Numerische Mathematik (2019) 143:749–780
https://doi.org/10.1007/s00211-019-01071-5

Numerische Mathematik

Numerical analysis of quasilinear parabolic equations under low regularity assumptions

Eduardo Casas[1] · Konstantinos Chrysafinos[2,3]

Received: 24 July 2018 / Revised: 25 April 2019 / Published online: 30 August 2019

Abstract
In this paper, we carry out the numerical analysis of a class of quasilinear parabolic equations, where the diffusion coefficient depends on the solution of the partial differential equation. The goal is to prove error estimates for the fully discrete equation using discontinuous Galerkin discretization in time DG(0) combined with piecewise linear finite elements in space. This analysis is performed under minimal regularity assumptions on the data. In particular, we omit any assumption regarding existence of a second derivative in time of the solution.

Mathematics Subject Classification Primary 35K59, 65M60; Secondary 35B65

1 Introduction

In this paper, we analyze a fully-discrete scheme for the following quasilinear parabolic equation

$$\begin{cases} \dfrac{\partial u}{\partial t} - \operatorname{div}_x[a(x,t,u(x,t))\nabla_x u] + a_0(x,t,u(x,t)) = f \ \text{ in } Q, \\ u(x,t) = 0 \ \text{ on } \Sigma, \quad u(x,0) = u_0(x) \ \text{ in } \Omega, \end{cases} \tag{1.1}$$

The first author was partially supported by Spanish Ministerio de Economía, Industria y Competitividad under research projects MTM2014-57531-P and MTM2017-83185-P.

✉ Konstantinos Chrysafinos
chrysafinos@math.ntua.gr

Eduardo Casas
eduardo.casas@unican.es

1 Departamento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, Av. Los Castros s/n, 39005 Santander, Spain

2 Department of Mathematics, School of Applied Mathematics and Physical Sciences, National Technical University of Athens, Zografou Campus, 15780 Athens, Greece

3 IACM, FORTH, 20013 Heraklion, Crete, Greece

where $\Omega \subset \mathbb{R}^n$, $n \leq 3$, is a bounded open convex set with a C^2 boundary Γ, $0 < T < \infty$ is given, $Q = \Omega \times (0, T)$ and $\Sigma = \Gamma \times (0, T)$. The following assumptions are assumed.

Assumption 1.1 The function $a : \bar{Q} \times \mathbb{R} \longrightarrow \mathbb{R}$ is continuous and

$$\exists \Lambda > 0 \text{ such that } a(x, t, u) \geq \Lambda \quad \forall (x, t, u) \in \bar{Q} \times \mathbb{R}; \tag{1.2}$$

$$\forall M > 0\, \exists L_M : \forall x_1, x_2 \in \bar{\Omega}, \ \forall t \in \mathbb{R}, \text{ and } \forall |u_1|, |u_2| \leq M$$
$$|a(x_2, t, u_2) - a(x_1, t, u_1)| \leq L_M(|x_2 - x_1| + |u_2 - u_1|). \tag{1.3}$$

Assumption 1.2 We suppose that $a_0 : Q \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function, monotone nondecreasing with respect to u, and satisfying

$$a_0(x, t, 0) = 0 \text{ for a.a. } (x, t) \in Q, \tag{1.4}$$

$$\forall M > 0\, \exists C_M : \forall |u_1|, |u_2| \leq M \text{ and for a.a. } (x, t) \in Q$$
$$|a_0(x, t, u_2) - a_0(x, t, u_1)| \leq L_{0,M}|u_2 - u_1|. \tag{1.5}$$

Under these assumptions, the following maximal parabolic regularity holds (see [9, Corollary 1]):

$$\|u\|_{W_p^{2,1}(Q)} \leq M_p\big(\|f\|_{L^{2p}(0,T;L^p(\Omega))} + \|u_0\|_{W^{2-\frac{2}{p},p}(\Omega)}\big),$$

for all p with $p > \max\{n, 2\}$. Our goal is to prove a fully-discrete error estimate without imposing additional regularity assumptions on the solution u. Our motivation to analyze the problem under $W_p^{2,1}(Q)$ regularity of u follows from control theory related to these equations, where this is typically the maximal regularity that we can expect.

