rates of convergence have been derived, with respect to the L^2-norm at the time partition points, for smooth solutions. An abstract nonlinear parabolic equation with coercive and monotone (up to a shift) operator in Gelfand triples, that includes the p-Laplacian, has been discretized in time by Emmrich and Thalhammer [18] with implicit Runge–Kutta schemes. Convergence in the weak and weak* senses has been established and suboptimal norm semi-discrete (in time) error estimates for $p \in (2,\infty)$, have been derived. In the same framework, Emmrich [16, 17] has established similar semi-discrete error estimates for a θ-scheme and a discontinuous Galerkin time-stepping scheme, respectively when $p \in (2,\infty)$.

For computational studies regarding a local discontinuous Galerkin scheme in space, together with diagonally implicit Runge–Kutta time discretization we refer the reader to Kröner *et al.* [24,25] for the parabolic p-Laplacian and the p-Navier–Stokes equations, respectively. In addition, Toulopoulos [34] has applied the interior penalty discontinuous-in space-Galerkin schemes in space along with diagonally implicit Runge–Kutta methods in time to the parabolic p-Laplacian equation proving error estimates for semi-discrete in space approximations along with numerical experiments for respective fully-discrete schemes.

A backward Euler/finite element scheme with semi-discrete-in-time error estimates has been considered by Emmrich and Wróblewska-Kamińska [19] for general quasilinear parabolic equations, including the parabolic p-Laplacian. Finally, for an *a posteriori* error estimate for the backward Euler/finite element discretization of the p-Laplacian, we refer the reader to the work of Kreuzer [23].

1.3. Outline

In Section 2, we set the notation and recall some preliminary results that are being applied throughout the paper. In Section 3, we define the fully-discrete scheme and provide stability estimates. In Section 4, we derive symmetric quasi-norm error estimates and the corresponding convergence results, under additional regularity. In Section 5, we develop estimates in $L^\infty(I;L^2(\Omega))$ and their respective symmetric error estimates. Finally, in Section 6, we present the respective numerical experiments.

2. Preliminaries

2.1. Notation

Let $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$ be an open, bounded set. We assume that the boundary $\partial\Omega$ is sufficiently piecewise smooth for our purposes. For all $s \in [1, \infty]$, $l \geq 0$, we denote by $L^s(\Omega)$, $W^{l,s}(\Omega)$ the standard Lebesgue and Sobolev spaces endowed with norms $\|\cdot\|_{L^s(\Omega)}$, and $\|\cdot\|_{W^{l,s}(\Omega)}$ respectively. The associated semi-norm, with $l = 1$, is denoted by $\|\nabla\cdot\|_{L^s(\Omega)}$. In the special case $l = 0$, we have that $L^s(\Omega) := W^{0,s}(\Omega)$. We also use the notation $(\cdot,\cdot)$ for the inner product in $L^2(\Omega)$. For $s = 2$, we denote by $H^l(\Omega) := W^{l,2}(\Omega)$, with norm $\|\cdot\|_{H^l(\Omega)}$. Moreover, $W_0^{1,s}(\Omega)$, and $H_0^1(\Omega)$ denote the spaces of functions in $W^{1,s}(\Omega)$ and $H^1(\Omega)$ with zero traces on the boundary $\partial\Omega$ respectively. Note that the semi-norm $\|\nabla\cdot\|_{L^s(\Omega)}$ becomes a norm in $W_0^{1,s}(\Omega)$ and it is equivalent to the $\|\cdot\|_{W^{1,s}(\Omega)}$-norm. For $s \in (1,\infty)$, we define by $W^{-1,s'}(\Omega)$ the dual space of $W_0^{1,s}(\Omega)$, where $s' = s/(s-1)$. For $s = 2$, we adopt the notation $W^{-1,2}(\Omega) = H^{-1}(\Omega)$. Their norms and associated duality pairings are denoted by $\|\cdot\|_{W^{-1,s'}(\Omega)}$, $\|\cdot\|_{H^{-1}(\Omega)}$, and $\langle\cdot,\cdot\rangle_{W^{-1,s'}(\Omega),W_0^{1,s}(\Omega)}$, $\langle\cdot,\cdot\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}$ respectively, which will be frequently abbreviated to $\langle\cdot,\cdot\rangle$, when there is no danger of confusion.

Furthermore, for an interval $I \subset \mathbb{R}$ and for a given Banach space X, we denote by $L^s(I;X)$ the standard Bochner space with norm $\|\cdot\|_{L^s(I;X)}$. The space of functions in $L^s(I;X)$, whose distributional derivatives lie in $L^s(I;X)$ is denoted by $W^{1,s}(I;X)$ and its norm by $\|\cdot\|_{W^{1,s}(I;X)}$. If $s = 2$, we denote the latter space by $H^1(I;X)$. We also denote by $C(\bar{I};X)$ the space of continuous functions $u : \bar{I} \to X$, with norm $\|\cdot\|_{C(\bar{I};X)}$. We refer the reader to [20] for more information. For a non-negative integer k, we denote by $\mathcal{P}_k(I;X)$ the space of polynomial functions of degree at most k, which are defined in I and take values in X. We also abbreviate $\mathcal{P}_k(I;\mathbb{R})$ by $\mathcal{P}_k(I)$. In all the above spaces, we denote the identity operator by $\mathcal{I}$.

For $p \in (1,\infty)$, we define the p-Laplacian operator $A : W_0^{1,p}(\Omega) \to W^{-1,p'}(\Omega)$ by

$$\langle Au, v\rangle = (|\nabla u|^{p-2}\nabla u, \nabla v) := \int_\Omega |\nabla u|^{p-2}\nabla u \cdot \nabla v \,\mathrm{d}x \quad \forall u, v \in W_0^{1,p}(\Omega). \tag{2.1}$$

Another useful tool that is necessary for our purposes is a quasi-norm in $W_0^{1,p}(\Omega)$. For any $p > 1$, and $\psi \in W_0^{1,p}(\Omega)$, we define the quasi-norm $\|\cdot\|_{(p,\psi)}$ by

$$\|w\|_{(p,\psi)}^2 = \int_\Omega (|\nabla\psi| + |\nabla w|)^{p-2}|\nabla w|^2 \,\mathrm{d}x \quad \forall w \in W_0^{1,p}(\Omega). \tag{2.2}$$

The corresponding space-time quasi-norm, with $\psi, w \in L^p(I; W_0^{1,p}(\Omega))$, is defined by

$$\|w\|^2_{L^2(I;p,\psi)} := \int_I \|w\|^2_{(p,\psi)} \,\mathrm{d}t = \int_I \int_\Omega (|\nabla\psi| + |\nabla w|)^{p-2} |\nabla w|^2 \,\mathrm{d}x\,\mathrm{d}t. \tag{2.3}$$

2.2. Continuity and monotonicity properties

The following fundamental technical lemma has been proved in [3].

Lemma 2.1 ([3], Lem. 2.2)**.** *For all $p \in (1,\infty)$ and $\delta \geq 0$ there exist positive constants $C_1(p,\delta)$ and $C_2(p,\delta)$ such that for all $y, z \in \mathbb{R}^d$,*

$$\big||y|^{p-2}y - |z|^{p-2}z\big| \leq C_1(p,\delta)|y-z|^{1-\delta}(|y|+|z|)^{p-2+\delta}, \tag{2.4}$$

$$\big(|y|^{p-2}y - |z|^{p-2}z\big)\cdot(y-z) \geq C_2(p,\delta)|y-z|^{2+\delta}(|y|+|z|)^{p-2-\delta}. \tag{2.5}$$

Furthermore, we collect the following generalised Young inequalities.

Lemma 2.2. *Let $\varepsilon > 0$, $a, \sigma_1, \sigma_2 \geq 0$, and $p, p' \in (1,\infty)$, with $1/p + 1/p' = 1$. Then, with $\Lambda_\varepsilon = \max(1, (4\varepsilon)^{-1})$, there holds,*

$$\sigma_1\sigma_2 \leq \frac{\varepsilon^p}{p}\sigma_1^p + \frac{1}{p'\varepsilon^{p'}}\sigma_2^{p'}, \tag{2.6}$$

$$(a+\sigma_1)^{p-2}\sigma_1\sigma_2 \leq \varepsilon^{-\frac{1}{p-1}}(a+\sigma_1)^{p-2}\sigma_1^2 + \varepsilon(a+\sigma_2)^{p-2}\sigma_2^2 \qquad \text{for all } p \in (1,2), \tag{2.7}$$

$$(a+\sigma_1)^{p-2}\sigma_1\sigma_2 \leq \Lambda_\varepsilon(a+\sigma_1)^{p-2}\sigma_1^2 + \varepsilon(a+\sigma_2)^{p-2}\sigma_2^2 \qquad \text{for all } p \in [2,\infty). \tag{2.8}$$

Proof. The proof of (2.6) is standard. The case $p \in (1,2)$ has been proved in Lemma 5.2 of [27]. The case of $p \in [2,\infty)$ is simpler, but for completeness we present the proof. For $p = 2$, (2.8) is the standard Young inequality, so it suffices to prove (2.8) for $p \in (2,\infty)$. Without loss of generality we can assume that $\sigma_1, \sigma_2 > 0$. Clearly, if $\sigma_1 > \sigma_2$ we have

$$(a+\sigma_1)^{p-2}\sigma_1\sigma_2 \leq (a+\sigma_1)^{p-2}\sigma_1^2 < (a+\sigma_1)^{p-2}\sigma_1^2 + \varepsilon(a+\sigma_2)^{p-2}\sigma_2^2 \qquad \forall\varepsilon > 0.$$

If $\sigma_1 \leq \sigma_2$, we have $(a+\sigma_1)^{(p-2)/2} \leq (a+\sigma_2)^{(p-2)/2}$. Hence, there holds $\forall\varepsilon > 0$,

$$(a+\sigma_1)^{p-2}\sigma_1\sigma_2 \leq (a+\sigma_1)^{\frac{p-2}{2}}\sigma_1(a+\sigma_2)^{\frac{p-2}{2}}\sigma_2 \leq \frac{1}{4\varepsilon}(a+\sigma_1)^{p-2}\sigma_1^2 + \varepsilon(a+\sigma_2)^{p-2}\sigma_2^2,$$

which implies the desired estimate. □

Moreover, the following inequality has been established.

Lemma 2.3 ([26], Lem. 2.4)**.** *For $p \in (1,\infty)$, $\sigma_1, \sigma_2 \in \mathbb{R}^d$ and $a \geq 0$, it holds,*

$$(a + |\sigma_1+\sigma_2|)^{p-2}|\sigma_1+\sigma_2|^2 \leq \max(2, 2^{p-1})\big((a+|\sigma_1|)^{p-2}|\sigma_1|^2 + (a+|\sigma_2|)^{p-2}|\sigma_2|^2\big). \tag{2.9}$$

Lemma 2.3 implies the following triangle-type inequality for the quasi-norm.

Lemma 2.4. *Let $p \in (1,\infty)$ and $w_1, w_2, \psi \in W_0^{1,p}(\Omega)$. Then, we have*

$$\|w_1 + w_2\|^2_{(p,\psi)} \leq \max(2, 2^{p-1})\Big(\|w_1\|^2_{(p,\psi)} + \|w_2\|^2_{(p,\psi)}\Big). \tag{2.10}$$

In addition, the quasi-norm defined in (2.2) satisfies the following quasi-equivalence relations with the classical $W^{1,p}(\Omega)$ semi-norm.

Lemma 2.5. *Let $p \in (1,\infty)$, $\psi \in W^{1,p}(\Omega)$. Then, there exist positive constants $\hat{C}_1$ and $\hat{C}_2$, depending only upon p, such that for all $w \in W^{1,p}(\Omega)$,*

(i) *if $p \in (1,2)$, then,*

$$\|w\|^2_{(p,\psi)} \leq \|\nabla w\|^p_{L^p(\Omega)}, \ \|\nabla w\|^2_{L^p(\Omega)} \leq \hat{C}_1\big(\|\nabla \psi\|_{L^p(\Omega)} + \|\nabla w\|_{L^p(\Omega)}\big)^{2-p}\|w\|^2_{(p,\psi)}, \tag{2.11}$$

(ii) *if $p \in (2,\infty)$, then,*

$$\|\nabla w\|^p_{L^p(\Omega)} \leq \|w\|^2_{(p,\psi)} \leq \hat{C}_2\big(\|\nabla \psi\|_{L^p(\Omega)} + \|\nabla w\|_{L^p(\Omega)}\big)^{p-2}\|\nabla w\|^2_{L^p(\Omega)}. \tag{2.12}$$

Here $\hat{C}_1 := 2^{(p-1)(2-p)}$, and $\hat{C}_2 := 2^{(p-1)(p-2)}$.

Proof. (i) The first inequality follows by the definition of the quasi-norm and the fact that for $p \in (1,2)$, $(a+b)^{p-2} \leq b^{p-2}$, for all $a, b \geq 0$. The second inequality follows by an application of Hölder's inequality. (ii) Similarly, the left inequality follows by the fact that $b^{p-2} \leq (a+b)^{p-2}$, for $p > 2$ and $a, b \geq 0$ and the definition of the quasi-norm. The right inequality can be obtained by Hölder's inequality. □

The next two results provide the monotonicity and continuity properties for the operator associated with the p-Laplacian. For completeness, we state the main steps of the proofs.

Lemma 2.6. *Let A be the p-Laplacian operator, defined in (2.1). Then,*

$$\langle Au - Av, u - v\rangle \geq \gamma_p \|\nabla(u-v)\|^p_{L^p(\Omega)}, \qquad \text{for all } p \in (2,\infty),\ u, v \in W^{1,p}_0(\Omega), \tag{2.13}$$

$$\langle Au - Av, u - v\rangle \geq \lambda_p \|u-v\|^2_{(p,u)}, \qquad \text{for all } p \in (1,\infty),\ u, v \in W^{1,p}_0(\Omega). \tag{2.14}$$

Here, $\gamma_p := C_2(p, p-2)$, $\lambda_p := \min(2^{p-2}, 2^{2-p})C_2(p,0)$ and $C_2(p,p-2)$, $C_2(p,0)$ are the constants of (2.5) with $\delta = p-2$ and $\delta = 0$, respectively.

Proof. Inequality (2.13) follows immediately from (2.5) with $\delta = p-2$. To prove (2.14), let $p \in (1,\infty)$ and $u, v \in W^{1,p}_0(\Omega)$. Then, (2.5) with $\delta = 0$ shows

$$\langle Au - Av, u - v\rangle \geq C_2(p,0)\int_\Omega (|\nabla u| + |\nabla v|)^{p-2}|\nabla(u-v)|^2 \,\mathrm{d}x. \tag{2.15}$$

For $p \in (1,2]$ there holds $|\nabla u| + |\nabla v| = |\nabla u| + |\nabla(u-v) - \nabla u| \leq 2(|\nabla u| + |\nabla(u-v)|)$, and hence

$$(|\nabla u| + |\nabla v|)^{p-2} \geq 2^{p-2}(|\nabla u| + |\nabla(u-v)|)^{p-2}. \tag{2.16}$$

For $p \in [2,\infty)$, there holds $|\nabla u| + |\nabla(u-v)| \leq 2(|\nabla u| + |\nabla v|)$, and hence

$$(|\nabla u| + |\nabla v|)^{p-2} \geq 2^{2-p}(|\nabla u| + |\nabla(u-v)|)^{p-2}. \tag{2.17}$$

We combine (2.15)–(2.17) and the proof is complete. □

Lemma 2.7. *Let $p \in (1,\infty)$, $u, v, w \in W^{1,p}_0(\Omega)$. Then, for every $\varepsilon > 0$ there holds*

$$|\langle Au - Av, w\rangle| \leq \kappa_p\Big(\mu_\varepsilon \|u-v\|^2_{(p,u)} + \varepsilon\|w\|^2_{(p,u)}\Big), \tag{2.18}$$

where $\kappa_p = \max(2^{p-2}, 2^{2-p})C_1(p,0)$, $\mu_\varepsilon = \max(1, \varepsilon^{-1/(p-1)}, (4\varepsilon)^{-1})$ and $C_1(p,0)$ is the constant of (2.4) with $\delta = 0$.

Proof. For all $p \in (1,\infty)$, (2.4) with $\delta = 0$ shows

$$|\langle Au - Av, w\rangle| \le C_1(p,0)\int_\Omega (|\nabla u| + |\nabla v|)^{p-2}|\nabla(u-v)|\cdot|\nabla w|\,\mathrm{d}x. \tag{2.19}$$

We assume that $p \in (1,2)$, and we note that since, $|\nabla u| + |\nabla(u-v)| \le 2(|\nabla u| + |\nabla v|)$, we get, $(|\nabla u| + |\nabla v|)^{p-2} \le 2^{2-p}(|\nabla u| + |\nabla(u-v)|)^{p-2}$. Substituting the above inequality into (2.19), and using (2.7), we get

$$\begin{aligned}|\langle Au - Av, w\rangle| &\le 2^{2-p}C_1(p,0)\int_\Omega (|\nabla u| + |\nabla(u-v)|)^{p-2}|\nabla(u-v)|\cdot|\nabla w|\,\mathrm{d}x\\ &\le 2^{2-p}C_1(p,0)\Big(\varepsilon^{-\frac{1}{p-1}}\|u-v\|^2_{(p,u)} + \varepsilon\|w\|^2_{(p,u)}\Big).\end{aligned}$$

Now, for $p \in [2,\infty)$, we observe that $(|\nabla u| + |\nabla v|)^{p-2} \le 2^{p-2}(|\nabla u| + |\nabla(u-v)|)^{p-2}$. Combining the above inequality with (2.19) and applying (2.8), we infer (2.18), similarly to the case $p \in (1,2)$. □

2.3. Weak solution

We define the solution space for problem (1.1) as

$$W := L^p(I; W_0^{1,p}(\Omega)) \cap W^{1,p'}(I; W^{-1,p'}(\Omega)) \cap L^\infty(I; L^2(\Omega)),$$

where $p' = p/(p-1)$ is the dual exponent of p. Then, for an initial condition $u_0 \in L^2(\Omega)$ and a right-hand side $f \in L^{p'}(I; W^{-1,p'}(\Omega))$, the weak formulation of (1.1) reads us: find $u \in W$, such that $\forall v \in W$

$$(u(T), v(T)) + \int_I \big(-\langle\partial_t v, u\rangle + (|\nabla u|^{p-2}\nabla u, \nabla v)\big)\,\mathrm{d}t = (u_0, v(0)) + \int_I \langle f, v\rangle\,\mathrm{d}t. \tag{2.20}$$

The following result for existence and uniqueness of weak solution has been established in [30]:

Theorem 2.8. *If $f \in L^{p'}(I; W^{-1,p'}(\Omega))$, $u_0 \in L^2(\Omega)$ and $p > \max(1, 2d/(d+2))$, then the problem (1.1) has a unique weak solution $u \in W$ that satisfies the estimate*

$$\|u\|^2_{L^\infty(I;L^2(\Omega))} + \frac{1}{p'}\|u\|^p_{L^p(I;W^{1,p}(\Omega))} \le \|u_0\|^2_{L^2(\Omega)} + \frac{C^{p'}}{p'}\|f\|^{p'}_{L^{p'}(I;W^{-1,p'}(\Omega))}, \tag{2.21}$$

where $C > 0$ denotes a constant, depending only on the domain Ω. If in addition, $f \in L^2(I; L^2(\Omega))$, and $u_0 \in W_0^{1,p}(\Omega) \cap L^p(\Omega)$ then, $u \in L^\infty(I, W^{1,p}(\Omega)) \cap H^1(I; L^2(\Omega))$.

Proof. The proof of existence (applicable in a more general setting including also a semilinear term) is given in detail in [30] by using an abstract framework of pseudomonotone, semi-coercive operators in a Gelfand triple $V \subset H \subset V^*$, that fits the p-Laplacian operator for $V = W_0^{1,p}(\Omega)$ and $H = L^2(\Omega)$. Uniqueness of the weak solution can then be proved easily, using the monotonicity condition (2.14). The estimate follows by testing with u. The enhanced regularity is proved in [30]. □

3. The fully-discrete scheme

3.1. Weak formulation and discretization

We consider a family of partitions $0 = t_0 < t_1 < \cdots < t_N = T$, $N \in \mathbb{N}$ of $\bar{I}$, with time-stepping sizes $\tau_n = t_n - t_{n-1}$ and maximal time-stepping size $\tau = \max_n \tau_n$. We denote by $I_n := (t_{n-1}, t_n]$, $n = 1,\ldots,N$ and we assume that the family of time-partitions is quasi-uniform, in the sense that there exists a constant $\theta \ge 1$, independent of N, such that

$$\tau = \max_{1\le n\le N}(t_n - t_{n-1}) \le \theta \min_{1\le n\le N}(t_n - t_{n-1}). \tag{3.1}$$

For any function $w : \bar{I} \to W_0^{1,p}(\Omega)$, and for all $0 \leq n \leq N$ we use the notation $w^n = w(t_n)$, and $w_+^n = w(t_n^+) = \lim_{\varepsilon \to 0^+} w(t_n + \varepsilon)$, with associated jump at partition points denoted by $[w^n] = w_+^n - w^n$. Note that if w is continuous at t_n, then the jump $[w^n] = 0$.