For some earlier works regarding the numerical analysis of quasilinear parabolic equations, we refer the reader to the papers [17,22,24–26]. A key difference of our work compared to these classical results is that in these papers more smoothness on u is required. In contrast to the above mentioned papers, here the solution u of (1.1) has only one derivative in time $\frac{\partial u}{\partial t} \in L^p(Q)$ for some $p > n + 2$. Despite this fact, assuming boundedness and global Lipschitz properties for nonlinear terms $a(.,.,.)$ and $a_0(.,.,.)$, as assumed in all the above papers, we still derive error estimates in $L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H^1(\Omega))$ without imposing any relation between the temporal and spatial discretization parameters τ and h, respectively. Moreover, for $n \leq 2$, assuming that $\tau \approx h^\kappa$ for some $\frac{3}{2-\frac{1}{p}} < \kappa \leq 2$, the error estimates hold under the weaker assumption of local boundedness and local Lipschitz regularity of a and a_0 (Assumptions 1.1 and 1.2). Relationships of this type between h and τ were also assumed in [22] and [24]. The reader is also referred to the papers [14–16] where nonlinearity on ∇u is also considered, and to the works of [1,3] where implicit-explict multistep methods and Galerkin time-stepping methods are studied for a smooth solution u respectively.

Here, we analyze the lowest order DG(0) discontinuous (in time) Galerkin scheme, combined with standard conforming finite element spaces. When time-dependent coefficients in the divergence form are absent, then the scheme can be viewed as the Implicit Euler scheme with an averaging right hand side. In presence of time-dependent coefficients through the nonlinear term, the DG(0) scheme maintains the structure of the weak formulation, which is an essential asset in development of error estimates The choice of the scheme is also dictated by the lack of regularity of our PDE solution.

Our proof for error estimates is based on a construction of a linear space-time projection with time-dependent coefficients, in the spirit of [13]. We show that this projection exhibits optimal approximation rates according to the available regularity. In addition, we also obtain an error estimate in the $L^\infty(Q)$ norm in the case $n \le 2$. This result is crucial to remove the usual boundedness and global Lipschitz regularity of a. To obtain these results we use maximal discrete parabolic regularity recently proved in [20]. The reader is also referred to [2], where the concept of discrete maximal parabolic regularity is considered for semi-discrete (in time) approximations of various time-stepping schemes. Finally, the reader is referred to [10,12,18,21] for the corresponding analysis in the elliptic case.

1.1 Notation

Before finishing this section, let us introduce some notation. Given $1 \le p, q \le \infty$, we denote

$$W^{1,0}_{p,q} = L^p(0,T; W^{1,q}_0(\Omega)) \cap W^{1,p}(0,T; W^{-1,q}(\Omega))$$

and

$$W^{2,1}_{p,q}(Q) = L^p(0,T; W^{2,q}(\Omega) \cap W^1_0(\Omega)) \cap W^{1,p}(0,T; L^q(\Omega)).$$

We take $\|\cdot\|_X = \|\cdot\|_{X_1} + \|\cdot\|_{X_2}$ as the norm in the spaces $X = X_1 \cap X_2$. In the case that $p = q$, we set $W^{2,1}_p(Q) = W^{2,1}_{p,p}(Q)$ and $W^{1,0}_p(Q) = W^{1,0}_{p,p}(Q)$, respectively. Moreover, if $p = q = 2$ we denote $H^{2,1}(Q) = W^{2,1}_2(Q)$ and $W(0,T) = W^{1,0}_2(Q)$ as usual. We also consider the Besov spaces ($1 < q < \infty$ and $1 \le p \le \infty$),