We also consider a family of conforming finite element spaces $\{V_h^n\}_{n=0}^N$, $h > 0$, such that $V_h^n \subset W_0^{1,p}(\Omega)$ (see, *e.g.*, [8]). Further assumptions for these spaces will be specified below. Note that for each $0 \leq n \leq N$, we may use a different finite element space V_h^n. This allows us to consider different meshes of Ω at each time step, or every few other steps. For a non-negative integer k, the fully-discrete space is defined by

$$\mathcal{U}_h := \Big\{ w_h \in L^p(I; W_0^{1,p}(\Omega)) : \; w_h|_{I_n} \in \mathcal{P}_k(I_n; V_h^n), n = 1, \dots, N \Big\}.$$

Then, the associated fully discrete discontinuous-in-time Galerkin scheme (dG(k)) reads: given $f \in L^{p'}(I; W^{-1,p'}(\Omega))$, with $p' = p/(p-1)$, and $u_h^0 = P_h^0 u_0$, where $P_h^0 : L^2(\Omega) \to V_h^0$ denotes the standard $L^2(\Omega)$ projection, find $u_h \in \mathcal{U}_h$ such that, for all $1 \leq n \leq N$, and for all $v_h \in \mathcal{P}_k(I_n; V_h^n)$,

$$(u_h^n, v_h^n) + \int_{I_n} \big(-(u_h, \partial_t v_h) + (|\nabla u_h|^{p-2} \nabla u_h, \nabla v_h) \big) \, \mathrm{d}t = (u_h^{n-1}, v_{h+}^{n-1}) + \int_{I_n} \langle f, v_h \rangle \, \mathrm{d}t. \tag{3.2}$$

Therefore, using integration by parts in time, (3.2) becomes: find $u_h \in \mathcal{U}_h$ such that for all $1 \leq n \leq N$, and for all $v_h \in \mathcal{P}_k(I_n; V_h^n)$

$$([u_h^{n-1}], v_{h+}^{n-1}) + \int_{I_n} \big((\partial_t u_h, v_h) + (|\nabla u_h|^{p-2} \nabla u_h, \nabla v_h) \big) \, \mathrm{d}t = \int_{I_n} \langle f, v_h \rangle \, \mathrm{d}t. \tag{3.3}$$

It is clear that when $k = 0$, *i.e.*, for the case of dG(0)-implicit Euler-time discretization scheme existence and uniqueness can be proved by standard techniques. For $k \geq 1$, existence can be proved by a standard Brouwer's fixed point argument, see for instance, [1, 20]. Uniqueness can then be obtained by using the monotonicity condition (2.14).

3.2. Stability estimates in energy norm and energy quasi-norm

We turn our attention to main stability estimate with respect to the natural energy $\|u_h\|_{L^p(I;W^{1,p}(\Omega))}$ norm and its relation to stability properties of quasi-norm.

Theorem 3.1. *Let u_h be the solution to (3.3). Then, there exists a positive constant C, depending only on Ω such that:*

(i) *For $p \in (1, \infty)$, and for all $n = 1, \dots, N$,*

$$\|u_h^n\|_{L^2(\Omega)}^2 + \sum_{j=0}^{n-1} \|[u_h^j]\|_{L^2(\Omega)}^2 + \frac{2}{p'} \|\nabla u_h\|_{L^p(0,t_n;L^p(\Omega))}^p \leq \|u_h^0\|_{L^2(\Omega)}^2 + \frac{2C^{p'}}{p'} \|f\|_{L^{p'}(0,t_n;W^{-1,p'}(\Omega))}^{p'}. \tag{3.4}$$

(ii) *For $p \in (1, 2)$ and for all $w \in L^p(I; W_0^{1,p}(\Omega))$,*

$$\|u_h\|_{L^2(I;p,w)}^2 \leq \frac{p'}{2} \|u_h^0\|_{L^2(\Omega)}^2 + C^{p'} \|f\|_{L^{p'}(I;W^{-1,p'}(\Omega))}^{p'}. \tag{3.5}$$

(iii) *For $p \in (2, \infty)$, and for all $w \in L^p(I; W_0^{1,p}(\Omega))$,*

$$\|u_h\|_{L^2(I;p,w)}^2 \leq \max\left\{ \frac{2^{p-1}(p-2)\hat{C}_2^{\frac{p}{p-2}}}{p}, \frac{2}{p} \right\} \Big(p' \|u_h^0\|_{L^2(\Omega)}^2 + 2C^{p'} \|f\|_{L^{p'}(I;W^{-1,p'}(\Omega))}^{p'} + \|\nabla w\|_{L^p(I;L^p(\Omega))}^p \Big). \tag{3.6}$$

Here $\hat{C}_2$ denotes the constant of Lemma 2.5.

Proof. We substitute $v_h = u_h$ in (3.3) to obtain, for all $j = 1, \dots, N$,

$$\begin{aligned}\frac{1}{2}\|[u_h^{j-1}]\|^2_{L^2(\Omega)} + \frac{1}{2}\|u_h^j\|^2_{L^2(\Omega)} + \int_{I_j}\|\nabla u_h\|^p_{L^p(\Omega)}\,\mathrm{d}t &= \frac{1}{2}\|u_h^{j-1}\|^2_{L^2(\Omega)} + \int_{I_j}\langle f, u_h\rangle\,\mathrm{d}t \\ &\le \frac{1}{2}\|u_h^{j-1}\|^2_{L^2(\Omega)} + \frac{1}{p}\int_{I_j}\|\nabla u_h\|^p_{L^p(\Omega)}\,\mathrm{d}t + \frac{C^{p'}}{p'}\int_{I_j}\|f\|^{p'}_{W^{-1,p'}(\Omega)}\,\mathrm{d}t.\end{aligned}\tag{3.7}$$

The proof of (3.4) is complete after noting that $1 - \frac{1}{p} = \frac{1}{p'}$, and summation over $j = 1, \dots, n$. Inequality (3.5) follows from (2.11) and (3.4). Inequality (3.6) follows from (2.12) and (3.4). Indeed, applying Young's inequality (2.6), with $p/(p-2)$ and $p/2$ we get

$$\begin{aligned}\int_I \|u_h\|^2_{(p,w)}\,\mathrm{d}t &\le \hat{C}_2\int_I\big(\|\nabla w\|_{L^p(\Omega)} + \|\nabla u_h\|_{L^p(\Omega)}\big)^{p-2}\|\nabla u_h\|^2_{L^p(\Omega)}\,\mathrm{d}t \\ &\le \frac{2^{p-1}(p-2)\hat{C}_2^{p/(p-2)}}{p}\Big(\|\nabla w\|^p_{L^p(I,L^p(\Omega))} + \|\nabla u_h\|^p_{L^p(I;L^p(\Omega))}\Big) + \frac{2}{p}\int_I\|\nabla u_h\|^p_{L^p(\Omega)}\,\mathrm{d}t.\end{aligned}$$

Hence, substituting the estimate (3.4), we obtain the desired estimate, by Lemma 2.5, Hölder's inequality and stability estimate (3.4). □

3.3. Stability estimates in $L^\infty(I; L^2(\Omega))$

The stability in $L^\infty(I; L^2(\Omega))$, is trivial when considering low order schemes with $k = 0$, and $k = 1$ (see, *e.g.*, [9], for the linear parabolic problem). To include high order schemes we employ an idea from [9, 10], based on the discrete characteristic function and its approximation.

Theorem 3.2. *There exists a positive constant C_k independent of τ, h and u (depending only on k), such that, for $n = 1, \dots, N$,*

$$\|u_h\|^2_{L^\infty(I_n;L^2(\Omega))} \le C_k\left((p'+1)\|u_h^0\|^2_{L^2(\Omega)} + \left(\frac{C^{p'}}{p'} + C^{p'}\right)\int_0^{t_n}\|f\|^{p'}_{W^{-1,p'}(\Omega)}\,\mathrm{d}t\right),\tag{3.8}$$

where C denotes a constant depending only on the domain Ω.

Proof. Let $1 \le n \le N$. Stability for $k = 0$ or $k = 1$ is trivial, since (3.4) shows stability for $\|u_h^n\|^2_{L^2(\Omega)} + \|[u_h^{n-1}]\|^2_{L^2(\Omega)}$. Indeed, the latter quantity is an upper bound for $\|u_h\|_{L^\infty(I_n;L^2(\Omega))}$, when u_h is piecewise constant or linear with respect to time. Thus, it suffices to assume that $k \ge 2$. Now, for an arbitrary (but fixed for the moment) $t \in (t_{n-1}, t_n)$, we define $\phi \in \mathcal{P}_k(I_n)$ by

$$\phi(t_{n-1}^+) = 1 \quad \text{and} \quad \int_{I_n}\phi\psi\,\mathrm{d}s = \int_{t_{n-1}}^t \psi\,\mathrm{d}s, \quad \forall \psi \in \mathcal{P}_{k-1}(I_n).$$

Due to Lemma 3.2 of [10], there exists a positive constant C_k (depending on k), such that $\|\phi\|_{L^\infty(I_n)} \le C_k$. For any $z_h \in V_h^n$, we define the function $v_h : I_n \to V_h^n$, by $v_h(s) = \phi(s)z_h$, for $s \in I_n$. Note that $\phi \in \mathcal{P}_k(I_n)$ and $z_h \in V_h^n$. Thus, $v_h \in \mathcal{P}_k(I_n; V_h^n)$, and hence, we have

$$\int_{I_n}(\partial_s u_h, v_h)\,\mathrm{d}s = \int_{I_n}(\partial_s u_h, \phi z_h)\,\mathrm{d}s = \int_{I_n}\phi(\partial_s u_h, z_h)\,\mathrm{d}s.$$

Now, $\partial_s u_h \in \mathcal{P}_{k-1}(I_n; V_h^n)$ and z_h is independent of the integration variable $s \in I_n$, and hence, $(\partial_s u_h, z_h) \in \mathcal{P}_{k-1}(I_n)$. Then, the definition of ϕ, together with the above equation show

$$\int_{I_n}(\partial_s u_h, v_h)\,\mathrm{d}s = \int_{I_n}\phi(\partial_s u_h, z_h)\,\mathrm{d}s = \int_{t_{n-1}}^t(\partial_s u_h, z_h)\,\mathrm{d}s.$$

On choosing $v_h(s) = \phi(s)z_h$ as a test function into (3.3), we observe that

$$(u_{h+}^{n-1} - u_h^{n-1}, v_{h+}^{n-1}) + \int_{I_n} (\partial_s u_h, v_h)\, \mathrm{d}s = (u_{h+}^{n-1} - u_h^{n-1}, z_h) + \int_{t_{n-1}}^{t} (\partial_s u_h, z_h)\, \mathrm{d}s$$
$$= (u_{h+}^{n-1} - u_h^{n-1}, z_h) + (u_h(t) - u_{h+}^{n-1}, z_h) = (u_h(t) - u_h^{n-1}, z_h).$$

Hence, with this particular choice of v_h in (3.3), and using the above equality we infer,

$$\begin{aligned}(u_h(t) - u_h^{n-1}, z_h) &= \int_{I_n} \langle f, v_h\rangle\, \mathrm{d}s - \int_{I_n} (|\nabla u_h|^{p-2}\nabla u_h, \nabla v_h)\, \mathrm{d}s \\ &\le C_k C \int_{I_n} \|f\|_{W^{-1,p'}(\Omega)} \|\nabla z_h\|_{L^p(\Omega)}\, \mathrm{d}s + C_k \int_{I_n} \|\nabla u_h\|^{p-1}_{L^p(\Omega)} \|\nabla z_h\|_{L^p(\Omega)}\, \mathrm{d}s \\ &\le C_k C \|\nabla z_h\|_{L^p(\Omega)} \int_{I_n} \|f\|_{W^{-1,p'}(\Omega)}\, \mathrm{d}s + C_k \|\nabla z_h\|_{L^p(\Omega)} \int_{I_n} \|\nabla u_h\|^{p-1}_{L^p(\Omega)}\, \mathrm{d}s,\end{aligned}$$

where at the last step we have used the fact that $z_h \in V_h^n$ is independent of s. Here the constants C_k, and C depend only on k and on Ω respectively. Therefore, using Hölder's inequality, with $s_1 = p$ and $s_2 = p' = p/(p-1)$, and standard algebra

$$(u_h(t) - u_h^{n-1}, z_h) \le C_k \tau_n^{1/p} \|\nabla z_h\|_{L^p(\Omega)} \Big(C\|f\|_{L^{p'}(I_n;W^{-1,p'}(\Omega))} + \|\nabla u_h\|^{p-1}_{L^p(I_n;L^p(\Omega))}\Big).$$

We now choose $z_h = u_h(t)$ (for the arbitrary fixed $t \in I_n$), and integrate the resulting inequality with respect to $t \in I_n$ to infer

$$\begin{aligned}\int_{I_n} \|u_h\|^2_{L^2(\Omega)}\, \mathrm{d}t &\le \|u_h^{n-1}\|_{L^2(\Omega)} \int_{I_n} \|u_h\|_{L^2(\Omega)}\, \mathrm{d}t \\ &\quad + C_k \tau_n^{1/p} \int_{I_n} \|\nabla u_h\|_{L^p(\Omega)}\, \mathrm{d}t \Big(C\|f\|_{L^{p'}(I_n;W^{-1,p'}(\Omega))} + \|\nabla u_h\|^{p-1}_{L^p(I_n;L^p(\Omega))}\Big) \\ &\le \tau_n^{1/2} \|u_h^{n-1}\|_{L^2(\Omega)} \|u_h\|_{L^2(I_n;L^2(\Omega))} \\ &\quad + C_k \tau_n^{1/p} \tau_n^{1/p'} \|\nabla u_h\|_{L^p(I_n;L^p(\Omega))} \Big(C\|f\|_{L^{p'}(I_n;W^{-1,p'}(\Omega))} + \|\nabla u_h\|^{p-1}_{L^p(I_n;L^p(\Omega))}\Big),\end{aligned}$$

where at the last step we have used Hölder's inequalities, with $s_1 = s_2 = 2$ and $s_1 = p' = p/(p-1)$ and $s_2 = p$ respectively. Note that $(1/p) + (1/p') = 1$, and hence Young's inequalities, result to

$$\frac{1}{2}\|u_h\|^2_{L^2(I_n;L^2(\Omega))} \le C_k \tau_n \Big(\|u_h^{n-1}\|^2_{L^2(\Omega)} + \frac{C^{p'}}{p'} \|f\|^{p'}_{L^{p'}(I_n;W^{-1,p'}(\Omega))} + 2\|\nabla u_h\|^p_{L^p(I_n;L^p(\Omega))}\Big),$$

where we have used $\frac{1}{p} + 1 \le 2$. Then, we apply the inverse estimate $\|u_h\|^2_{L^\infty(I_n;L^2(\Omega))} \le C_k \tau_n^{-1} \|u_h\|^2_{L^2(I_n;L^2(\Omega))}$, (with C_k denoting a constant depending on k) to obtain the estimate after using (3.4). □

4. Error estimates

We present the main results starting from the key "almost symmetric estimate", in spirit of the classical Céa Lemma, developed under minimal regularity assumptions. Then, using the quasi-equivalence between the quasi-norm and the classical norms, we derive various estimates, that can be also applicable in the case of high-order schemes. Finally, estimates in $L^\infty(I, L^2(\Omega))$ are also obtained, by combining the techniques of [9,10], and a "boot-strap" argument.

4.1. Projections

We need to define various projection operators. We begin from the definition of the orthogonal projectors $P_h^n : L^2(\Omega) \to V_h^n$, for $n = 0, \ldots, N$,

$$(P_h^n w, v_h) = (w, v_h) \quad \forall v_h \in V_h^n \ \forall w \in L^2(\Omega),\ n = 0, \ldots, N. \tag{4.1}$$

From now on, we assume the following:

Assumption 4.1. *The finite element spaces $\{V_h^n\}_{n=0}^N$ are such that the orthogonal projections are uniformly stable in $W^{1,p}(\Omega)$, i.e., there exists a constant $C_{Proj} > 0$ (depending on p) which is independent of h and N, such that,*

$$\|P_h^n w\|_{W^{1,p}(\Omega)} \le C_{Proj} \|w\|_{W^{1,p}(\Omega)}, \quad n = 0, \ldots, N.$$

A variety of finite element configurations satisfies this property cf. [6, 11, 14]. Furthermore, we define a time-discrete projector $\pi^k : C(\bar{I}; L^2(\Omega)) \to U_k$, associated to discontinuous-in-time Galerkin schemes by: For all $v_k \in \mathcal{P}_{k-1}(I_n, W_0^{1,p}(\Omega))$,

$$(\pi^k w)(t_n) = w(t_n) \text{ and } \int_{I_n} (\pi^k w - w, v_k) \,\mathrm{d}t = 0,\ 1 \le n \le N, \tag{4.2}$$

where $U_k := \{u_k \in L^p(I; W_0^{1,p}(\Omega)) : \ u_k|_{I_n} \in \mathcal{P}_k(I_n; W_0^{1,p}(\Omega))\}$. For $k = 0$, we just define $\pi^k w$ by the first equation in (4.2). Finally, we define the fully-discrete projector $P_h^{\mathrm{loc}} : C(\bar{I}; L^2(\Omega)) \to \mathcal{U}_h$ by

$$P_h^{\mathrm{loc}} w|_{I_n} = P_h^n (\pi^k w|_{I_n}),\ 1 \le n \le N,\ \forall w \in C(\bar{I}; L^2(\Omega)). \tag{4.3}$$

These projections are classical in the analysis of discontinuous in time Galerkin schemes (see, *e.g,* [33], where various results regarding stability and approximation properties are presented). Below, we state a stability and an error estimate for the projector π^k in $L^p(I; W^{1,p}(\Omega))$ norm, which will be subsequently used for the derivation of error estimates. These estimates have been proved in [16], but we present the proof in the appendix, for the sake of completeness.

Lemma 4.2 ([16], Lems. 5.3, 5.4)**.** *Let $w \in W^{1,p}(I; W^{1,p}(\Omega)) \cap C(\bar{I}; L^2(\Omega))$. Then, there exists a constant $C_{k,p}$ depending only on p and k such that,*

$$\|\pi^k w\|_{L^p(I;W^{1,p}(\Omega))} \le C_{k,p} \big(\|w\|_{L^p(I;W^{1,p}(\Omega))} + \tau \|\partial_t w\|_{L^p(I;W^{1,p}(\Omega))} \big). \tag{4.4}$$

In addition, if $\partial_t^{k+1} w \in L^p(I; W^{1,p}(\Omega))$, there exists a constant $C_{k,p} > 0$, depending only upon k and p, such that

$$\|w - \pi^k w\|_{L^p(I;W^{1,p}(\Omega))} \le C_{k,p} \left(\sum_{n=1}^N \tau_n^{p(k+1)} \|\partial_t^{k+1} w\|^p_{L^p(I_n;W^{1,p}(\Omega))} \right)^{1/p}. \tag{4.5}$$

The latter lemma implies the following stability and error estimates for P_h^{loc}:

Theorem 4.3. *Let $w \in W^{1,p}(I; W^{1,p}(\Omega)) \cap C(\bar{I}; L^2(\Omega))$. Then, there exists a constant $C_{k,p}$, depending only on k and p, such that,*

$$\|P_h^{\mathrm{loc}} w\|_{L^p(I;W^{1,p}(\Omega))} \le C_{k,p} \big(\|w\|_{L^p(I;W^{1,p}(\Omega))} + \tau \|\partial_t w\|_{L^p(I;W^{1,p}(\Omega)} \big). \tag{4.6}$$

If $\partial_t^{k+1} w \in L^p(I; W^{1,p}(\Omega))$ then there exists constant $C_{k,p}$ depending on k and p, such that,

$$\|w - P_h^{\mathrm{loc}} w\|_{L^p(I;W^{1,p}(\Omega))} \le C_{k,p} \left(\sum_{n=1}^N \Big(\tau_n^{p(k+1)} \|\partial_t^{k+1} w\|^p_{L^p(I_n;W^{1,p}(\Omega))} + \|(\mathcal{I} - P_h^n) w\|^p_{L^p(I_n;W^{1,p}(\Omega))} \Big) \right)^{\frac{1}{p}}. \tag{4.7}$$

Proof. Inequality (4.6) is an immediate consequence of the fact that $(P_h^{\rm loc}w)(t) = P_h^n((\pi^k w)(t))$ for $t \in I_n$ and the stability of the L^2-projections in $W^{1,p}(\Omega)$. For (4.7), using triangle inequality, we deduce,

$$\begin{aligned}\|w - P_h^{\rm loc}w\|^p_{L^p(I_n;W^{1,p}(\Omega))} &\le 2^{p-1}\big(\|w - P_h^n w\|^p_{L^p(I_n;W^{1,p}(\Omega))} + \|P_h^n(w-\pi^k w)\|^p_{L^p(I_n;W^{1,p}(\Omega))}\big)\\ &\le 2^{p-1}\Big(\|w - P_h^n w\|^p_{L^p(I_n;W^{1,p}(\Omega))} + C^p_{Proj}\|w-\pi^k w\|^p_{L^p(I_n;W^{1,p}(\Omega))}\Big).\end{aligned}$$

Summing the above inequalities over $n = 1, \ldots, N$, we obtain

$$\|w - P_h^{\rm loc}w\|^p_{L^p(I;W^{1,p}(\Omega))} \le 2^{p-1}\left(C^p_{Proj}\|w-\pi^k w\|^p_{L^p(I;W^{1,p}(\Omega))} + \sum_{n=1}^N \|w - P_h^n w\|^p_{L^p(I_n;W^{1,p}(\Omega))}\right).$$

Thus, (4.7) immediately follows from (4.5). □

Remark 4.4. In the remaining of the paper, and similar to Lemma 4.2 and Theorem 4.3, we employ the notation $C_{k,p}$, that might be different in each occurance, for constants depending only on p and k, through the stability and approximation properties of the projections π_k and $P_h^{\rm loc}$. We only specify the additional dependence upon p in the constants appearing in the analysis of the stability and error estimates of the dG(k) scheme.