$$W_{q,p}(\Omega) = \left(W^{-1,q}(\Omega), W^{1,q}_0(\Omega)\right)_{1-\frac{1}{p},p}$$

and

$$B_{q,p}(\Omega) = \left(L^q(\Omega), W^{2,q}(\Omega) \cap W^{1,q}_0(\Omega)\right)_{1-\frac{1}{p},p}$$

where $\left(X, Y\right)_{1-\frac{1}{p},p}$ denotes the real interpolation between the Banach spaces X and Y. We recall that for $p = q \ge 2$, the following identities hold (see [8, (14.2.4) and (14.2.5)])

$$W^{1-\frac{2}{p},p}_0(\Omega) = W_{p,p} = \left(W^{-1,p}(\Omega), W^{1,p}_0(\Omega)\right)_{1-\frac{1}{p},p}$$

and

$$W^{2-\frac{2}{p},p}(\Omega) \cap W^{1,p}_0(\Omega) = B_{p,p} = \left(L^p(\Omega), W^{2,p}(\Omega) \cap W^{1,p}_0(\Omega)\right)_{1-\frac{1}{p},p}.$$

We finally quote the following embedding result

$$W_p^{2,1}(Q) \subset W^{(1-\theta),p}(0,T;W^{2\theta,p}(\Omega)), \quad \forall\theta \in [0,1], \tag{1.6}$$

see [5, Theorem 5.2]. Hence, if $p > n+2$ then there exists $\alpha \in (0,1)$ depending on p such that

$$W^{(1-\theta),p}(0,T;W^{2\theta,p}(\Omega)) \subset C^{\alpha}([0,T];C^{1,\alpha}(\bar{\Omega})). \tag{1.7}$$

This can be proved by choosing θ satisfying $(1-\theta)p > 1$ and $(2\theta-1)p > n$ in (1.6) and using the classical Sobolev embeddings, see [2] or [6, Theorem 3] and [4, Theorem III.4.10.2].

Furthermore, we note that:

1. When $n = 2$, selecting $\theta = \frac{1}{4}$ in (1.6) we observe that

$$W_p^{2,1}(Q) \subset W^{\frac{3}{4},p}(0,T;W^{\frac{1}{2},p}(\Omega)).$$

   Hence, we conclude with $\alpha_T = \frac{3}{4} - \frac{2}{p} \in (\frac{1}{2},1)$, and $\alpha_\Omega = \frac{1}{2} - \frac{2}{p} \in (0,1)$ that

$$W_p^{2,1}(Q) \subset C^{\alpha_T}([0,T];C^{\alpha_\Omega}(\bar{\Omega})). \tag{1.8}$$

2. When $n = 3$, choosing $\theta = \frac{3}{10}$, (1.8) holds with $\alpha_T = \frac{7}{10} - \frac{1}{p} \in (\frac{1}{2},1)$ and $\alpha_\Omega = \frac{3}{5} - \frac{3}{p} \in (0,1)$.

2 Existence, uniqueness and regularity

The goal of this section is to state the basic solvability and uniqueness result of solution of (1.1) as well as the maximal parabolic regularity.

Theorem 2.1 [9, Theorem 2.1] *Let us assume that $f \in L^p(0,T;W^{-1,q}(\Omega))$ and $u_0 \in W_{q,p}(\Omega) \cap C_0(\Omega)$ with $p,q \in [2,+\infty)$ satisfying that $\frac{1}{p} + \frac{n}{q} < 1$. Then, (1.1) has a unique solution $u \in C(\bar{Q}) \cap W_{p,q}^{1,0}(Q)$. Moreover, there exists a constant $M_{p,q}$ depending on $\|u_0\|_{C_0(\Omega)}$, $\|f\|_{L^p(0,T;W^{-1,q}(\Omega))}$, p and q such that*

$$\begin{aligned}&\|u\|_{C(\bar{Q})} + \|u\|_{W_{p,q}^{1,0}(Q)} \\ &\quad \le M_{p,q}\big(\|f\|_{L^p(0,T;W^{-1,q}(\Omega))} + \|u_0\|_{W_{q,p}(\Omega)} + \|u_0\|_{C_0(\Omega)}\big).\end{aligned} \tag{2.1}$$

The following lemma, concerning the solvability of the linear parabolic PDE, with space-time dependent coefficients will be also needed.