In the analysis below it is crucial to use the quasi-norm $\|\cdot\|_{L^2(I;p,w)}$ with weight $w = P_h^{\rm loc}u$. Below we collect its basic stability properties.

Lemma 4.5. *Let $u \in W^{1,p}(I;W^{1,p}(\Omega))\cap C(\bar{I};L^2(\Omega))$ and $u_h \in \mathcal{U}_h$ denote the solutions of (2.20) and (3.3), respectively. Let $w := P_h^{\rm loc}u$ be defined by (4.3) and suppose that Assumption 4.1 holds. Then, there exists constant $C{>}0$, depending on Ω such that:*

(i) *For $p \in (1,2)$,*

$$\|u_h\|^2_{L^2(I;p,P_h^{\rm loc}u)} \le \frac{p'}{2}\|u_h^0\|^2_{L^2(\Omega)} + C^{p'}\|f\|^{p'}_{L^{p'}(I;W^{-1,p'}(\Omega))}. \tag{4.8}$$

(ii) *For $p \in (2,\infty)$,*

$$\|u_h\|^2_{L^2(I;p,P_h^{\rm loc}u)} \le C_{k,p}\Big(\|u_h^0\|^2_{L^2(\Omega)} + \|f\|^{p'}_{L^{p'}(I;W^{-1,p'}(\Omega))} + \big(\|\nabla u\|_{L^p(I;L^p(\Omega))} + \tau\|\partial_t u\|_{L^p(I;W^{1,p}(\Omega))}\big)^p\Big). \tag{4.9}$$

Here, $C_{k,p}$ is the constant that appears in (3.6).

Proof. The corresponding bound for $p \in (1,2)$ is independent of the stability bounds of $w := P_h^{\rm loc}u$ and hence, it is stated in (3.5). For the case $p \in (2,\infty)$, using the stability estimate (4.6) into the estimate (3.6) of Theorem 3.1, we obtain (4.9). □

We close this subsection with the following classical error estimate for $P_h^{\rm loc}$ in $L^\infty(I;L^2(\Omega))$, see, *e.g.*, [33].

Theorem 4.6. *If $\partial_t^{k+1}w \in L^\infty(I;L^2(\Omega))$, there exists constant $C_k{>}0$, depending only on k such that*

$$\|w - P_h^{\rm loc}w\|_{L^\infty(I;L^2(\Omega))} \le C_k\Big(\tau^{k+1}\|\partial_t^{k+1}w\|_{L^\infty(I;L^2(\Omega))} + \max_{1\le n\le N}\|(\mathcal{I}-P_h^n)w\|_{L^\infty(I_n;L^2(\Omega))}\Big). \tag{4.10}$$

4.2. Symmetric quasi-norm error estimates

We are now ready to prove a symmetric error estimate of the fully-discrete scheme, with respect to the quasi-norm defined in (2.2) with the fully-discrete projection $P_h^{\rm loc}u$ as the weight function of this quasi-norm. The following estimate is valid for any $p \in (1,\infty)$. We will distinguish two cases. The first one corresponds to the general finite element space setting, where $V_h^n \subset H_0^1(\Omega)$ might change every few time steps, and the classical one where $V_h^n := V_h^0$ for every $n = 1, \ldots, N$.

Theorem 4.7. *Let $p \in (1,\infty)$, and $u \in W$, $u_h \in \mathcal{U}_h$, be the solutions of (1.1) and (3.2) respectively, with $u_h^0 = P_h^0 u_0$. Let $e := u_h - u = e_h + \hat{e}$, where $\hat{e} = P_h^{\mathrm{loc}} u - u$. Then, the following estimates hold:*

(1) *There exists a constant $C_p = \max\{2, 2^{p-1}\}$ such that*

$$\max_{1\le n\le N} \|e(t_n)\|_{L^2(\Omega)}^2 + \lambda_p \|e\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)} \le C_p \Bigg(\sum_{n=0}^{N-1} \|(\mathcal{I} - P_h^n)u(t_n)\|^2_{L^2(\Omega)} + \max_{1\le n\le N} \|(\mathcal{I} - P_h^n)u(t_n)\|^2_{L^2(\Omega)} + C_{\kappa_p,\lambda_p} \|\hat{e}\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)} \Bigg), \tag{4.11}$$

where $C_{\kappa_p,\lambda_p} := 2\kappa_p \max(1, (2\kappa_p/\lambda_p)^{1/(p-1)}, \kappa_p/(2\lambda_p)) + \lambda_p$, with λ_p, κ_p from Lemmas 2.6 and 2.7 respectively.

(2) *In addition, if $V_h^n = V_h^0$ for all $n = 1, \dots, N$, then we have,*

$$\max_{1\le n\le N} \|e(t_n)\|_{L^2(\Omega)}^2 + \lambda_p \|e\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)} \le C_p \Bigg(\max_{1\le n\le N} \|(\mathcal{I} - P_h^n)u(t_n)\|^2_{L^2(\Omega)} + C_{\kappa_p,\lambda_p} \|\hat{e}\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)} \Bigg). \tag{4.12}$$

Proof. We define $w_h = P_h^{\mathrm{loc}} u$ and split the error as $e = u_h - u = u_h - w_h + w_h - u = e_h + \hat{e}$. Let $1 \le n \le N$. On multiplying (1.1) with $v_h \in \mathcal{P}_k(I_n; V_h^n)$, integrating in space and time and then integrating by parts, we obtain

$$(u(t_n), v_h^n) + \int_{I_n} (-(u, \partial_t v_h) + \langle Au, v_h \rangle)\, \mathrm{d}t = (u(t_{n-1}^+), v_{h+}^{n-1}) + \int_{I_n} \langle f, v_h \rangle \, \mathrm{d}t,$$

for all $v_h \in \mathcal{P}_k(I_n, V_h^n)$. However, $u \in W$ and W is continuously embedded in $C(\overline{I}; L^2(\Omega))$, and hence, $(u(t_{n-1}^+), v_{h+}^{n-1}) = (u(t_{n-1}), v_{h+}^{n-1})$. Then, we obtain

$$(u(t_n), v_h^n) + \int_{I_n} (-(u, \partial_t v_h) + \langle Au, v_h \rangle)\, \mathrm{d}t = (u(t_{n-1}), v_{h+}^{n-1}) + \int_{I_n} \langle f, v_h \rangle \, \mathrm{d}t. \tag{4.13}$$

Now, for $p \in (1,\infty)$, we substract (4.13) from (3.2) to infer, for all $v_h \in \mathcal{P}_k(I_n, V_h^n)$, and for $n = 1, \dots, N$, the Galerkin orthogonality condition

$$(e(t_n), v_h^n) - (e(t_{n-1}), v_{h+}^{n-1}) + \int_{I_n} (-(e, \partial_t v_h) + \langle Au_h - Au, v_h \rangle)\, \mathrm{d}t = 0. \tag{4.14}$$

Using integration by parts in time, the relation $e_h = e - \hat{e}$, and (4.14), we get

$$\begin{aligned}
([e_h^{n-1}], v_{h+}^{n-1}) &+ \int_{I_n} ((\partial_t e_h, v_h) + \langle Au_h - Aw_h, v_h \rangle)\mathrm{d}t \\
&= (e_h^n, v_h^n) - (e_h^{n-1}, v_{h+}^{n-1}) + \int_{I_n} (-(e_h, \partial_t v_h) + \langle Au_h - Aw_h, v_h \rangle)\, \mathrm{d}t \\
&= -(\hat{e}(t_n), v_h^n) + (\hat{e}(t_{n-1}), v_{h+}^{n-1}) - \int_{I_n} (-(\hat{e}, \partial_t v_h) + \langle Aw_h - Au, v_h \rangle)\, \mathrm{d}t.
\end{aligned}$$

By the definition of $\hat{e}$, the first term and the third term on the right-hand side vanish, resulting to

$$\begin{aligned}
([e_h^{n-1}], v_{h+}^{n-1}) &+ \int_{I_n} ((\partial_t e_h, v_h) + \langle Au_h - Aw_h, v_h \rangle)\, \mathrm{d}t \\
&= ((P_h^{n-1} - \mathcal{I})u(t_{n-1}), v_{h+}^{n-1}) + \int_{I_n} \langle Au - Aw_h, v_h \rangle \, \mathrm{d}t.
\end{aligned} \tag{4.15}$$

We choose $v_h = e_h$ and apply (2.14) to obtain

$$\frac{1}{2}\|e_h^n\|^2_{L^2(\Omega)} + \frac{1}{2}\|[e_h^{n-1}]\|^2_{L^2(\Omega)} + \lambda_p \int_{I_n} \|e_h\|^2_{(p,w_h)}\,\mathrm{d}t$$
$$\leq \frac{1}{2}\|e_h^{n-1}\|^2_{L^2(\Omega)} + ((P_h^{n-1} - \mathcal{I})u(t_{n-1}), e_{h+}^{n-1}) + \int_{I_n} \langle Au - Aw_h, e_h\rangle\,\mathrm{d}t.$$

Summing the above inequalities, from 1 to n and using the fact that $e_h^0 = 0$ by definition, and (2.18) to bound the last term of the right hand side, we infer

$$\frac{1}{2}\|e_h^n\|^2_{L^2(\Omega)} + \frac{1}{2}\sum_{i=0}^{n-1}\|[e_h^i]\|^2_{L^2(\Omega)} + \lambda_p \int_0^{t_n} \|e_h\|^2_{(p,w_h)}\,\mathrm{d}t$$
$$\leq \sum_{i=0}^{n-1}((P_h^i - \mathcal{I})u(t_i), e_{h+}^i) + \int_0^{t_n}\langle Au - Aw_h, e_h\rangle\,\mathrm{d}t.$$
$$\leq \sum_{i=0}^{n-1}((P_h^i - \mathcal{I})u(t_i), e_{h+}^i) + \kappa_p\Big(\mu_\varepsilon \int_0^{t_n}\|\hat{e}\|^2_{(p,w_h)}\,\mathrm{d}t + \varepsilon\int_0^{t_n}\|e_h\|^2_{(p,w_h)}\,\mathrm{d}t\Big). \tag{4.16}$$

We select $\varepsilon = \kappa_p^{-1}\lambda_p/2$ and hide the quasi-norm of e_h on the left-hand side of (4.16), and we obtain

$$\|e_h^n\|^2_{L^2(\Omega)} + \sum_{i=0}^{n-1}\|[e_h^i]\|^2_{L^2(\Omega)} + \lambda_p\int_0^{t_n}\|e_h\|^2_{(p,w_h)}\,\mathrm{d}t$$
$$\leq 2\sum_{i=0}^{n-1}((P_h^i - \mathcal{I})u(t_i), e_{h+}^i) + M_{\kappa_p,\lambda_p}\int_0^{t_n}\|\hat{e}\|^2_{(p,w_h)}\,\mathrm{d}t, \tag{4.17}$$

where $M_{\kappa_p,\lambda_p} := 2\kappa_p \max(1, (2\kappa_p/\lambda_p)^{1/(p-1)}, \kappa_p/(2\lambda_p))$. Here λ_p, κ_p denote the constants of Lemma 2.6 and 2.7 respectively. It remains to bound the first term of the right-hand side of (4.17). Using the definition of P_h^i and Young's inequality to bound

$$((P_h^i - \mathcal{I})u(t_i), e_{h+}^i) = ((P_h^i - \mathcal{I})u(t_i), e_{h+}^i - e_h^i) \leq \frac{1}{2}\|(P_h^i - \mathcal{I})u(t_i)\|^2_{L^2(\Omega)} + \frac{1}{2}\|[e_h^i]\|^2_{L^2(\Omega)}, \tag{4.18}$$

and substituting (4.18) into (4.17) we deduce the estimate (4.11) by triangle inequality (2.10), the observation that $\|\hat{e}(t_n)\|_{L^2(\Omega)} = \|(\mathcal{I} - P_h^{\mathrm{loc}})u(t_n)\|_{L^2(\Omega)} = \|(\mathcal{I} - P_h^n)u(t_n)\|_{L^2(\Omega)}$ and standard algebra. If $V_h^n = V_h^0\ \forall 1 \leq n \leq N$, the first term of the right hand side of (4.17) vanishes, and hence the estimate (4.12) follows by triangle inequality (2.10). □

Remark 4.8. We observe that the definitions of the constants C_p and C_{κ_p,λ_p} in the above theorem, show that for large values of p, the constant in the upper bound is of order $\mathcal{O}(2^{3p-4}/C_2(p,0))$, with respect to p. Here, $C_2(p,0)$ is the monotonicity constant in (2.14).

Remark 4.9. As stated in Section 1, the quasi-norm that we used to obtain symmetric error estimates has the projection $P_h^{\mathrm{loc}}u$ as a weight, instead of the exact solution u. However, a similar estimate can be obtained for the quasi-norms $\|e\|_{L^2(I;p,u)}$ and $\|e\|_{L^2(I;p,u_h)}$ as well. This estimate is stated in the corollary below.

Corollary 4.10. *Let $p \in (1,\infty)$ and suppose that the assumptions of Theorem 4.7 are satisfied. Then the following estimates hold:*

(1) *There exists constant $\tilde{C}_p > 0$, depending only upon p, such that*

$$\max_{1\le n\le N} \|e(t_n)\|^2_{L^2(\Omega)} + \|e\|^2_{L^2(I;p,u)} + \|e\|^2_{L^2(I;p,u_h)}$$
$$\le \tilde{C}_p\Bigg(\sum_{n=0}^{N-1} \|(\mathcal{I} - P^n_h)u(t_n)\|^2_{L^2(\Omega)} + \max_{1\le n\le N} \|(\mathcal{I} - P^n_h)u(t_n)\|^2_{L^2(\Omega)}$$
$$+ \|\hat{e}\|^2_{L^2(I;p,P^{\mathrm{loc}}_h u)} + \|\hat{e}\|^2_{L^2(I;p,u_h)}\Bigg). \tag{4.19}$$

(2) *If in addition, $V^n_h = V^0_h$ for all $1 \le n \le N$, then the estimate takes the form*

$$\max_{1\le n\le N} \|e(t_n)\|^2_{L^2(\Omega)} + \|e\|^2_{L^2(I;p,u)} + \|e\|^2_{L^2(I;p,u_h)}$$
$$\le \tilde{C}_p\Big(\max_{1\le n\le N} \|(\mathcal{I} - P^n_h)u(t_n)\|^2_{L^2(\Omega)} + \|\hat{e}\|^2_{L^2(I;p,P^{\mathrm{loc}}_h u)} + \|\hat{e}\|^2_{L^2(I;p,u_h)}\Big). \tag{4.20}$$

Proof. For $p \in (1,\infty)$, (2.10) shows

$$\|e\|^2_{L^2(I;p,u_h)} \le \min(2, 2^{p-1})\Big(\|e_h\|^2_{L^2(I;p,u_h)} + \|\hat{e}\|^2_{L^2(I;p,u_h)}\Big). \tag{4.21}$$

Now, for $p \in (2,\infty)$, we have

$$\|e_h\|^2_{L^2(I;p,u_h)} = \int_I\int_\Omega \big(|\nabla u_h| + |\nabla(u_h - P^{\mathrm{loc}}_h u)|\big)^{p-2}|\nabla e_h|^2 \mathrm{d}x\,\mathrm{d}t$$
$$\le \int_I\int_\Omega \big(|\nabla P^{\mathrm{loc}}_h u| + 2|\nabla(u_h - P^{\mathrm{loc}}_h u)|\big)^{p-2}|\nabla e_h|^2 \mathrm{d}x\,\mathrm{d}t \le 2^{p-2}\|e_h\|^2_{L^2(I;p,P^{\mathrm{loc}}_h u)}. \tag{4.22}$$

We combine the latter inequality with (4.21), to infer

$$\|e\|^2_{L^2(I;p,u_h)} \le 2^{p-1}\Big(2^{p-2}\|e_h\|^2_{L^2(I;p,P^{\mathrm{loc}}_h u)} + \|\hat{e}\|^2_{L^2(I;p,u_h)}\Big). \tag{4.23}$$

Moreover, we can work exactly as in (4.22) to obtain

$$\|e\|^2_{L^2(I;p,u)} \le 2^{p-2}\|e\|^2_{L^2(I;p,u_h)} \le 2^{2p-3}\Big(2^{p-2}\|e_h\|^2_{L^2(I;p,P^{\mathrm{loc}}_h u)} + \|\hat{e}\|^2_{L^2(I;p,u_h)}\Big), \tag{4.24}$$

where at the last step, we applied (4.23). Similarly, we can show that for $p \in (1,2)$, there holds

$$\|e\|^2_{L^2(I;p,u_h)} \le 2\Big(2^{2-p}\|e_h\|^2_{L^2(I;p,P^{\mathrm{loc}}_h u)} + \|\hat{e}\|^2_{L^2(I;p,u_h)}\Big), \tag{4.25}$$

starting from $\|e_h\|^2_{L^2(I;p,P^{\mathrm{loc}}_h u)}$, bounding it from below by $2^{p-2}\|e_h\|^2_{L^2(I;p,u_h)}$ and then combining the resulting inequality with (4.21). In a similar fashion, we can show

$$\|e\|^2_{L^2(I;p,u)} \le 2^{3-p}\Big(2^{2-p}\|e_h\|^2_{L^2(I;p,P^{\mathrm{loc}}_h u)} + \|\hat{e}\|^2_{L^2(I;p,u_h)}\Big), \tag{4.26}$$

starting from $\|e\|^2_{L^2(I;p,u_h)}$ and bounding it from below by $\|e\|^2_{L^2(I;p,u)}$. Then, an application of (4.17), (4.18) and (4.11) to (4.24) and (4.24), or (4.26) and (4.26) yields the desired estimate. For the proof of (4.20) we simply observe that as in (4.11), the sum on the right-hand side of (4.19) vanishes, when $V^n_h = V^0_h$ for all $1 \le n \le N$. □

The importance of such estimates, in particular in the "almost symmetric" form established in Theorem 4.7 and Corollary 4.10, is evident since it also allows to handle schemes of high order (in both time and space) for nonlinear problems.

4.3. Convergence rates

In this section, we exploit the symmetric estimate of Theorem 4.7 to derive convergence rates. The estimates are of optimal order for $p \in (2,\infty)$, contrary to the singular case of $p \in (1,2)$, in which are of suboptimal order by a multiplicative factor of $p/2$. In order to obtain convergence rates, in addition to satisfy Assumption 4.1, we require that L^2-projections $P_h^n : L^2(\Omega) \to V_h^n$ associated to the finite element spaces V_h^n satisfy the following standard approximation properties.

Assumption 4.11. *Let $\{P_h^n\}_{n=0}^N$ be the orthogonal projections defined by (4.1). Then, for $p \in (1,\infty)$, there exist positive constants $C_{r,p}$ independent of τ, h and N, such that for a positive integer $r \geq 1$, the following estimates hold: For all $1 \leq n \leq N$,*

$$\|w - P_h^n w\|_{W^{1,p}(\Omega)} \leq C_{r,p} h^r \|w\|_{W^{r+1,p}(\Omega)}, \qquad \forall w \in W_0^{1,p}(\Omega) \cap W^{r+1,p}(\Omega), \tag{4.27}$$

$$\|w - P_h^n w\|_{L^2(\Omega)} \leq C_{r,2} h^{r+1} \|w\|_{H^{r+1}(\Omega)}, \qquad \forall w \in H_0^1(\Omega) \cap H^{r+1}(\Omega). \tag{4.28}$$

For example, for a bounded domain $\Omega \subset \mathbb{R}^d$ with suitably piecewise smooth boundary, the standard Lagrange finite element spaces of the form,

$$V_h^n := \{v_h \in C(\overline{\Omega}) : \; v_h|_K \in \mathcal{P}_r(K) \; \forall K \in \mathcal{T}_h^n \text{ and } v_h|_{\partial\Omega} = 0\}, \tag{4.29}$$

on quasi-uniform triangulations $\{\mathcal{T}_h^n\}_{n=0}^N$ in space, with K denoting a triangle on $\mathcal{T}_h^n$, satisfy Assumptions 4.1 and 4.11, with maximal element diameter h cf. [6, 11, 14]. We first consider the $p \in (1,2)$ case.

Theorem 4.12. *Let $p \in (1,2)$ and $u \in W$, $u_h \in \mathcal{U}_h$ be the solutions of (2.20) and (3.3) respectively, with $u_h^0 = P_h^0 u_0$. Suppose that the finite element spaces $\{V_h^n\}_{n=0}^N$ satisfy Assumptions 4.1 and 4.11 and let u satisfy*

$$\partial_t^{k+1} u \in L^p(I; W^{1,p}(\Omega)), \; u \in L^p(I; W^{r+1,p}(\Omega)) \cap C(\bar{I}; H^{r+1}(\Omega)).$$

Then the following estimates hold:

(1) *There exists a constant $\hat{C}_{k,r,p}$ depending on constant C_{κ_p,λ_p}, of Theorem 4.7 and on constants $C_{k,p}, C_{r,p}$ and $C_{r,2}$ of Theorem 4.3 and Assumption 4.11 respectively such that,*

$$\max_{1\leq n\leq N} \|e(t_n)\|_{L^2(\Omega)}^2 + \lambda_p \|e\|_{L^2(I;p,P_h^{\mathrm{loc}}u)}^2 \leq \hat{C}_{k,r,p}\Big(\tau^{p(k+1)} \|\partial_t^{k+1} u\|_{L^p(I;W^{1,p}(\Omega))}^p + h^{pr}\|u\|_{L^p(I;W^{r+1,p}(\Omega))}^p + (1+\theta\tau^{-1}) h^{2(r+1)} \|u\|_{C(\bar{I};H^{r+1}(\Omega))}^2\Big), \tag{4.30}$$

where θ is the constant in (3.1).

(2) *In addition, if $V_h^n = V_h^0$ for all $1 \leq n \leq N$, then it holds*

$$\max_{1\leq n\leq N} \|e(t_n)\|_{L^2(\Omega)}^2 + \lambda_p \|e\|_{L^2(I;p,P_h^{\mathrm{loc}}u)}^2 \leq \hat{C}_{k,r,p}\Big(\tau^{p(k+1)} \|\partial_t^{k+1} u\|_{L^p(I;W^{1,p}(\Omega))}^p + h^{pr}\|u\|_{L^p(I;W^{r+1,p}(\Omega))}^p + h^{2(r+1)} \|u\|_{C(\bar{I};H^{r+1}(\Omega))}^2\Big). \tag{4.31}$$

Proof. For $p \in (1,\infty)$, from (4.28) we have,

$$\sum_{n=0}^{N-1} \|(\mathcal{I} - P_h^n) u(t_n)\|_{L^2(\Omega)}^2 \leq C_{r,2}\theta\tau^{-1} h^{2(r+1)} \|u\|_{C(\bar{I};H^{r+1}(\Omega))}^2, \tag{4.32}$$

using the quasi-uniformity of the time steps, and

$$\max_{1\leq n\leq N} \|(\mathcal{I} - P_h^n) u(t_n)\|_{L^2(\Omega)}^2 \leq C_{r,2} h^{2(r+1)} \|u\|_{C(\bar{I};H^{r+1}(\Omega))}. \tag{4.33}$$

Now, for $p \in (1,2)$, from (2.11), (4.7) and (4.27) we easily infer,

$$\begin{aligned}\int_I \|\hat{e}\|^2_{(p,P_h^{\mathrm{loc}}u)}\,\mathrm{d}t &\leq \int_I \|\hat{e}\|^p_{W^{1,p}(\Omega)}\,\mathrm{d}t \\ &\leq C^p_{k,p}\left(\tau^{p(k+1)}\|\partial_t^{k+1}u\|^p_{L^p(I;W^{1,p}(\Omega))} + \sum_{n=1}^{N}\|(\mathcal{I}-P_h^n)u\|^p_{L^p(I_n;W^{1,p}(\Omega))}\right) \\ &\leq C^p_{k,p}\left(\tau^{p(k+1)}\|\partial_t^{k+1}u\|^p_{L^p(I;W^{1,p}(\Omega))} + C^p_{r,p}h^{pr}\|u\|^p_{L^p(I;W^{r+1,p}(\Omega))}\right).\end{aligned} \tag{4.34}$$

Hence, using (4.32), (4.33), (4.34) into (4.11) we obtain (4.30) and from (4.33), (4.34) and (4.12) we deduce (4.31). □

Remark 4.13. The above error estimates are suboptimal up to a multiplicative factor $p/2$. This suboptimality occurs, due to the lack of sharpness of the inequality (2.11) – which is present even at the continuous setting of (2.20) – that connects the quasi-norm to the Sobolev norm. The convergence rates could possibly be improved if the error estimate of the projection in the quasi-norm could be directly derived, without using (2.11). An estimate of this form has been established for the low-order-in-space case (see, *e.g.*, [12]). However, for high-order-in-space schemes, the numerical results of Section 6 suggest the possibility of a slight suboptimality when $p \in (1,2)$.

Finally, next theorem addresses the $p \in (2,\infty)$ case.

Theorem 4.14. *Let $p \in (2,\infty)$ and $u \in W$, $u_h \in \mathcal{U}_h$ be the solutions of (2.20) and (3.3) respectively, with $u_h^0 = P_h^0 u_0$. Suppose that the finite element spaces $\{V_h^n\}_{n=0}^N$ satisfy Assumptions 4.1 and 4.11 and let u satisfy*

$$\partial_t^{k+1}u \in L^p(I;W^{1,p}(\Omega)),\ u \in L^p(I;W^{r+1,p}(\Omega)) \cap C(\bar{I};H^{r+1}(\Omega)).$$

Then, the following estimates hold:

(1) *There exists a constant $\hat{C}_{k,r,p}$ depending on C_{κ_p,λ_p}, $C_{k,p}$, $C_{r,p}$ and $C_{r,2}$ such that,*

$$\begin{aligned}\max_{1\leq n\leq N}\|e(t_n)\|^2_{L^2(\Omega)} + \lambda_p\|e\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)} &\leq \hat{C}_{k,r,p}\Big(C_{u,k,p}\big(\tau^{2(k+1)}\|\partial_t^{k+1}u\|^2_{L^p(I;W^{1,p}(\Omega))} \\ &\quad + h^{2r}\|u\|^2_{L^p(I;W^{r+1,p}(\Omega))}\big) + (1+\theta\tau^{-1})h^{2(r+1)}\|u\|^2_{C(\bar{I};H^{r+1}(\Omega))}\Big),\end{aligned} \tag{4.35}$$

where $C_{u,k,p} = 2^{2(p-1)(p-2)}\big((1+C_{k,p})\|u\|_{L^p(I;W^{1,p}(\Omega))} + C_{k,p}\tau\|u_t\|_{L^p(I;W^{1,p}(\Omega))}\big)^{p-2}$. Here, $\hat{C}_{k,r,p}$ denotes a constant depending on the constant C_{κ_p,λ_p}, of Theorem 4.7 and on constants $C_{k,p}, C_{r,p}$ and $C_{r,2}$ of Theorem 4.3 and Assumption 4.11 respectively.

(2) *In addition, if $V_h^n = V_h^0$ for all $1 \leq n \leq N$, there holds,*

$$\begin{aligned}\max_{1\leq n\leq N}\|e(t_n)\|^2_{L^2(\Omega)} + \lambda_p\|e\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)} &\leq \hat{C}_{k,r,p}\Big(C_{u,k,p}\big(\tau^{2(k+1)}\|\partial_t^{k+1}u\|^2_{L^p(I;W^{1,p}(\Omega))} \\ &\quad + h^{2r}\|u\|^2_{L^p(I;W^{r+1,p}(\Omega))}\big) + h^{2(r+1)}\|u\|^2_{C(\bar{I};H^{r+1}(\Omega))}\Big).\end{aligned} \tag{4.36}$$

Proof. For $p \in (2,\infty)$, using (2.12), Hölder's inequality with $s_1 = p/(p-2)$ and $s_2 = p/2$, in space, and then once more in time, we infer by using standard algebra,

$$\begin{aligned}\int_I \|\hat{e}\|^2_{(p,P_h^{\mathrm{loc}}u)}\,\mathrm{d}t &\leq \hat{C}_2\int_I \big(\|\nabla P_h^{\mathrm{loc}}u\|_{L^p(\Omega)} + \|\nabla\hat{e}\|_{L^p(\Omega)}\big)^{p-2}\|\nabla\hat{e}\|^2_{L^p(\Omega)}\,\mathrm{d}t \\ &\leq 2^{(p-1)}\hat{C}_2\big(\|P_h^{\mathrm{loc}}u\|_{L^p(I;W^{1,p}(\Omega))} + \|\hat{e}\|_{L^p(I;W^{1,p}(\Omega))}\big)^{p-2}\|\hat{e}\|^2_{L^p(I;W^{1,p}(\Omega))}.\end{aligned}$$

Note that $\hat{e} = P_h^{\mathrm{loc}} u - u$. Hence triangle inequality, (4.6), (4.7) and (4.27), imply

$$\begin{aligned}
\int_I \|\hat{e}\|^2_{(p,P_h^{\mathrm{loc}}u)}\,\mathrm{d}t &\le 2^{(p-1)}\hat{C}_2\Big((1+C_{k,p})\|u\|_{L^p(I;W^{1,p}(\Omega))} + C_{k,p}\tau\|u_t\|_{L^p(I;W^{1,p}(\Omega))}\Big)^{p-2}\|\hat{e}\|^2_{L^p(I;W^{1,p}(\Omega))}\\
&\le C_{u,k,p}C_{k,p}^{2/p}\left(\tau^{2(k+1)}\|\partial_t^{k+1}u\|^2_{L^p(I;W^{1,p}(\Omega))} + \left(\sum_{n=1}^N \|(\mathcal{I}-P_h^n)u\|^p_{L^p(I_n;W^{1,p}(\Omega))}\right)^{2/p}\right)\\
&\le C_{u,k,p}C_{k,p}^{2/p}\Big(\tau^{2(k+1)}\|\partial_t^{k+1}u\|^2_{L^p(I;W^{1,p}(\Omega))} + h^{2r}C_{r,p}^{2/p}\|u\|^2_{L^p(I;W^{r+1,p}(\Omega))}\Big). \qquad (4.37)
\end{aligned}$$

The substitution of (4.32), (4.33), (4.37) into (4.11) implies (4.35) while the substitution of (4.33), (4.37) into (4.12) implies (4.36). □

Remark 4.15. We note that the quasi-norm in the estimates of Theorems 4.12 and 4.14 can be replaced by the quasi-norm $\|e\|_{L^2(I;p,u)}$ with slightly different constants on the right-hand side. This can be achieved using (4.19), instead of (4.11) whenever is necessary, and observing that for $p \in (1,2)$, (2.11) shows

$$\|\hat{e}\|^2_{L^2(I;p,u_h)} \le C_p\|\hat{e}\|^p_{L^p(I;W^{1,p}(\Omega))},$$

and that for $p \in (2,\infty)$, (2.12) shows

$$\begin{aligned}
\|\hat{e}\|^2_{L^2(I;p,u_h)} &\le C_p\big(\|\hat{e}\|_{L^p(I;W^{1,p}(\Omega))} + \|u_h\|_{L^p(I;W^{1,p}(\Omega))}\big)^{p-2}\|\hat{e}\|^2_{L^p(I;W^{1,p}(\Omega))}\\
&\le C_{k,p}\Big(\|u\|_{L^p(I;W^{1,p}(\Omega))} + \tau\|\partial_t u\|_{L^p(I;W^{1,p}(\Omega))}\\
&\qquad + \|u_0\|^{2/p}_{L^2(\Omega)} + \|f\|^{p'/p}_{L^{p'}(I;W^{-1,p'}(\Omega))}\Big)^{p-2}\|\hat{e}\|^2_{L^p(I;W^{1,p}(\Omega))},
\end{aligned}$$

where in the last step, we have used the stability estimates (4.6) and (3.4) for $P_h^{\mathrm{loc}}u$ and u_h, respectively. Thus, the additional term $\|\hat{e}\|^2_{L^2(I;p,u_h)}$ exhibits the convergence rates that appear in Theorems 4.12 and 4.14, requiring the same regularity assumptions for u.

Remark 4.16. For the lowest order scheme, *i.e.*, $k=0$, $r=1$, when $V_h^n = V_h^0$, for $n=1,\dots,N$, observe that for $p \in (2,\infty)$, our estimate implies convergence rate of order $\mathcal{O}(\tau+h)$, when $u \in L^p(I;W^{2,p}(\Omega)) \cap W^{1,p}(I;W^{1,p}(\Omega))$. This estimate is in accordance with the optimal estimates of [7,12]. In contrast, the estimates (4.30) and (4.31) for $p \in (1,2)$ are of lower order $\mathcal{O}(\tau^{p/2}+h^{p/2})$ (see, Rem. 4.13), than the corresponding optimal error estimates that have been established in [7,12]. However, unlike [7,12] our estimate is still applicable in the case of higher order schemes -which is the main goal of this paper- for all $p \in (1,\infty)$ and of optimal order $\mathcal{O}(\tau^{k+1}+h^r)$, when $p \in (2,\infty)$. We refer the reader to the computational examples in Section 6.

Remark 4.17. It is straightforward to derive error estimates under lower regularity assumptions in Theorems 4.12 and 4.14. In particular, if

$$\partial_t^{\ell+1}u \in L^p(I;W^{1,p}(\Omega)),\ u \in L^p(I;W^{s+1,p}(\Omega)) \cap C(\bar{I};H^{s+1}(\Omega)),$$

for some integer $0 \le \ell \le k$ and some $0 \le s \le r$, then we can replace k and r, with ℓ and s, respectively in (4.30)–(4.35). For example, in such case, the optimal estimate would be replaced by an error bound of order $\mathcal{O}(\tau^{\ell+1}+h^s)$. This also applies to all the following theorems, that consist of error bounds with convergence rates.

4.4. Norm error bounds

We derive error bounds in $L^p(I;W^{1,p}(\Omega))$-norm, using quasi-norm error estimates.

Theorem 4.18. *Suppose that the assumptions of Theorem 4.12 hold, and let $p \in (1,2)$. Then the following estimates hold:*

(1) *There exists $C > 0$, depending only on Ω, such that*

$$\lambda_p \|e\|^2_{L^p(I;W^{1,p}(\Omega))} \leq \hat{C}_{k,r,p} \hat{D}_1 \mathcal{K}^{2-p}_{f,u_0,u} \Big(\tau^{p(k+1)} \|\partial_t^{k+1} u\|^p_{L^p(I;W^{1,p}(\Omega))} + h^{pr} \|u\|^p_{L^p(I;W^{r+1,p}(\Omega))} + (1+\theta\tau^{-1}) h^{2(r+1)} \|u\|^2_{C(\bar{I};H^{r+1}(\Omega))} \Big). \tag{4.38}$$

(2) *If in addtion, $V_h^n = V_h^0$ for all $1 \leq n \leq N$, there exists $C > 0$ depending on Ω such that*

$$\lambda_p \|e\|^2_{L^p(I;W^{1,p}(\Omega))} \leq \hat{C}_{k,p} \mathcal{K}^{2-p}_{f,u_0,u} \Big(\tau^{p(k+1)} \|\partial_t^{k+1} u\|^p_{L^p(I;W^{1,p}(\Omega))} + h^{pr} \|u\|^p_{L^p(I;W^{r+1,p}(\Omega))} + h^{2(r+1)} \|u\|^2_{C(\bar{I};H^{r+1}(\Omega))} \Big). \tag{4.39}$$

Here, $\hat{C}_{k,r,p}$, denotes the constant of Theorem 4.12, $\hat{D}_1 := 2^{(p-1)(2-p)/p} \hat{C}_1 C^2$, with $\hat{C}_1$ the constant of (2.11), and

$$\mathcal{K}_{f,u_0,u} = \left(\frac{Cp'}{2} \|u_0\|^2_{L^2(\Omega)} + C^{p'+1} \|f\|^{p'}_{L^{p'}(I;W^{-1,p'}(\Omega))} \right)^{1/p} + (C_{k,p}+1) \|u\|_{L^p(I;W^{1,p}(\Omega))} + C_{k,p} \tau \|\partial_t u\|_{L^p(I;W^{1,p}(\Omega))}.$$

Proof. For $p \in (1,2)$, the equivalence of the $W^{1,p}(\Omega)$-seminorm with the $W^{1,p}(\Omega)$-norm in $W^{1,p}_0(\Omega)$, (2.11) and Hölder's inequality with exponents $s_1 = 2/(2-p)$, $s_2 = 2/p$, show

$$\begin{aligned} \int_I \|e\|^p_{W^{1,p}(\Omega)} \mathrm{d}t &\leq \hat{C}_1^{p/2} C^p \int_I \big(\|\nabla P_h^{\mathrm{loc}} u\|_{L^p(\Omega)} + \|\nabla e\|_{L^p(\Omega)} \big)^{\frac{(2-p)p}{2}} \|e\|^p_{(p,P_h^{\mathrm{loc}} u)} \mathrm{d}t \\ &\leq \hat{C}_1^{p/2} C^p \left(\int_I \big(\|\nabla P_h^{\mathrm{loc}} u\|_{L^p(\Omega)} + \|\nabla e\|_{L^p(\Omega)} \big)^p \mathrm{d}t \right)^{\frac{2-p}{2}} \left(\int_I \|e\|^2_{(p,P_h^{\mathrm{loc}} u)} \mathrm{d}t \right)^{p/2} \\ &\leq \hat{C}_1^{p/2} 2^{(p-1)(2-p)/2} \big(\|P_h^{\mathrm{loc}} u\|_{L^p(I;W^{1,p}(\Omega))} + \|e\|_{L^p(I;W^{1,p}(\Omega))} \big)^{\frac{(2-p)p}{2}} \|e\|^p_{L^2(I;p,P_h^{\mathrm{loc}} u)}, \end{aligned}$$

where $\hat{C}_1$ denotes the constant of (2.11). Hence, with $\hat{D}_1 := 2^{(p-1)(2-p)/p} \hat{C}_1 C^2$ the latter inequality implies,

$$\begin{aligned} \|e\|^2_{L^p(I;W^{1,p}(\Omega))} &\leq \hat{D}_1 \big(\|P_h^{\mathrm{loc}} u\|_{L^p(I;W^{1,p}(\Omega))} + \|e\|_{L^p(I;W^{1,p}(\Omega))} \big)^{2-p} \|e\|^2_{L^2(I;p,P_h^{\mathrm{loc}} u)} \\ &= \hat{D}_1 \mathcal{K}^{2-p}_{f,u_0,u} \|e\|^2_{L^2(I;p,P_h^{\mathrm{loc}} u)}. \end{aligned}$$

Here, we have used the stability estimates in Theorems 3.1 and 4.3, and we have estimated u_h^0 by the inequality $\|u_h^0\|_{L^2(\Omega)} = \|P_h^0 u_0\|_{L^2(\Omega)} \leq \|u_0\|_{L^2(\Omega)}$. Hence, from (4.30) and (4.31) we obtain (4.38) and (4.39), respectively. ☐

The next theorem covers the $p \in (2,\infty)$ case.

Theorem 4.19. *Suppose that the assumptions of Theorem 4.14 hold and let $p \in (2,\infty)$. Then, the following estimates hold:*

(1) *There exists a constant $C > 0$, depending only on Ω, such that:*

$$\lambda_p \|e\|^p_{L^p(I;W^{1,p}(\Omega))} \leq \hat{C}_{k,r,p} \Big(C_{u,k,p} \big(\tau^{2(k+1)} \|\partial_t^{k+1} u\|^2_{L^p(I;W^{1,p}(\Omega))} + h^{2r} \|u\|^2_{L^p(I;W^{r+1,p}(\Omega))} \big) + (1+\theta\tau^{-1}) h^{2(r+1)} \|u\|^2_{C(\bar{I};H^{r+1}(\Omega))} \Big). \tag{4.40}$$

(2) *If in addition, $V_h^n = V_h^0$ for all $1 \leq n \leq N$, there exists $C > 0$, depending on Ω such that,*

$$\lambda_p \|e\|^p_{L^p(I;W^{1,p}(\Omega))} \leq \hat{C}_{k,r,p}\Big(C_{u,k,p}(\tau^{2(k+1)}\|\partial_t^{k+1}u\|^2_{L^p(I;W^{1,p}(\Omega))} + h^{2r}\|u\|^2_{L^p(I;W^{r+1,p}(\Omega))}) + h^{2(r+1)}\|u\|^2_{C(\bar{I};H^{r+1}(\Omega))}\Big). \tag{4.41}$$

Here $\hat{C}_{k,r,p}, C_{u,k,p}$ denote the constants of Theorem 4.12.

Proof. The estimate follows directly from (4.35) and (4.36) respectively, after noting that (2.12) shows

$$\int_I \|e\|^p_{W^{1,p}(\Omega)}\,\mathrm{d}t \leq \int_I \|e\|^2_{(p,P_h^{\mathrm{loc}}u)}\,\mathrm{d}t.$$

□

Remark 4.20. The above theorem improves the corresponding $L^p(I;W^{1,p}(\Omega))$-error estimates of orders $\mathcal{O}(\tau^{1/2} + h^{p/2})$ and $\mathcal{O}(\tau^{1/p} + h^{2/p})$ for $p \in (1,2)$ and $p \in (2,\infty)$, respectively, established in [3] for the implicit Euler method in time, together with linear finite elements in space. Indeed, from Theorem 4.18 we conclude $L^p(I;W^{1,p}(\Omega))$-error estimates of orders $\mathcal{O}(\tau^{p/2} + h^{p/2})$ and $\mathcal{O}(\tau^{2/p} + h^{2/p})$ for $p \in (1,2)$ and $p \in (2,\infty)$, respectively, if we choose $k = 0$ and $r = 1$, *i.e.*, the piecewise constant discontinuous-in-time Galerkin scheme with linear finite elements in space. Furthermore, our norm error estimates are extended to arbitrary order fully-discrete schemes. Moreover, in [3], there has been established an error estimate of order $\mathcal{O}(\tau^{1/2} + h)$ for $p \in (1,2)$, with the additional regularity assumption $u \in L^p(I;C^{2,(2-p)/p}(\overline{\Omega}) \cap W^{3,1}(\Omega)) \cap L^2(I;H^2(\Omega))$. This result has not been extended in high order discontinuous-in-time Galerkin schemes so far.

5. Symmetric error estimates and estimates in $L^\infty(I;L^2(\Omega))$

Throughout this section, we will specify the integration variables in time integrals, in order to prevent any confusion in multiple time integrations. We observe that for $k = 0,1$, Theorem 4.7, implies $L^\infty(I;L^2(\Omega))$-estimates for all $p \in (1,\infty)$, without any conditionality between the discretization parameters. Hence, we restrict our dG schemes to the case $k \geq 2$.

Assumption 5.1. *There exists a constant $C_{inv} > 0$, independent of h, such that*

$$\|v_h\|_{H^1(\Omega)} \leq C_{inv}h^{-1}\|v_h\|_{L^2(\Omega)}, \quad \text{for all } v_h \in V_h^n,\ n = 1,\dots,N. \tag{5.1}$$

Assumption 5.1 is a classical inverse estimate that holds for a large variety of finite element spaces, including Lagrange finite element spaces, which are defined by (4.29).

Theorem 5.2. *Let $p \in (2,\infty)$, and let $u \in W$ and $u_h \in \mathcal{U}_h$ be the solutions of (1.1) and (3.3) respectively. Suppose that the assumptions of Theorem 4.7 hold. Let $e := u_h - u = e_h + \hat{e}$, where $\hat{e} = P_h^{\mathrm{loc}}u - u$, and $P_h^{\mathrm{loc}}u$ defined by (4.3).*

(1) *Then, there exist constants $\hat{D}^{(1)}_{k,p} := 2^{p-1}C_k\Big(1 + \hat{D}_2/\lambda_p\Big)\max\{2, 2^{p-1}\}$, $\hat{D}^{(2)}_{k,p} = C_k\lambda_p$ with C_k depending only on k, and $\hat{C}_u = 2^{p-2}\|u\|^{p-2}_{L^p(I;W^{1,p}(\Omega))}$ such that,*

$$\|e\|^2_{L^\infty(I;L^2(\Omega))} \leq \hat{D}^{(1)}_{k,p}\Bigg(\sum_{n=0}^{N-1}\|(\mathcal{I} - P_h^n)u(t_n)\|^2_{L^2(\Omega)} + \|\hat{e}\|^2_{L^\infty(I;L^2(\Omega))} + C_{\kappa_p,\lambda_p}\|\hat{e}\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)}\Bigg) + \hat{C}_u\hat{D}^{(2)}_{k,p}\Big(\|\hat{e}\|^{4/p}_{L^2(I;p,P_h^{\mathrm{loc}}u)} + \|e\|^{4/p}_{L^2(I;p,P_h^{\mathrm{loc}}u)}\Big). \tag{5.2}$$

Here, C_{κ_p,λ_p} is given in Theorem 4.7, and $\hat{D}_2 = \begin{cases} C_1(p,0)C_k & \text{when } p \in (2,3], \\ 2^{p-3}C_1(p,0)C_k & \text{when } p \in (3,\infty). \end{cases}$

(2) *If in addition* (5.1) *holds,* $u \in L^\infty(I; W^{1,\infty}(\Omega))$, *and the discretization parameters* τ *and* h *satisfy* $\tau \le C_k 2^{2-p} \hat{D}_2^{2-p} C_{inv}^{2-p} \|\nabla u\|^{2-p}_{L^\infty(I;L^\infty(\Omega))} h^2$ *for a suitable constant* C_k *depending upon* k*, then there exists a constant* $\hat{D}_{k,p}$ *depending on* k *and* p *through constants* $\hat{D}_2$, $1/\lambda_p$ *and* C_{κ_p,λ_p}, *and* $\hat{D}_3 = \begin{cases} 2^{p-3} & \text{when } p \in (3,\infty), \\ 1 & \text{when } p \in (2,3]. \end{cases}$ *such that*

$$\|e\|^2_{L^\infty(I;L^2(\Omega))} + \|e\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)} \le \hat{D}_{k,p} \Bigg(\sum_{n=0}^{N-1} \|(\mathcal{I} - P_h^n)u(t_n)\|^2_{L^2(\Omega)} + \|\hat{e}\|^2_{L^\infty(I;L^2(\Omega))} + \hat{D}_3 \Big(\|e\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)} + \|\hat{e}\|^2_{L^2(I;p,P_h^{\mathrm{loc}}u)} \Big) \Bigg). \tag{5.3}$$

(3) *Furthermore, if* $\partial_t u \in L^\infty(I; W^{1,\infty}(\Omega))$, *and the discretization parameters* τ *and* h *are selected to satisfy* $\tau^{p-1} \le C_k 2^{2-p} \hat{D}_2^{2-p} C_{inv}^{2-p} \|\partial_t u\|^{2-p}_{L^\infty(I;W^{1,\infty}(\Omega))} h^2$, *then* (5.3) *holds.*

Remark 5.3. Working as in the proof of Corollary 4.10, we can derive a similar estimate to (5.3), for the $L^\infty(I; L^2(\Omega))$-norm together with the quasi-norm $\|e\|_{L^2(I;p,u)}$, with an additional term $\|\hat{e}\|_{L^2(I;p,u_h)}$ on the right-hand side.

Remark 5.4. We note that the estimate for $k \ge 2$ does not follow from already established estimates at partition points and the jumps. On the other hand, given $t \in I_n$, it is not possible to use as a test function $\chi_{(t_{n-1},t]}(u_h - w_h)$. To circumvent this difficulty, we will use the approach of [10], combined with the previously developed estimate of Section 4.2. It turns out that the particular choice of the quasi-norm plays a crucial role. We observe that the first estimate is derived under minimal regularity assumptions, and it states that the error at arbitrary time points exhibits similar rates as $L^p(I; W^{1,p}(\Omega))$ norm, without imposing any condition between the temporal and spatial discretization parameters. If more regularity is assumed, we recover a symmetric error estimate, and hence an improved rate, under a restriction between h and τ.

Proof of Theorem 5.2. We will use again the notation $w_h = P_h^{\mathrm{loc}} u$ and we split the error as $e = u_h - u = u_h - w_h + w_h - u = e_h + \hat{e}$. Now, let $1 \le n \le N$ and let an arbitrary $t \in I_n$. We proceed similarly to Theorem 3.2. From the orthogonality condition (4.15), we obtain,

$$([e_h^{n-1}], v_{h+}^{n-1}) + \int_{I_n} (\partial_s e_h, v_h)\,\mathrm{d}s = ((P_h^{n-1} - \mathcal{I})u(t_{n-1}), v_{h+}^{n-1}) + \int_{I_n} \langle Au - Au_h, v_h \rangle\,\mathrm{d}s.$$

We set $v_h(s) = \phi(s) z_h$, $s \in I_n$, for an arbitrary $z_h \in V_h^n$, where $\phi \in \mathcal{P}_k(I_n)$ is defined by

$$\phi(t_{n-1}^+) = 1 \text{ and } \int_{I_n} \phi(s)\psi(s)\,\mathrm{d}s = \int_{t_{n-1}}^t \psi(s)\,\mathrm{d}s, \text{ for all } \psi \in \mathcal{P}_{k-1}(I_n).$$

We note that there exists a positive constant C_k, depending only upon k, such that $\|\phi\|_{L^\infty(I_n)} \le C_k$. Therefore working identically as in Theorem 3.2,

$$\begin{aligned}(e_h(t), z_h) &= (e_h^{n-1}, z_h) + ((P_h^{n-1} - \mathcal{I})u(t_{n-1}), z_h) + \int_{I_n} \langle Au(s) - Au_h(s), v_h(s) \rangle\,\mathrm{d}s \\ &\le \frac{1}{2}\|z_h\|^2_{L^2(\Omega)} + \|e_h^{n-1}\|^2_{L^2(\Omega)} + \|(P_h^{n-1} - \mathcal{I})u(t_{n-1})\|^2_{L^2(\Omega)} + \int_{I_n} |\langle Au(s) - Au_h(s), v_h(s) \rangle|\,\mathrm{d}s. \end{aligned} \tag{5.4}$$

For the last term, we employ (2.4) (with $\delta = 0$) and the fact that $\|\phi\|_{L^\infty(I_n)} \le C_k$,

$$\int_{I_n} |\langle Au(s) - Au_h(s), v_h(s) \rangle|\,\mathrm{d}s \le C_1(p,0) C_k \int_{I_n} \int_\Omega (|\nabla u(s)| + |\nabla u_h(s)|)^{p-2} |\nabla e(s)| |\nabla z_h|\,\mathrm{d}x\,\mathrm{d}s$$

$$\leq C_1(p,0)C_k\int_{I_n}\int_\Omega (2|\nabla u(s)|+|\nabla e(s)|)^{p-2}|\nabla e(s)||\nabla z_h|\mathrm{d}x\,\mathrm{d}s$$
$$\leq \hat{D}_2\int_{I_n}\int_\Omega\big(|\nabla e(s)|^{p-1}|\nabla z_h|+2^{p-2}|\nabla u(s)|^{p-2}|\nabla e(s)|\big)|\nabla z_h|\,\mathrm{d}x\,\mathrm{d}s, \tag{5.5}$$

where $\hat{D}_2 = C_1(p,0)C_k$ when $p\in(2,3]$ and $\hat{D}_2 = 2^{p-3}C_1(p,0)C_k$ when $p\in(3,\infty)$. For the first integral of (5.5), we apply Hölder's inequality with $s_1 = p/(p-1)$, $s_2 = p$, the fact that z_h is independent of s, and Young's inequality to get,

$$\int_{I_n}\int_\Omega |\nabla e(s)|^{p-1}\cdot|\nabla z_h|\mathrm{d}x\,\mathrm{d}s \leq \|\nabla e\|^{p-1}_{L^p(I_n;L^p(\Omega))}\tau_n^{1/p}\|\nabla z_h\|_{L^p(\Omega)}$$
$$\leq \frac{p-1}{p}\|\nabla e\|^p_{L^p(I_n;L^p(\Omega))}+\frac{\tau_n}{p}\|\nabla z_h\|^p_{L^p(\Omega)}\leq\frac{p-1}{p}\|e\|^2_{L^2(I_n;p,w_h)}+\frac{\tau_n}{p}\|z_h\|^2_{(p,w_h(t))}, \tag{5.6}$$

where at the last step we applied (2.12). For the second integral, we observe that using Hölder's inequality in space with $s_1 = p/(p-2)$, $s_2 = p$, $s_3 = p$, the fact that z_h is independent of s, and Hölder's inequality in time with $s_1 = p/(p-2)$, $s_2 = s_3 = p$,

$$\int_{I_n}\int_\Omega |\nabla u(s)|^{p-2}|\nabla e(s)|\cdot|\nabla z_h|\,\mathrm{d}x\,\mathrm{d}s \leq \|\nabla u\|^{p-2}_{L^p(I_n;L^p(\Omega))}\|\nabla e\|_{L^p(I_n;L^p(\Omega))}\tau_n^{1/p}\|\nabla z_h\|_{L^p(\Omega)}. \tag{5.7}$$

Substituting (5.6), (5.7) into (5.5) and the resulting inequality into (5.4), we infer,

$$(e_h(t),z_h)\leq\frac{1}{2}\|z_h\|^2_{L^2(\Omega)}+\|e_h^{n-1}\|^2_{L^2(\Omega)}+\|(P_h^{n-1}-\mathcal{I})u(t_{n-1})\|^2_{L^2(\Omega)}+\hat{D}_2\bigg(\frac{1}{p'}\|e\|^2_{L^2(I_n;p,w_h)}$$
$$+\frac{\tau_n}{p}\|z_h\|^2_{(p,w_h(t))}+2^{p-2}\|\nabla u\|^{p-2}_{L^p(I_n;L^p(\Omega))}\|\nabla e\|_{L^p(I_n;L^p(\Omega))}\tau_n^{1/p}\|\nabla z_h\|_{L^p(\Omega)}\bigg). \tag{5.8}$$

Setting $z_h = e_h(t)$, integrating with respect to t, using Hölder's inequality in time with $s_1 = p/(p-1), s_2 = p$, at the last term of (5.8), and then Young's inequality, we infer,

$$\frac{1}{2}\|e_h\|^2_{L^2(I_n;L^2(\Omega))}\leq\tau_n\|e_h^{n-1}\|^2_{L^2(\Omega)}+\tau_n\|(P_h^{n-1}-\mathcal{I})u(t_{n-1})\|^2_{L^2(\Omega)}+\hat{D}_2\bigg(\frac{\tau_n}{p'}\|e\|^2_{L^2(I_n;p,w_h)}$$
$$+\frac{\tau_n}{p}\|e_h\|^2_{L^2(I_n;p,w_h)}+\hat{C}_u\|\nabla e\|_{L^p(I_n;L^p(\Omega))}\tau_n^{1/p}\int_{I_n}\|\nabla e_h(t)\|_{L^p(\Omega)}\,\mathrm{d}t\bigg)$$
$$\leq\tau_n\|e_h^{n-1}\|^2_{L^2(\Omega)}+\tau_n\|(P_h^{n-1}-\mathcal{I})u(t_{n-1})\|^2_{L^2(\Omega)}+\hat{D}_2\bigg(\frac{\tau_n}{p'}\|e\|^2_{L^2(I_n;p,w_h)}$$
$$+\frac{\tau_n}{p}\|e_h\|^2_{L^2(I_n;p,w_h)}+\hat{C}_u\tau_n\|\nabla e_h\|_{L^p(I_n;L^p(\Omega))}\|\nabla e\|_{L^p(I_n;L^p(\Omega))}\bigg)$$
$$\leq\tau_n\Big(\|e_h^{n-1}\|^2_{L^2(\Omega)}+\|(P_h^{n-1}-\mathcal{I})u(t_{n-1})\|^2_{L^2(\Omega)}+\frac{\hat{D}_2}{\lambda_p}\Big(\lambda_p\|e\|^2_{L^2(I_n;p,w_h)}$$
$$+\lambda_p\|e_h\|^2_{L^2(I_n;p,w_h)}+\frac{\hat{C}_u\lambda_p}{2}\Big(\|\nabla e_h\|^2_{L^p(I_n;L^p(\Omega))}+\|\nabla e\|^2_{L^p(I_n;L^p(\Omega))}\Big)\Big)\Big), \tag{5.9}$$

where $\hat{C}_u := 2^{p-2}\|\nabla u\|^{p-2}_{L^p(I;L^p(\Omega))}$. Diving by τ_n, using an inverse estimate in time, and substituting $e_h = e-\hat{e}$, we obtain the analogue of (5.2) for e_h, by an application of triangle inequality (2.10) and the estimate of Theorem 4.7 that allows to bound the first four terms of (5.9). For the remaining last two terms of (5.9) we have used triangle inequality and (2.12) that shows $\int_I\|\nabla e\|^p_{L^p(\Omega)}\,\mathrm{d}t\leq\int_I\|e\|^2_{(p,P_h^{\mathrm{loc}}u)}\,\mathrm{d}t$.

If we impose a restriction between the sizes of the temporal and spatial discretization parameters, we can remove the suboptimal term that comes through (5.7). Indeed, if $\nabla u \in L^\infty(I; L^\infty(\Omega))$ then, it is clear that (5.7) can be bounded as follows: Using standard algebra along with Hölder's and Young's inequalities, we obtain,

$$\begin{aligned}\int_{I_n}\int_\Omega |\nabla u(s)|^{p-2}|\nabla e(s)||\nabla z_h| \,\mathrm{d}x\,\mathrm{d}s &= \int_{I_n}\int_\Omega |\nabla u(s)|^{\frac{p-2}{2}}|\nabla e(s)||\nabla u(s)|^{\frac{p-2}{2}}|\nabla z_h| \,\mathrm{d}x\,\mathrm{d}s \\ &\leq \frac{1}{2}\int_{I_n}\int_\Omega \big(|\nabla u(s)|^{p-2}|\nabla e(s)|^2 + |\nabla u(s)|^{p-2}|\nabla z_h|^2\big)\mathrm{d}x\,\mathrm{d}s \\ &\leq \frac{1}{2}\int_{I_n} \|e(s)\|^2_{(p,u(s))}\,\mathrm{d}s + \frac{1}{2}\int_{I_n}\int_\Omega |\nabla u(s)|^{p-2}|\nabla z_h|^2\,\mathrm{d}x\,\mathrm{d}s. \end{aligned} \tag{5.10}$$

Now, the second integral on the right-hand side of (5.10) becomes

$$\int_{I_n}\int_\Omega |\nabla u(s)|^{p-2}|\nabla z_h|^2\,\mathrm{d}x\,\mathrm{d}s \leq \tau_n \|\nabla u\|^{p-2}_{L^\infty(I_n;L^\infty(\Omega))}\|\nabla z_h\|^2_{L^2(\Omega)} \leq \frac{C_{inv}\tau}{h^2}\|\nabla u\|^{p-2}_{L^\infty(I;L^\infty(\Omega))}\|z_h\|^2_{L^2(\Omega)}, \tag{5.11}$$

where at the last step we applied the inverse estimate (5.1). Then, the combination of (5.4), (5.6), (5.10) and (5.11) gives

$$\begin{aligned}(e_h(t), z_h) \leq \frac{1}{2}\|z_h\|^2_{L^2(\Omega)} + \|e_h^{n-1}\|^2_{L^2(\Omega)} + \|(P_h^{n-1} - \mathcal{I})u(t_{n-1})\|^2_{L^2(\Omega)} + \hat{D}_2\Big(\|e\|^2_{L^2(I_n;p,w_h)} \\ + \tau_n\|z_h\|^2_{(p,w_h(t))} + \|e\|^2_{L^2(I_n;p,u)} + \frac{2^{p-2}C_{inv}\tau}{h^2}\|\nabla u\|^{p-2}_{L^\infty(I;L^\infty(\Omega))}\|z_h\|^2_{L^2(\Omega)}\Big).\end{aligned} \tag{5.12}$$

As before, we set $z_h = e_h(t)$ and we integrate with respect to time to deduce,

$$\begin{aligned}\left(\frac{1}{2} - \frac{C_{u,k,p}\tau}{h^2}\right)\int_{I_n}\|e_h(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t &\leq \tau_n\Big(\|e_h^{n-1}\|^2_{L^2(\Omega)} + \|(P_h^{n-1} - \mathcal{I})u(t_{n-1})\|^2_{L^2(\Omega)} \\ &\quad + \hat{D}_2\Big(\|e\|^2_{L^2(I_n;p,w_h)} + \|e_h\|^2_{L^2(I_n;p,w_h)} + \|e\|^2_{L^2(I_n;p,u)}\Big)\Big) \\ &\leq \tau_n\Big(\|e_h^{n-1}\|^2_{L^2(\Omega)} + \|(P_h^{n-1} - \mathcal{I})u(t_{n-1})\|^2_{L^2(\Omega)} \\ &\quad + \hat{D}_2\Big(\max(2, 2^{p-1})(\|e\|^2_{L^2(I_n;p,w_h)} + \|\hat{e}\|^2_{L^2(I_n;p,w_h)}) + \|e\|^2_{L^2(I_n;p,u)}\Big)\Big),\end{aligned} \tag{5.13}$$

where $C_{u,k,p} = 2^{p-2}\hat{D}_2 C_{inv}\|\nabla u\|^{p-2}_{L^\infty(I;L^\infty(\Omega))}$, and at the last step we have used the identity $e_h = e - \hat{e}$ and then triangle inequality (2.10). For the last term in (5.13), we observe that adding and subtracting appropriate terms, and using Hölder's inequality with $s_1 = p/(p-2)$, $s_2 = p/2$, in space and time, and then Young's inequality, we obtain the following estimate,

$$\begin{aligned}\|e\|^2_{L^2(I_n;p,u)} &= \int_{I_n}\int_\Omega (|\nabla e| + |\nabla u|)^{p-2}|\nabla e|^2\,\mathrm{d}x\,\mathrm{d}t \\ &\leq \int_{I_n}\int_\Omega (|\nabla e| + |\nabla \hat{e}| + |\nabla w_h|)^{p-2}|\nabla e|^2\,\mathrm{d}x\,\mathrm{d}t \\ &\leq \hat{D}_3\int_{I_n}\int_\Omega \Big((|\nabla e| + |\nabla w_h|)^{p-2}|\nabla e|^2 + |\nabla \hat{e}|^{p-2}|\nabla e|^2\Big)\,\mathrm{d}x\,\mathrm{d}t \\ &\leq \hat{D}_3\Big(\|e\|^2_{L^2(I_n,p;w_h)} + \|\nabla \hat{e}\|^{p-2}_{L^p(I_n;L^p(\Omega))}\|\nabla e\|^2_{L^p(I_n;L^p(\Omega))}\Big) \\ &\leq \hat{D}_3\Big(\|e\|^2_{L^2(I_n,p;w_h)} + \frac{p-2}{p}\|\nabla \hat{e}\|^p_{L^p(I_n;L^p(\Omega))} + \frac{2}{p}\|\nabla e\|^p_{L^p(I_n;L^p(\Omega))}\Big)\end{aligned}$$

$$\leq \hat{D}_3\Big(2\|e\|^2_{L^2(I_n,p;w_h)} + \|\hat{e}\|^2_{L^2(I_n,p;w_h)}\Big), \tag{5.14}$$

where at the last step we have used (2.12). Now, we suppose that τ and h are selected to satisfy $\tau \leq C^{-1}_{u,k,p}h^2/4$ and combine (5.13), and (5.14) to obtain,

$$\frac{1}{4}\|e_h\|^2_{L^2(I_n;L^2(\Omega))} \leq C_{k,p}\tau_n\big(\|e_h^{n-1}\|^2_{L^2(\Omega)} + \|(P_h^{n-1} - \mathcal{I})u(t_{n-1})\|^2_{L^2(\Omega)} + \|e\|^2_{L^2(I_n;p,w_h)} + \|\hat{e}\|^2_{L^2(I_n;p,w_h)}\big).$$

We divide the above estimate by τ_n and apply an inverse estimate in time. Moreover, we apply Theorem 4.7 and the triangle inequality in the resulting inequality, together with the respective estimates of e_h in the quasi-norms that appears in (5.3). This completes the proof of the second part of the theorem.

For the third part of the theorem, we return to the inequality (5.10) and observe that for all $s \in I_n$ and a.e. $x \in \Omega$, from the fundamental theorem of calculus, we have

$$\begin{aligned}|\nabla u(x,s)|^{p-2} &\leq \left(|\nabla u(x,t)| + \left|\int_s^t \partial_\sigma \nabla u(x,\sigma)\mathrm{d}\sigma\right|\right)^{p-2}\\ &\leq \big(|\nabla u(x,t)| + \|\partial_t u\|_{L^\infty(I;W^{1,\infty}(\Omega))}|t-s|\big)^{p-2} \leq C_p\big(|\nabla u(x,t)|^{p-2} + \tau_n^{p-2}\|\partial_t u\|^{p-2}_{L^\infty(I;W^{1,\infty}(\Omega))}\big).\end{aligned}$$

Then, the above inequality shows

$$\begin{aligned}\int_{I_n}\int_\Omega |\nabla u(s)|^{p-2}|\nabla z_h|^2\,\mathrm{d}x\,\mathrm{d}s &\leq C_p\left(\tau_n\int_\Omega |\nabla u(t)|^{p-2}|\nabla z_h|^2\,\mathrm{d}x + \tau_n^{p-1}\|\partial_t u\|^{p-2}_{L^\infty(I;W^{1,\infty}(\Omega))}\|\nabla z_h\|^2_{L^2(\Omega)}\right)\\ &\leq C_p\left(\tau_n\|z_h\|^2_{(p,u(t))} + C_{inv}\frac{\tau^{p-1}}{h^2}\|\partial_t u\|^{p-2}_{L^\infty(I;W^{1,\infty}(\Omega))}\|z_h\|^2_{L^2(\Omega)}\right),\end{aligned} \tag{5.15}$$

where in the last step we applied the inverse estimate (5.1). Now, we work exactly as in the proof of the second part of this theorem, but in this case we replace (5.11) with (5.15) and we impose the restriction $\tau^{p-1} \leq C_k 2^{2-p}\hat{D}_2^{2-p}C_{inv}^{2-p}\|\partial_t u\|^{2-p}_{L^\infty(I;W^{1,\infty}(\Omega))}h^2$. This modification concludes our proof. □

Remark 5.5. Similar to Theorem 4.12, if $V_h^n = V_h^0$, $\forall 1 \leq n \leq N$, then the first term of the right hand side of (5.2) and (5.3) can be dropped. In particular, under the assumptions of Theorem 5.2 (1), we deduce,

$$\begin{aligned}\|e\|^2_{L^\infty(I;L^2(\Omega))} \leq \hat{D}^{(1)}_{k,p}\Big(&\|\hat{e}\|^2_{L^\infty(I;L^2(\Omega))} + C_{\kappa_p,\lambda_p}\|\hat{e}\|^2_{L^2(I;p,P_h^{\rm loc}u)}\Big)\\ &+ \hat{C}_u\hat{D}^{(2)}_{k,p}\Big(\|\hat{e}\|^{4/p}_{L^2(I;p,P_h^{\rm loc}u)} + \|e\|^{4/p}_{L^2(I;p,P_h^{\rm loc}u)}\Big)\Big).\end{aligned} \tag{5.16}$$

while under the assumptions of Theorem 5.2 (2) and (3), we obtain

$$\|e\|^2_{L^\infty(I;L^2(\Omega))} + \|e\|^2_{L^2(I;p,P_h^{\rm loc}u)} \leq \hat{D}_{k,p}\Big(\|\hat{e}\|^2_{L^\infty(I;L^2(\Omega))} + \hat{D}_3\Big(\|e\|^2_{L^2(I;p,P_h^{\rm loc}u)} + \|\hat{e}\|^2_{L^2(I;p,P_h^{\rm loc}u)}\Big)\Big). \tag{5.17}$$

Working exactly as in the proof of Theorems 4.12 and 4.18, the following rates of convergence in $L^\infty(I;L^2(\Omega))$ are proved:

Theorem 5.6. *Let $p \in (2,\infty)$, $u \in W$ and $u_h \in \mathcal{U}_h$ be the solutions of (1.1) and (3.3), respectively. Suppose that the assumptions of Theorems 4.7, 4.12 and 5.2 hold.*

(1) *Then, there exists a constant $C > 0$, such that*

$$\|u - u_h\|_{L^\infty(I;L^2(\Omega))} \leq C\Big(\tau^{\frac{2}{p}(k+1)} + h^{\frac{2r}{p}} + \xi\max(\tau^{-1/2}h^{r+1}, \tau^{-1/p}h^{\frac{2}{p}(r+1)})\Big), \tag{5.18}$$

where $\xi = 0$ if $V_h^n = V_h^0$ for all $1 \leq n \leq N$ and $\xi = 1$ otherwise, i.e., ξ specifies whether or not the same finite element spaces are being used at each time step.

(2) *If Assumption 5.1 holds, and* $u \in L^\infty(I;W^{1,\infty}(\Omega))$ *with* τ *and* h *satifying* $\tau \leq C_k 2^{2-p}\hat{D}_2^{2-p}C_{inv}^{2-p}\|\nabla u\|_{L^\infty(I;L^\infty(\Omega))}^{2-p}h^2$, *or if* $\partial_t u \in L^\infty(I;W^{1,\infty}(\Omega))$ *with* τ *and* h *satisfying* $\tau^{p-1} \leq C_k 2^{2-p}\hat{D}_2^{2-p}C_{inv}^{2-p}\|\partial_t u\|_{L^\infty(I;W^{1,\infty}(\Omega))}^{2-p}h^2$, *then there exists a constant* $C > 0$, *such that*

$$\|u-u_h\|_{L^\infty(I;L^2(\Omega))} + \|u-u_h\|_{L^2(I;p,P_h^{\text{loc}}u)} + \|u-u_h\|_{L^2(I;p,u)} \leq C\Big(\tau^{k+1} + h^r + \xi\tau^{-1/2}h^{r+1}\Big), \tag{5.19}$$

where in both cases, the constants C *depend upon* k, p, Ω *and the norms of* u *appearing in Theorem 4.12.*

Remark 5.7. The estimates of Theorem 5.2 imply suboptimal rates of convergence under no restriction between the discretization parameters, while the symmetric error estimates that were obtained under restrictions between τ and h, yield the optimal rates. We also note that in the two above estimates, we have omitted the higher order terms and the dependence of the constant on the upper bounds.

So far, we have established error estimates in $L^\infty(I;L^2(\Omega))$, only for $p \in (2,\infty)$. Now, we turn our attention to the respective estimates for $p \in (1,2)$.

Theorem 5.8. *Let* $p \in (1,2)$, $u \in W$ *and* $u_h \in \mathcal{U}_h$ *be the solutions of* (1.1) *and* (3.3), *respectively. Suppose that the assumptions of Theorem 4.7 hold, and let* e, $\hat{e}$ *be the errors defined as in Theorem 5.6. Then, the following estimates hold:*

(1) *There exists a constant* $\hat{D}_{k,p} := C_k \max(2, 4C_{\kappa_p,\lambda_p}, C_1(p,2-p))$, *such that*

$$\|e\|_{L^\infty(I;L^2(\Omega))}^2 \leq \hat{D}_{k,p}\Bigg(\sum_{n=0}^{N-1}\|(\mathcal{I}-P_h^n)u(t_n)\|_{L^2(\Omega)}^2 + \|\hat{e}\|_{L^\infty(I;L^2(\Omega))}^2 + \|\hat{e}\|_{L^2(I;p,P_h^{\text{loc}}u)}^2 + \|\nabla e\|_{L^p(I;L^p(\Omega))}^p + \|\nabla\hat{e}\|_{L^p(I;L^p(\Omega))}^p\Bigg). \tag{5.20}$$

(2) *If in addition,* $V_h^n = V_h^0$, *for all* $1 \leq n \leq N$, *then we have*

$$\|e\|_{L^\infty(I;L^2(\Omega))}^2 \leq \hat{D}_{k,p}\Big(\|\hat{e}\|_{L^\infty(I;L^2(\Omega))}^2 + \|\hat{e}\|_{L^2(I;p,P_h^{\text{loc}}u)}^2 + \|\nabla e\|_{L^p(I;L^p(\Omega))}^p + \|\nabla\hat{e}\|_{L^p(I;L^p(\Omega))}^p\Big). \tag{5.21}$$

Proof. Let $t \in I_n$ be arbitrary. We select $v_h(s) = \phi(s)z_h$, as in the proof of Theorem 5.2 in (4.15). Now, we recall (5.4), which is valid for $p \in (1,2)$ as well, and we apply (2.4) with $\delta = 2-p$, together with Hölder's inequality with $s_1 = p/(p-1)$, $s_2 = p$, and Young's inequality, to obtain,

$$\begin{aligned}\int_{I_n}|\langle Au(s)-Au_h(s), v_h(s)\rangle|\,\mathrm{d}s &\leq C_kC_1(p,2-p)\int_{I_n}\int_\Omega|\nabla e(s)|^{p-1}|\nabla z_h|\mathrm{d}x\,\mathrm{d}s\\ &\leq C_kC_1(p,2-p)\|\nabla e\|_{L^p(I_n;L^p(\Omega))}^{p-1}\tau_n^{1/p}\|\nabla z_h\|_{;L^p(\Omega)}\\ &\leq C_kC_1(p,2-p)\left(\frac{1}{p'}\|\nabla e\|_{L^p(I_n;L^p(\Omega))}^p + \frac{\tau_n}{p}\|\nabla z_h\|_{L^p(\Omega)}^p\right).\end{aligned}$$

Now, we substitute the above inequality into (5.4), select $z_h = e_h(t)$ and integrate the resulting inequality over $t \in I_n$. Then, the proof is complete, after using an inverse estimate in time and the application of Theorem 4.7, similar to Theorem 5.2. □

Theorem 5.8 implies the following rates of convergence

Theorem 5.9. *Let $p \in (1,2)$, $u \in W$ and $u_h \in \mathcal{U}_h$ be the solutions of (1.1) and (3.3), respectively. Suppose that the Assumptions of Theorems 4.18 and 5.8 hold. Then, the error $u - u_h$ satisfies*

$$\|u-u_h\|_{L^\infty(I;L^2(\Omega))} \leq C\left(\tau^{\frac{p^2}{4}(k+1)} + h^{\frac{p^2 r}{4}} + \xi \max(\tau^{-1/2}h^{r+1}, \tau^{-p/2}h^{\frac{p}{2}(r+1)})\right), \tag{5.22}$$

where the constant C depends upon k, p, Ω and the norms of u appearing in Theorem 4.18, and ξ is defined as in Theorem 5.6.

Remark 5.10. We note that the arguments that have been applied to the case $p \in (2,\infty)$ for the optimal error estimates cannot be applied for $p \in (1,2)$, since $p-2$ is negative, and hence only suboptimal errors have been obtained for high order discontinuous-in-time Galerkin schemes. As before, we have omitted the higher order terms at the upper bounds and the dependence of the constant in (5.22).

6. Numerical results

We present some numerical experiments in which we have implemented the discontinuous-in-time Galerkin scheme in two spatial dimensions, for $k = 0,1,2$, using the finite element software Netgen/NGSolve [32]. We use uniform time partitions and standard conforming Lagrange simplicial finite element spaces for the spatial discretization with $V_h^n = V_h^0$, for all $1 \leq n \leq N$. We consider the two following examples.

Example 6.1. We define a smooth solution $u : [0,1]\times[0,1]\times[0,1]\to\mathbb{R}$ by

$$u(x_1,x_2,t) = e^{-t}x_1x_2(1-x_1)(1-x_2)\sin(2\pi x_1x_2), \tag{6.1}$$

for $(x_1,x_2) \in [0,1]\times[0,1],\ t \in [0,1]$, and we choose the data u_0 and f to be such that u is the solution to (1.1) with $\Omega = (0,1)\times(0,1)$ and $T = 1$.

Example 6.2. We define a solution $u : [-1,1]\times[-1,1]\times[0,1]\to\mathbb{R}$ of restricted regularity, by

$$u(x_1,x_2,t) = (t^{3/2}+1)(1-x_1^2)(1-x_2^2)(x_1^2\operatorname{sgn}(x_1) + x_2^2\operatorname{sgn}(x_2)), \tag{6.2}$$

for $(x_1,x_2) \in [-1,1]\times[-1,1],\ t \in [0,1]$, and we choose the data u_0 and f to be such that u is the solution to (1.1) with $\Omega = (-1,1)\times(-1,1)$ and $T = 1$. We observe that $u \in C^\infty(I;W^{2,\infty}(\Omega)) \cap W^{2,q}(I;W^{2,\infty}(\Omega))$, for $q \in [1,2)$, but $\partial_t^2 u \notin L^2(I;V)$ and $\partial_t^3 u \notin L^1(I;V)$, for any superspace V of $W^{2,\infty}(\Omega)$. Moreover, we observe that even the first- and second-order spatial derivatives of u are in $L^\infty(\Omega)$ for $t \in I$, the third-order spatial derivatives are only defined in the distributional sense.

Let us mention that in both examples, the gradients take zero values at least at the extremal points of u. In all tables of this section, we compute the errors $E_p := \|\nabla(u_h-u)\|_{L^p(I;L^p(\Omega))}$, $E_u := \|u_h-u\|_{L^2(I;p,u)}$ and $E_{P_h^{\rm loc}u} := \|u_h-u\|_{L^2(I;p,P_h^{\rm loc}u)}$ and their experimental rates of convergence with respect to τ, for $p = 3/2, 3, 5$.

The rates of convergence are being computed with the following strategy: for two different choices (τ_i,h_i) and (τ_{i+1},h_{i+1}) of discretization parameters, we calculate the corresponding errors $E(\tau_i,h_i)$ and $E(\tau_{i+1},h_{i+1})$ (with respect to the energy seminorm and the quasi-norm). Then, the respective rate of convergence with respect to time is being computed by $R_{i,i+1} = \frac{\log(E(\tau_i,h_i)/E(\tau_{i+1},h_{i+1}))}{\log(\tau_i/\tau_{i+1})}$. Therefore, if we compute the errors for different choices (τ_i,h_i), $i = 1,\ldots,m$, the rates that are presented in the following tables are being computed by the formula

$$R = \frac{1}{m-1}\sum_{i=1}^{m-1} R_{i,i+1} := \frac{1}{m-1}\sum_{i=1}^{m-1} \frac{\log(E(\tau_i,h_i)/E(\tau_{i+1},h_{i+1}))}{\log(\tau_i/\tau_{i+1})}.$$

TABLE 1. Example 6.1: Quasi-norm errors and temporal rates of convergence for $k = 0$ and piecewise linear finite elements.

τ	h	$p = 3/2$		$p = 3$		$p = 5$	
		E_u	$E_{P_h^{\text{loc}}u}$	E_u	$E_{P_h^{\text{loc}}u}$	E_u	$E_{P_h^{\text{loc}}u}$
1/20	1/20	2.5391e−02	2.5388e−02	1.7663e−03	1.7671e−03	8.3319e−05	8.3480e−05
1/40	1/40	1.4431e−02	1.4431e−02	7.7076e−04	7.7087e−04	3.0180e−05	3.0198e−05
1/80	1/80	7.9834e−03	7.9835e−03	3.4381e−04	3.4382e−04	1.2282e−05	1.2284e−05
1/160	1/160	4.3691e−03	4.3691e−03	1.6240e−04	1.6240e−04	5.6682e−06	5.6684e−06
Rate		0.8462	0.8462	1.1476	1.1479	1.2925	1.2935

TABLE 2. Example 6.1: $L^p(I; W^{1,p}(\Omega))$-errors and temporal rates of convergence for $k = 0$ and piecewise linear finite elements.

τ	h	$p = 3/2$	$p = 3$	$p = 5$
		E_p	E_p	E_p
1/20	1/20	8.9810e−03	1.2049e−02	1.6170e−02
1/40	1/40	4.5934e−03	6.1533e−03	8.2011e−03
1/80	1/80	2.3030e−03	3.0871e−03	4.1108e−03
1/160	1/160	1.1598e−03	1.5554e−03	2.0679e−03
Rate		0.9873	0.9844	0.9889

TABLE 3. Example 6.1: Quasi-norm errors and temporal rates of convergence for $k = 1$ and piecewise linear finite elements.

τ	h	$p = 3/2$		$p = 3$		$p = 5$	
		E_u	$E_{P_h^{\text{loc}}u}$	E_u	$E_{P_h^{\text{loc}}u}$	E_u	$E_{P_h^{\text{loc}}u}$
1/5	1/25	2.1540e−02	2.1940e−02	1.3785e−03	1.3069e−03	6.0246e−05	4.8948e−05
1/8	1/64	9.7310e−03	9.8769e−03	4.4612e−04	4.2558e−04	1.6157e−05	1.3616e−05
1/11	1/121	5.6034e−03	5.6729e−03	2.2102e−04	2.1250e−04	7.7915e−06	6.8233e−06
1/14	1/196	3.6567e−03	3.6951e−03	1.3086e−04	1.2664e−04	4.5301e−06	4.0720e−06
Rate		1.7312	1.7389	2.2596	2.2381	2.4463	2.3441

6.1. Discontinuous-in-time Galerkin schemes with linear finite elements

We present the results for the implementation of the discontinuous-in-time scheme with $k = 0, 1, 2$ combined with a piecewise linear Lagrange finite element method in space in Tables 1–6. These combinations of space and time discretizations have been chosen in order to show that for similar maximal element diameters h between higher and lower order discontinuous-in-time schemes, we can arrive at the same amounts of errors using coarser time-steppings in higher order discontinuous-in-time schemes. The optimal rates of convergence are expected to be $\mathcal{O}(\tau^{k+1} + h)$ with respect to the quasi-norms, and hence for a certain time-stepping size τ, we choose $h = \tau^{k+1}$ in Example 6.1 and $h = 2\tau^{k+1}$ in Example 6.2, so that the error will be of order at most $\mathcal{O}(\tau^{k+1})$.

From Tables 1–12, we observe that both examples exhibit the expected behaviour of the errors with respect to the quasi-norm, for $p = 3$ and $p = 5$. For $p = 3/2$ these rates are higher than the ones predicted in our estimates for $p \in (1, 2)$, possibly due to the fact that in both examples, the sets of zeros of the gradients have zero measures. Moreover, the errors with respect to the $L^p(I; W^{1,p}(\Omega))$ seminorm seem to be of order $\mathcal{O}(\tau^{k+1})$

TABLE 4. Example 6.1: $L^p(I;W^{1,p}(\Omega))$-errors and temporal rates of convergence for $k=1$ and piecewise linear finite elements.

τ	h	$p=3/2$ E_p	$p=3$ E_p	$p=5$ E_p
1/5	1/25	7.3884e−03	9.8548e−03	1.3086e−02
1/8	1/64	2.8963e−03	3.8702e−03	5.1438e−03
1/11	1/121	1.5392e−03	2.0566e−03	2.7321e−03
1/14	1/196	9.4915e−04	1.2685e−03	1.6863e−03
Rate		1.9940	1.9924	1.9914

TABLE 5. Example 6.1: Quasi-norm errors and temporal rates of convergence for $k=2$ and piecewise linear finite elements.

		$p=3/2$		$p=3$		$p=5$	
τ	h	E_u	$E_{P_h^{\mathrm{loc}}u}$	E_u	$E_{P_h^{\mathrm{loc}}u}$	E_u	$E_{P_h^{\mathrm{loc}}u}$
1/5	1/125	5.7133e−03	5.7130e−03	2.0021e−04	2.0022e−04	6.6578e−06	6.6580e−06
1/6	1/216	3.5134e−03	3.5133e−03	1.1372e−04	1.1372e−04	3.8442e−06	3.8442e−06
1/7	1/343	2.3138e−03	2.3138e−03	7.0893e−05	7.0894e−05	2.4094e−06	2.4094e−06
1/8	1/512	1.6057e−03	1.6057e−03	4.7329e−05	4.7329e−05	1.6156e−06	1.6156e−06
rate		2.7040	2.7039	3.0646	3.0647	3.0120	3.0121

TABLE 6. Example 6.1: $L^p(I;W^{1,p}(\Omega))$-errors and temporal rates of convergence for $k=2$ and piecewise linear finite elements.

τ	h	$p=3/2$ E_p	$p=3$ E_p	$p=5$ E_p
1/5	1/125	1.4875e−03	1.9877e−03	2.6412e−03
1/6	1/216	8.6201e−04	1.1524e−03	1.5323e−03
1/7	1/343	5.4285e−04	7.2576e−04	9.6517e−04
1/8	1/512	3.6389e−04	4.8658e−04	6.4717e−04
Rate		2.9958	2.9945	2.9926

TABLE 7. Example 6.2: Quasi-norm errors and temporal rates of convergence for $k=0$ and piecewise linear finite elements.

		$p=3/2$		$p=3$		$p=5$	
τ	h	E_u	$E_{P_h^{\mathrm{loc}}u}$	E_u	$E_{P_h^{\mathrm{loc}}u}$	E_u	$E_{P_h^{\mathrm{loc}}u}$
1/20	1/10	4.0172e−01	4.0187e−01	2.2881e−01	2.2891e−01	1.3850e−01	1.3912e−01
1/40	1/20	2.2143e−01	2.2142e−01	9.3620e−02	9.3659e−02	3.9491e−02	3.9569e−02
1/80	1/40	1.2034e−01	1.2034e−01	4.1001e−02	4.1009e−02	1.4462e−02	1.4470e−02
1/160	1/80	6.4479e−02	6.4479e−02	1.8907e−02	1.8908e−02	6.1911e−03	6.1919e−03
Rate		0.8797	0.8799	1.1990	1.1992	1.4945	1.4966

TABLE 8. Example 6.2: $L^p(I;W^{1,p}(\Omega))$-errors and temporal rates of convergence for $k=0$ and piecewise linear finite elements.

		$p=3/2$	$p=3$	$p=5$
τ	h	E_p	E_p	E_p
1/20	1/10	3.5960e−01	3.2047e−01	3.6039e−01
1/40	1/20	1.7872e−01	1.5900e−01	1.7943e−01
1/80	1/40	8.9265e−02	7.9230e−02	8.9084e−02
1/160	1/80	4.4659e−02	3.9612e−02	4.4416e−02
Rate		1.0031	1.0053	1.0068

TABLE 9. Example 6.2: Quasi-norm errors and temporal rates of convergence for $k=1$ and piecewise linear finite elements.

		$p=3/2$		$p=3$		$p=5$	
τ	h	E_u	$E_{P_h^{\text{loc}}u}$	E_u	$E_{P_h^{\text{loc}}u}$	E_u	$E_{P_h^{\text{loc}}u}$
1/5	2/25	3.2922e−01	3.2476e−01	1.6713e−01	1.7223e−01	8.7598e−02	9.6357e−02
1/8	2/64	1.4628e−01	1.4458e−01	5.2766e−02	5.4337e−02	1.9314e−02	2.1349e−02
1/11	2/121	8.2965e−02	8.2147e−02	2.5591e−02	2.6255e−02	8.4951e−03	9.2553e−03
1/14	2/196	5.3506e−02	5.3051e−02	1.5140e−02	1.5478e−02	4.8472e−03	5.2096e−03
Rate		1.7751	1.7700	2.3005	2.3099	2.7075	2.7380

TABLE 10. Example 6.2: $L^p(I;W^{1,p}(\Omega))$-errors and temporal rates of convergence for $k=1$ and piecewise linear finite elements.

		$p=3/2$	$p=3$	$p=5$
τ	h	E_p	E_p	E_p
1/5	2/25	2.8338e−01	2.5237e−01	2.8420e−01
1/8	2/64	1.1107e−01	9.8727e−02	1.1115e−01
1/11	2/121	5.8885e−02	5.2282e−02	5.8710e−02
1/14	2/196	3.6395e−02	3.2300e−02	3.6198e−02
Rate		1.9935	1.9966	2.0029

TABLE 11. Example 6.2: Quasi-norm errors and temporal rates of convergence for $k=2$ and piecewise linear finite elements.

		$p=3/2$		$p=3$		$p=5$	
τ	h	E_u	$E_{P_h^{\text{loc}}u}$	E_u	$E_{P_h^{\text{loc}}u}$	E_u	$E_{P_h^{\text{loc}}u}$
1/5	2/125	8.4914e−02	8.4909e−02	2.2817e−02	2.2818e−02	6.8825e−03	6.8810e−03
1/6	2/216	5.1279e−02	5.1277e−02	1.2917e−02	1.2918e−02	3.9189e−03	3.9186e−03
1/7	2/343	3.3246e−02	3.3246e−02	8.0636e−03	8.0636e−03	2.4560e−03	2.4559e−03
1/8	2/512	2.2745e−02	2.2745e−02	5.3806e−03	5.3806e−03	1.6427e−03	1.6426e−03
Rate		2.8067	2.8065	3.0690	3.0691	3.0440	3.0443

TABLE 12. Example 6.2: $L^p(I;W^{1,p}(\Omega))$-errors and temporal rates of convergence for $k=2$ and piecewise linear finite elements.

		$p=3/2$	$p=3$	$p=5$
τ	h	E_p	E_p	E_p
1/5	2/125	5.7015e−02	5.0617e−02	5.6829e−02
1/6	2/216	3.3032e−02	2.9322e−02	3.2874e−02
1/7	2/343	2.0819e−02	1.8481e−02	2.0700e−02
1/8	2/512	1.3954e−02	1.2392e−02	1.3884e−02
Rate		2.9947	2.9940	2.9978

TABLE 13. Example 6.1: Quasi-norm errors and temporal rates of convergence for $k=2$ and piecewise quadratic finite elements.

		$p=3/2$		$p=3$		$p=5$	
τ	h	E_u	$E_{P_h^{\text{loc}}u}$	E_u	$E_{P_h^{\text{loc}}u}$	E_u	$E_{P_h^{\text{loc}}u}$
1/8	$1/8^{3/2}$	2.0602e−03	2.0603e−03	6.6710e−05	6.6711e−05	4.8137e−06	4.8109e−06
1/16	$1/16^{3/2}$	3.6420e−04	3.6422e−04	7.4814e−06	7.4814e−06	6.1476e−07	6.1474e−07
1/32	$1/32^{3/2}$	8.4589e−05	8.4591e−05	9.7632e−07	9.7632e−07	6.7558e−08	6.7558e−08
Rate		2.3031	2.3031	3.0472	3.0472	3.0774	3.0770

in time, better than expected from Theorem 4.18, since it appears that the projection errors in $L^p(I;W^{1,p}(\Omega))$ are dominated by the projection errors in $L^\infty(I;L^2(\Omega))$, in these particular computational examples. Moreover, we observe that the quasi-norm errors with weights u and $P_h^{\text{loc}}u$ have close values. This justifies the choice of $\|\cdot\|_{L^2(I;p,P_h^{\text{loc}}u)}$ as the main quasi-norm for our error estimates.

Besides the calculations of the convergence rates in Tables 1–6, we consider the following numerical experiment: in Example 6.1 for $k=0,1,2$, we compute the dG(k) solutions with piecewise linear finite elements in space, with the same mesh sizes $h=1/20,1/40,1/60,\dots,1/280,1/300$ and uniform time-stepping sizes $\tau=1/\lfloor h^{-1/(k+1)}\rfloor$, where $\lfloor\cdot\rfloor$ denotes the integer part. With these choices of discretization parameters, we expect quasi-norm errors of order $\mathcal{O}(h)$. In Figure 1, we present the CPU times and the errors $\|e\|_{L^2(I;p,u)}$, $\|e\|_{L^2(I;p,P_h^{\text{loc}}u)}$ and $\|\nabla e\|_{L^p(I;L^p(\Omega))}$, as functions of h^{-1}, for $p=3$. The graphs of the errors are in logarithmic scale.

From the graphics in Figure 1, we obseve that in the dG(1) and dG(2) schemes we get almost the same errors in similar execution times. However, in the lower order dG(0) scheme, we obtain almost the same errors as in the two higher order schemes, in significantly larger CPU times. This indicates that if the exact solution regular enough, the coarse time-stepping for higher order discontinuous-in-time Galerkin schemes outbalances the increased computational cost that comes along with the increase of the polynomial degree in time and provides us with solutions of the same precision as in the dG(0) scheme, in notably smaller CPU times.

6.2. Quadratic discontinuous-in-time Galerkin scheme with quadratic finite elements

We now consider the discontinuous-in-time Galerkin scheme scheme with $k=2$, combined with piecewise quadratic Lagrange finite elements in space, *i.e.*, $r=2$. For a given time-stepping size τ, we choose $h=\tau^{3/2}$ in Example 6.1 and $h=2\tau^{2/3}$ for Example 6.2. With this choice of spatial discretization parameters, if the error has rate $\mathcal{O}(\tau^{3\alpha}+h^{2\alpha})$ for $\alpha=1,p/2$ or $2/p$, depending on the choices of p and the norm or the quasi-norm, then we should obtain a rate $\mathcal{O}(\tau^{3\alpha})$, when the solution is regular enough.

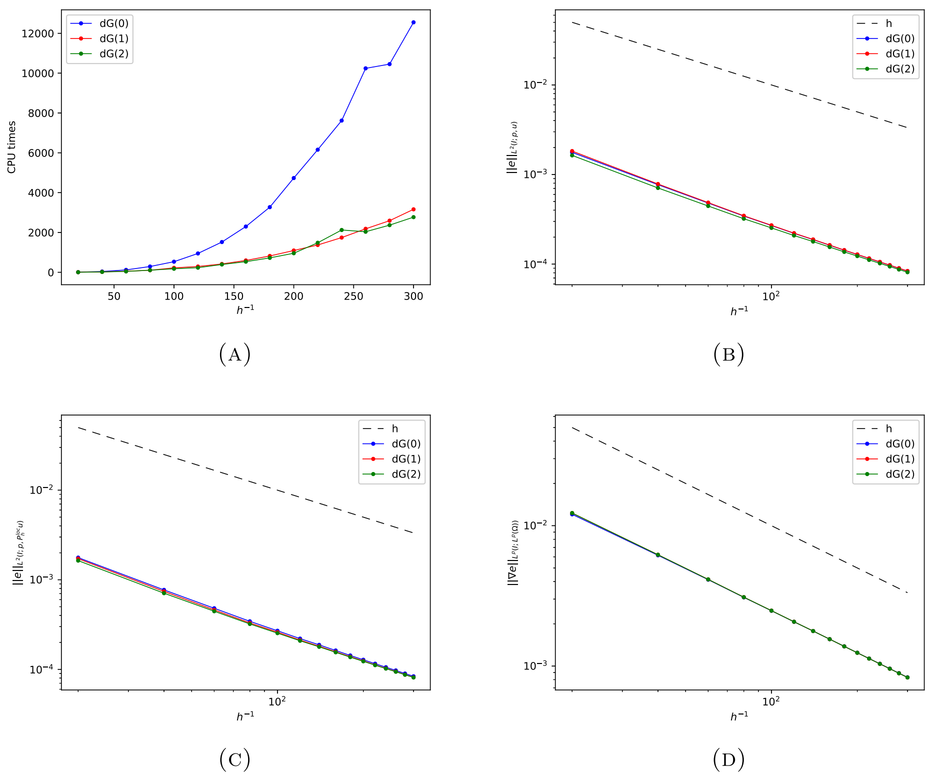

FIGURE 1. Example 6.1: (A) CPU times. (B) Quasi-norm errors with weight u. (C) Quasi-norm errors with weight $P_h^{\mathrm{loc}}u$. (D) Errors in the $L^p(I;W^{1,p}(\Omega))$-seminorm.

TABLE 14. Example 6.1: $L^p(I;W^{1,p}(\Omega))$-errors and temporal rates of convergence for $k=2$ and piecewise quadratic finite elements.

		$p=3/2$	$p=3$	$p=5$
τ	h	E_p	E_p	E_p
$1/8$	$1/8^{3/2}$	4.6370e−04	7.4918e−04	2.6734e−03
$1/16$	$1/16^{3/2}$	5.5960e−05	1.0554e−04	5.7906e−04
$1/32$	$1/32^{3/2}$	7.6378e−06	1.8584e−05	9.6124e−05
Rate		2.9619	2.6666	2.3988

From the results in Tables 13–16, we observe that contrary to the lower order space discretizations, the effects of the multiplicative factors $p/2$ and $2/p$ in the rates of convergence for $p\in(1,2)$ and $p\in(2,\infty)$, respectively, are visible in the computational results, although those rates mostly appear to be slightly better in practice. Moreover, the results confirm the optimality of the quasi-norm estimates for $p\in(2,\infty)$. The values of the quasi-norm errors with weights u and $P_h^{\mathrm{loc}}u$ are even closer in this case, because of the faster convergence of $P_h^{\mathrm{loc}}u$ to u. Furthermore, we observe that for example, in the quadratic discontinuous-in-time Galerkin scheme, if we choose $\tau=1/8$ for the linear finite element space discretization and for the quadratic finite element discretization, the errors of the two choices have close values for $h=\tau^3$ and $h=\tau^{3/2}$, respectively. This hints that for the same

TABLE 15. Example 6.2: Quasi-norm errors and temporal rates of convergence for $k = 2$ and piecewise quadratic finite elements.

τ	h	$p=3/2$ E_u	$p=3/2$ $E_{P_h^{loc}u}$	$p=3$ E_u	$p=3$ $E_{P_h^{loc}u}$	$p=5$ E_u	$p=5$ $E_{P_h^{loc}u}$
1/8	$2/8^{3/2}$	2.9553e−02	2.9551e−02	7.4445e−03	7.4433e−03	7.3821e−03	7.3625e−03
1/16	$2/16^{3/2}$	4.8437e−03	4.8436e−03	1.0718e−03	1.0718e−03	1.0062e−03	1.0059e−03
1/32	$2/32^{3/2}$	9.2607e−04	9.2606e−04	1.7789e−04	1.7789e−04	1.2089e−04	1.2088e−04
Rate		2.4980	2.4980	2.6935	2.6935	2.9661	2.9642

TABLE 16. Example 6.2: $L^p(I; W^{1,p}(\Omega))$-errors and temporal rates of convergence for $k = 2$ and piecewise quadratic finite elements.

τ	h	$p=3/2$ E_p	$p=3$ E_p	$p=5$ E_p
1/8	$2/8^{3/2}$	1.7714e−02	1.7843e−02	4.4117e−02
1/16	$2/16^{3/2}$	2.3175e−03	3.4416e−03	9.9125e−03
1/32	$2/32^{3/2}$	2.2669e−03	8.2044e−04	2.1582e−03
Rate		2.8804	2.2142	2.1766

time-stepping sizes, the quadratic finite elements give similar errors to the linear finite elements, for significantly coarser space meshes. Thus, if the exact solution is regular enough, the high order discretizations seem to be more efficient.

Moreover, comparing Examples 6.1 and 6.2, we observe similar rates among them, almost in all cases. We only observe slight reductions of the rates with respect to the quasi-norms, for $p = 3$ and with respect to the $L^p(I; W^{1,p}(\Omega))$, for all values of p. These rates might have been affected by the low regularity in space, since in order to exhibit optimal rates, one needs at least $W^{3,p}(\Omega)$ spatial regularity, while in this particular example, the solution is only $W^{2,\infty}(\Omega)$ in space.

Funding

The first author was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment" grant 3270. The second author was supported by a "H.F.R.I. PhD Fellowship grant", Project number 10688.

References

[1] G.D. Akrivis, V.A. Dougalis and O.A. Karakashian, On fully discrete Galerkin methods of second-order temporal accuracy for the nonlinear Schrödinger equation. *Numer. Math.* **59** (1991) 31–53.

[2] C. Atkinson and C.W. Jones, Similarity solutions in some non-linear diffusion problems and in boundary-layer flow of a pseudo plastic fluid. *Quart. J. Mech. Appl. Math.* **27** (1974) 193–211.

[3] J.W. Barrett and W.B. Liu, Finite element approximation of the parabolic p-Laplacian. *SIAM J. Numer. Anal.* **31** (1994) 413–428.

[4] L. Berselli and M. Růžička, Space-time discretization for nonlinear parabolic systems with p-structure. *IMA J. Numer. Anal.* **42** (2021) 260–299.

[5] L. Berselli, L. Diening and M. Růžička, Optimal error estimate for semi-implicit space-time dicretization for the equations describing incompressible generalized Newtonian fluids. *IMA J. Numer. Anal.* **55** (2015) 680–697.

[6] M. Boman, Estimates for the L_2-projection onto continuous finite element spaces in a weighted L_p-norm. *BIT Numer. Math.* **46** (2006) 249–260.

[7] D. Breit, L. Diening, J. Storn and J. Wichmann, The parabolic p-Laplacian with fractional differentiability. *IMA J. Numer. Anal.* **41** (2021) 2110–2138.

[8] S. Brenner and R. Scott, The mathematical theory of finite element methods, in Texts in Applied Mathematics. Vol. 15. Springer-Verlag, New York (1994).

[9] K. Chrysafinos and N.J. Walkington, Error estimates for the discontinuous Galerkin methods for parabolic equations. *SIAM J. Numer. Anal.* **44** (2006) 349–366.

[10] K. Chrysafinos and N.J. Walkington, Discontinuous Galerkin approximations of the Stokes and Navier–Stokes equations. *Math. Comput.* **79** (2010) 2135–2167.

[11] M. Crouzeix and V. Thomée, The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces. *Math. Comput.* **48** (1987) 521–532.

[12] L. Diening, C. Ebmeyer and M. Růžička, Optimal convergence for the implicit space-time discretization of parabolic systems with p-structure. *SIAM J. Numer. Anal.* **42** (2007) 457–472.

[13] L. Diening, C. Kreuzer and E. Süli, Finite element approximation of steady flows of incompressible fluids with implicit power-law-like rheology. *SIAM J. Numer. Anal.* **51** (2013) 984–1015.

[14] L. Diening, J. Storn and T. Tscherpel, On the Sobolev and L^p-stability of the L^2-projection. *SIAM J. Numer. Anal.* **59** (2021) 2571–2607.

[15] S. Eckstein and M. Růžička, On the full space-time discretization of the generalized Navier–Stokes equations: the Dirichlet case. *SIAM J. Numer. Anal.* **56** (2018) 2234–2261.

[16] E. Emmrich, Time discretisation of monotone nonlinear evolution problems by the discontinuous Galerkin method. *BIT Numer. Math.* **51** (2011) 581–607.

[17] E. Emmrich, Variable time-step θ-scheme for nonlinear evolution equations governed by a monotone operator. *Calcolo* **46** (2009) 187–210.

[18] E. Emmrich and M. Thalhammer, Stiffly accurate Runge–Kutta methods for nonlinear evolution equations governed by a monotone operator. *Math. Comput.* **79** (2010) 785–806.

[19] E. Emmrich and A. Wróblewska-Kamińska, Convergence of a full discretization of quasi-linear parabolic equations in isotropic and anisotropic Orlicz spaces. *SIAM J. Numer. Anal.* **51** (2013) 1163–1184.

[20] L.C. Evans, Partial Differential Equations. American Mathematical Society, Rhode Island (2010).

[21] E. Hansen, Galerkin/Runge–Kutta discretizations of nonlinear parabolic equations. *J. Comput. Appl. Math.* **205** (2007) 882–890.

[22] S. Ko, P. Pustějovská and E. Süli, Finite element approximation of an incompressible chemically reacting non-Newtonian fluid. *ESAIM: M2AN* **52** (2018) 509–541.

[23] C. Kreuzer, Reliable and efficient a posteriori error estimates for finite element approximations of the parabolic p-Laplacian. *Calcolo* **50** (2013) 50–79.

[24] D. Kröner, M. Růžička and I. Toulopoulos, Local discontinuous Galerkin numerical solutions of non-Newtonian incompressible flows modeled by p-Navier–Stokes equations. *J. Comput. Phys.* **270** (2014) 182–202.

[25] D. Kröner, M. Růžička and I. Toulopoulos, Numerical solutions os systems with (p, δ)-structure using local discontinuous finite element methods. *Int. J. Numer. Fluids* **76** (2014) 855–874.

[26] W. Liu and N. Yan, Quasi-norm a priori and a posteriori error estimates for the nonconforming approximation of p-Laplacian. *Numer. Math.* **84** (2001) 341–378.

[27] W. Liu and N. Yan, Quasi-norm local error estimators for p-Laplacian. *SIAM J. Numer. Anal.* **39** (2001) 100–127.

[28] J. Málek and K.R. Rajagopal, Mathematical issues concerning the Navier–Stokes equations and some of its generalizations , in *Evolution Equations.* Vol II of *Hanb. Differ. Equ.* Elsevier/North-Holland, Amsterdam (2005) 371–459.

[29] J.R. Philip, n-Diffusion. *Aust. J. Phys.* **14** (1961) 1–13.

[30] T. Roubiček, Nonlinear Partial Differential Equations with Applications. Springer Basel (2013).

[31] M. Růžička, Electrorheological Fluids: Modeling and Mathematical Theory. Lecture Notes in Math. 1748. Springer, Berlin (2000).

[32] J. Schöberl, A. Arnold, J. Erb, J. Melenk and T.P. Wihler, C++ 11 implementation of finite elements in NGSolve. Technical Report, Institute for Analysis and Scientific Computing, Vienna University of Technology, ASC Report 30/2014 (2014).

[33] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems. Springer, Berlin (2006).

[34] I. Toulopoulos, An interior penalty discontinuous Galerkin finite element method for quasilinear parabolic problems. *Finite Elements in Analysis and Design* **95** (2015) 42–50.

Appendix A.

A.1. Projection estimates

In this section, we present a similar proof of the projection estimates (*i.e.*, Lem. 4.2) to the ones that have been derived in [16], for the sake of completeness. We start with some notation. First of all, we denote by $\hat{J} = (-1, 1]$ a reference interval, and for $1 \leq n \leq N$, we denote by $G_n : \hat{J} \to I_n$, the affine mapping from $\hat{J}$ onto I_n, *i.e.*,

$$G_n(\hat{t}) = t_{n-1} + \frac{\tau_n}{2}(\hat{t}+1) \ \forall \hat{t} \in \hat{J} \text{ and } G_n^{-1}(t) = 2\tau_n^{-1}(t - t_{n-1}) - 1 \ \forall t \in I_n.$$

From now on, in order to prevent notational confusion, we will specify the dependece of a function on the temporal variable, since we will be transforming back and forth from $\hat{J}$ to I_n. Moreover, we shall denote by $\partial_t w$ the time derivative of a function $w \in W^{1,p}(I; W^{1,p}(\Omega))$, and by $\hat{\partial}_t \hat{w}$ the time derivative of a function $\hat{w} \in W^{1,p}(\hat{J}; W^{1,p}(\Omega))$.

Now, we observe that for $w \in C(\overline{I}; L^2(\Omega)) \cap L^p(I; W^{1,p}(\Omega))$, the projection $\pi^k w$, defined by (4.2), can equivalently be defined for $1 \leq n \leq N$, by

$$\pi^k w(t_n) = w(t_n) \text{ and } \int_{I_n} (\pi^k w(t) - w(t))\phi_k(t)\,\mathrm{d}t \qquad \forall \phi_k \in \mathcal{P}_{k-1}(I_n). \tag{A.1}$$

We now define the corresponding projection for functions defined in the reference interval, *i.e.*, for $\hat{w} \in C(\overline{\hat{J}}; L^2(\Omega)) \cap L^p(\hat{J}; W^{1,p}(\Omega))$, we define the projection $\hat{\pi}^k \hat{w} \in \mathcal{P}_k(\hat{J}; W_0^{1,p}(\Omega))$, *via*

$$\hat{\pi}^k \hat{w}(1) = \hat{w}(1) \text{ and } \int_{\hat{J}} (\hat{\pi}^k \hat{w}(\hat{t}) - \hat{w}(\hat{t}))\hat{\phi}_k(\hat{t})\mathrm{d}\hat{t} \qquad \forall \hat{\phi}_k \in \mathcal{P}_{k-1}(\hat{J}). \tag{A.2}$$

Upon applying change of variables through the affine mapping, to the integral on (A.2), one can easily show that the projection $\hat{\pi}^k \hat{w}$ satisfies

$$\hat{\pi}^k \hat{w}(\hat{t}) = \Big(\pi^k \big(\hat{w} \circ G_n^{-1}\big)\Big)\big(G_n(\hat{t})\big) \qquad \forall \hat{t} \in \hat{J}. \tag{A.3}$$

Our purpose is to derive stability and error estimates for the projection $\hat{\pi}^k$ and then apply the affine mapping G_n, to obtain estimates (4.4) and (4.5).

We shall start with the following approximation result.

Lemma A.1. *Let $\hat{w} \in L^p(\hat{J}; W_0^{1,p}(\Omega))$, such that $\hat{\partial}_t \hat{w} \in L^p(\hat{J}; W^{1,p}(\Omega))$. Then, there exists a polynomial $\hat{Q}_k \hat{w} \in \mathcal{P}_k(\hat{J}; W_0^{1,p}(\Omega))$, such that*

$$\|\hat{w} - \hat{Q}_k \hat{w}\|_{L^p(\hat{J}; W^{1,p}(\Omega))} \leq \hat{C}_{k,p} \|\hat{\partial}_t^{k+1} \hat{w}\|_{L^p(\hat{J}; W^{1,p}(\Omega))}, \tag{A.4}$$

and

$$\|\hat{\partial}_t(\hat{w} - \hat{Q}_k \hat{w})\|_{L^p(\hat{J}; W^{1,p}(\Omega))} \leq \hat{C}_{k-1,p} \|\hat{\partial}_t^{k+1} \hat{w}\|_{L^p(\hat{J}; W^{1,p}(\Omega))}, \tag{A.5}$$

where $\hat{C}_{k,p}$ is a positive constant independent of $\hat{w}$, but dependent on k and p.

Proof. We select

$$\hat{Q}^k \hat{w}(\hat{t}) = \sum_{i=0}^{k} \frac{(\hat{t}+1)^i}{i!} \hat{\partial}_t^i \hat{w}(-1),$$

i.e., a Taylor expansion of $\hat{w}$, around the point $\hat{t} = -1$. Then, Taylor's theorem shows

$$\hat{w}(\hat{t}) - \hat{Q}^k \hat{w}(\hat{t}) = \int_{-1}^{\hat{t}} \frac{(\hat{t}-s)^k}{k!} \hat{\partial}_t^{k+1} \hat{w}(s)\, \mathrm{d}s, \quad \hat{t} \in \hat{J}.$$

Now, for all $\hat{t} \in \hat{J}$, an application of Hölder's inequality and some straightforward integral calculations show

$$\|\hat{w}(\hat{t}) - \hat{Q}^k \hat{w}(\hat{t})\|_{W^{1,p}(\Omega)} \leq \left(\frac{(\hat{t}+1)^{kp'+1}}{(k!)^{p'}(kp'+1)} \right)^{1/p'} \|\hat{\partial}_t^{k+1} \hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))}.$$

Upon raising the above inequality to the power of p, integrating over $\hat{t} \in \hat{J}$ and calculating a simple integral that occurs, we obtain

$$\|\hat{w} - \hat{Q}^k \hat{w}\|^p_{L^p(\hat{J};W^{1,p}(\Omega))} \leq \frac{2^{(kp'+1)(p-1)+1}}{(k!)^p (kp'+1)^{p-1}((kp'+1)(p-1)+1)} \|\hat{\partial}_t^{k+1} \hat{w}\|^p_{L^p(\hat{J};W^{1,p}(\Omega))}.$$

Thus, the proof of (A.4) is complete. In order to prove (A.5), we assume that $k \geq 1$, since for $k = 0$, the left-hand side of (A.5) is trivially zero, and observe that

$$\hat{\partial}_t \Big(\hat{w}(\hat{t}) - \hat{Q}^k \hat{w}(\hat{t}) \Big) = \int_{-1}^{\hat{t}} \frac{(\hat{t}-s)^{k-1}}{(k-1)!} \hat{\partial}_t^{k+1} \hat{w}(s)\, \mathrm{d}s, \quad \hat{t} \in \hat{J}.$$

Therefore, we proceed as above, to complete the proof. □

We continue with a stability estimate for the projector $\hat{\pi}^k$.

Lemma A.2. *Let $\hat{w} \in C(\overline{\hat{J}}; L^2(\Omega)) \cap L^p(\hat{J}; W_0^{1,p}(\Omega))$, that satisfies $\hat{\partial}_t \hat{w} \in L^p(\hat{J}; W^{1,p}(\Omega))$. Then, there exists a constant $C_{k,p}$, depending only on k and p, such that*

$$\|\hat{\pi}^k \hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))} \leq C_{k,p} \Big(\|\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))} + \|\hat{\partial}_t \hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))} \Big). \tag{A.6}$$

Proof. First of all, the continuous embedding $W^{1,p}(\hat{J}; W^{1,p}(\Omega)) \subset C(\overline{\hat{J}}; W^{1,p}(\Omega))$ shows

$$\|\hat{w}(1)\|_{W^{1,p}(\Omega)} \leq C_p \Big(\|\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))} + \|\hat{\partial}_t \hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))} \Big), \tag{A.7}$$

which completes the proof for $k = 0$, since in that case, we have $\hat{\pi}^k \hat{w}(\hat{t}) = \hat{w}(1)$, for all $\hat{t} \in \hat{J}$. Now, for $k \geq 1$, we denote by $\{L_i\}_{i=0}^k$ the Legendre polynomials up to degree k in $[-1,1]$. It is easy to show, by setting $\hat{\phi}_k = L_i$ into (A.2) and employing the orthogonality condition $(L_i, L_j)_{L^2(\hat{J})} = 2\delta_{ij}/(2j+1)$, that

$$\hat{\pi}^k \hat{w}(\hat{t}) = \sum_{i=0}^{k-1} (L_i(\hat{t}) - L_k(\hat{t})) \hat{w}_i + L_k(\hat{t}) \hat{w}(1) \quad \forall \hat{t} \in \hat{J},$$

where

$$\hat{w}_i = \frac{2i+1}{2} \int_{\hat{J}} L_i(\hat{t}) \hat{w}(\hat{t}) \mathrm{d}\hat{t}, \quad i = 0, \ldots, k-1.$$

Thus, we obtain

$$\|\hat{\pi}^k \hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))} \leq \sum_{i=0}^{k-1} \|L_i - L_k\|_{L^p(\hat{J})} \|\hat{w}_i\|_{W^{1,p}(\Omega)} + \|L_k\|_{L^p(\hat{J})} \|\hat{w}(1)\|_{W^{1,p}(\Omega)}. \tag{A.8}$$

Moreover, an application of Hölder's inequality shows

$$\|\hat{w}_i\|_{W^{1,p}(\Omega)} \le \frac{2i+1}{2}\|\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))}\|L_i\|_{L^{p'}(\hat{J})}. \tag{A.9}$$

We combine (A.7)–(A.9), to infer

$$\begin{aligned}
\|\hat{\pi}^k\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))} &\le \|\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))}\sum_{i=0}^{k-1}\frac{2i+1}{2}\|L_i\|_{L^{p'}(\hat{J})}\|L_i-L_k\|_{L^p(\hat{J})}\\
&\quad + C_p\|L_k\|_{L^p(\hat{J})}\big(\|\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))}+\|\hat{\partial}_t\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))}\big)\\
&\le \left(\sum_{i=0}^{k-1}(2i+1)\|L_i\|_{L^\infty(\hat{J})}\|L_i-L_k\|_{L^\infty(\hat{J})}+2^{1/p}C_p\|L_k\|_{L^\infty(\hat{J})}\right)\|\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))}\\
&\quad + 2^{1/p}C_p\|L_k\|_{L^\infty(\hat{J})}\|\hat{\partial}_t\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))},
\end{aligned}$$

and the proof is complete, after the observation that the Legendre polynomials are bounded by a constant depending only on k. ☐

We now proceed to an error estimate for $\hat{\pi}^k$.

Lemma A.3. *Let $\hat{w}\in C(\overline{\hat{J}};L^2(\Omega))\cap L^p(\hat{J};W^{1,p}(\Omega))$, such that $\hat{\partial}_t^{k+1}\hat{w}\in L^p(\hat{J};W^{1,p}(\Omega))$. Then, there exists a constant $C_{k,p}$, depending only on k and p, such that*

$$\|\hat{w}-\hat{\pi}^k\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))}\le C_{k,p}\|\hat{\partial}_t^{k+1}\hat{\|}_{L^p(\hat{J};W^{1,p}(\Omega))}. \tag{A.10}$$

Proof. Let $\hat{Q}_k\hat{w}$ be the polynomial approximation of $\hat{w}$, from Lemma A.1. Then, an application of the triangle inequality shows

$$\begin{aligned}
\|\hat{w}-\hat{\pi}^k\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))} &\le \|\hat{w}-\hat{Q}_k\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))}+\|\hat{\pi}^k(\hat{Q}_k\hat{w}-\hat{w})\|_{L^p(\hat{J};W^{1,p}(\Omega))}\\
&\le \|\hat{w}-\hat{Q}_k\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))}+C_{k,p}\Big(\|\hat{Q}_k\hat{w}-\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))}\\
&\quad +\|\hat{\partial}_t(\hat{Q}_k\hat{w}-\hat{w})\|_{L^p(\hat{J};W^{1,p}(\Omega))}\Big)\\
&\le C_{k,p}\|\hat{\partial}_t^{k+1}\hat{w}\|_{L^p(\hat{J};W^{1,p}(\Omega))},
\end{aligned}$$

where in the second inequality, we applied the stability bound (A.6), and in the third inequality, we applied the estimates (A.4) and (A.5). Thus, the proof is complete. ☐

It remains to apply the above lemmas, in order to prove Lemma 4.2.

Proof of Lemma 4.2. For the stability estimate (4.4), it holds

$$\begin{aligned}
\|\pi^k w\|^p_{L^p(I;W^{1,p}(\Omega)} &= \sum_{n=1}^N\int_{I_n}\|\pi^k w(t)\|^p_{W^{1,p}(\Omega)}\,\mathrm{d}t \\
&= \sum_{n=1}^N\frac{\tau_n}{2}\int_{\hat{J}}\|\hat{\pi}^k(w\circ G_n)(\hat{t})\|^p_{W^{1,p}(\Omega)}\,\mathrm{d}\hat{t}\\
&\le C_{k,p}\sum_{n=1}^N\frac{\tau_n}{2}\Big(\|(w\circ G_n)\|^p_{L^p(\hat{J};W^{1,p}(\Omega))}+\|\hat{\partial}_t(w\circ G_n)\|^p_{L^p(\hat{J};W^{1,p}(\Omega))}\Big),
\end{aligned} \tag{A.11}$$

where we applied (A.6) for $w\circ G_n$. Now, we have

$$\|w\circ G_n\|^p_{L^p(\hat{J};W^{1,p}(\Omega))}=\int_{\hat{J}}\big\|w(G_n(\hat{t}))\big\|^p_{W^{1,p}(\Omega)}\mathrm{d}\hat{t}=2\tau_n^{-1}\int_{I_n}\|w(t)\|^p_{W^{1,p}(\Omega)}\,\mathrm{d}t, \tag{A.12}$$

and

$$\|\hat{\partial}_t(w \circ G_n)\|^p_{L^p(\hat{J};W^{1,p}(\Omega))} = \frac{\tau_n^p}{2^p}\int_{\hat{J}} \|\partial_t w(G_n(\hat{t}))\|^p_{W^{1,p}(\Omega)}\mathrm{d}\hat{t} = \frac{\tau_n^{p-1}}{2^{p-1}}\int_{I_n} \|\partial_t w(t)\|^p_{W^{1,p}(\Omega)}\,\mathrm{d}t. \tag{A.13}$$

We substitute (A.12) and (A.13) into (A.11), to infer

$$\|\pi^k w\|^p_{L^p(I;W^{1,p}(\Omega)} \le C_{k,p}\sum_{n=1}^N \bigg(\|w\|^p_{L^p(I_n;W^{1,p}(\Omega)} + \frac{\tau_n^p}{2^p}\|\partial_t w\|^p_{L^p(I_n;W^{1,p}(\Omega)}\bigg),$$

which surrenders the desired estimate. It remains to prove the error estimate (4.5). To that end, we observe that

$$\begin{aligned}\|w - \pi^k w\|^p_{L^p(I;W^{1,p}(\Omega))} &= \sum_{n=1}^N \frac{\tau_n}{2}\int_{\hat{J}} \|(w\circ G_n)(\hat{t}) - \hat{\pi}^k(w\circ G_n)(\hat{t})\|^p_{W^{1,p}(\Omega)}\mathrm{d}\hat{t} \\ &\le C_{k,p}\sum_{n=1}^N \frac{\tau_n}{2}\int_{\hat{J}} \|\hat{\partial}_t^{k+1}(w\circ G_n)(\hat{t})\|^p_{W^{1,p}(\Omega)}\mathrm{d}\hat{t},\end{aligned} \tag{A.14}$$

where we used (A.10). A consecutive application of the chain rule shows

$$\hat{\partial}_t^{k+1}(w\circ G_n)(\hat{t}) = \frac{\tau_n^{k+1}}{2^{k+1}}\partial_t^{k+1} w(G_n(\hat{t})).$$

We substitute the above equation into (A.14), to infer

$$\begin{aligned}\|w - \pi^k w\|^p_{L^p(I;W^{1,p}(\Omega)} &\le C_{k,p}\sum_{n=1}^N \frac{\tau_n}{2}\frac{\tau_n^{p(k+1)}}{2^{p(k+1)}}\int_{\hat{J}} \|\partial_t^{k+1} w(G_n(\hat{t}))\|^p_{W^{1,p}(\Omega)}\mathrm{d}\hat{t} \\ &= C_{k,p}\sum_{n=1}^N \frac{\tau_n^{p(k+1)}}{2^{p(k+1)}}\int_{I_n} \|\partial_t^{k+1} w(t)\|^p_{W^{1,p}(\Omega)}\,\mathrm{d}t,\end{aligned}$$

which completes the proof. □