Lemma 2.2 [9, Lemma 2.2] *Let us assume that $f \in L^p(0,T;W^{-1,q}(\Omega))$ and $u_0 \in W_{q,p}(\Omega)$ with $p,q \in [2,+\infty)$. Given a function $b \in C(\bar{Q})$ with $b(x,t) \ge \Lambda > 0$*

$\forall(x,t)\in Q$, we consider the problem

$$\begin{cases} \dfrac{\partial u}{\partial t} - \mathrm{div}_x[b(x,t)\nabla_x u] = f \;\; in\; Q, \\ u(x,t) = 0 \;\; on\; \Sigma, \quad u(x,0) = u_0(x) \;\; in\; \Omega. \end{cases} \tag{2.2}$$

Then, (2.2) has a unique solution $u \in W^{1,0}_{p,q}(Q)$. Moreover, there exists a constant $C_{p,q}$ depending on p and q such that

$$\|u\|_{W^{1,0}_{p,q}(Q)} \leq C_{p,q}\big(\|f\|_{L^p(0,T;W^{-1,q}(\Omega))} + \|u_0\|_{W_{q,p}(\Omega)}\big). \tag{2.3}$$

We will also use the following lemma concerning maximal parabolic regularity:

Lemma 2.3 [23] *Let $b \in C(\bar{Q})$ satisfy $b(x,t) \geq \Lambda > 0\; \forall(x,t)\in Q$. We assume that $f \in L^p(0,T;L^q(\Omega))$ and $u_0 \in B_{q,p}(\Omega)$ with $1 < p,q < +\infty$. Then the equation*

$$\begin{cases} \dfrac{\partial u}{\partial t} - b(x,t)\Delta_x u = f \;\; in\; Q, \\ u(x,t) = 0 \;\; on\; \Sigma, \quad u(x,0) = u_0(x) \;\; in\; \Omega \end{cases} \tag{2.4}$$

has a unique solution $u \in W^{2,1}_{p,q}(Q)$ and the following inequality holds

$$\|u\|_{W^{2,1}_{p,q}(Q)} \leq M_{p,q}\big(\|f\|_{L^p(0,T;L^q(\Omega))} + \|u_0\|_{B_{q,p}(\Omega)}\big). \tag{2.5}$$

Finally, we state the main theorem, which will be used subsequently.

Theorem 2.4 [9, Corollary 1] *Let us assume that $f \in L^{2p}(0,T;L^p(\Omega))$ with $p > \max\{n,2\}$, and $u_0 \in W^{2-\frac{2}{p},p}(\Omega)\cap W^{1,p}_0(\Omega)$, then the Eq. (1.1) has a unique solution $u \in W^{2,1}_p(Q)$ and there exists a constant $M_p > 0$ depending on $\|u_0\|_{C_0(\Omega)}$ such that*

$$\|u\|_{W^{2,1}_p(Q)} \leq M_p\big(\|f\|_{L^{2p}(0,T;L^p(\Omega))} + \|u_0\|_{W^{2-\frac{2}{p},p}(\Omega)}\big). \tag{2.6}$$

3 Error estimates

In this section we consider a fully-discrete scheme based on the discontinuous Galerkin time-stepping scheme of lowest order (piecewise constants in time) combined with standard conforming finite elements in space. First, we define the scheme, and then we consider error estimates. In this section, in addition to the hypotheses of Sect. 1 we assume the following assumptions:

Assumption 3.1 We assume that $f \in L^{2p}(0,T;L^p(\Omega))$ and $u_0 \in W^{2-\frac{2}{p},p}(\Omega)\cap W^{1,p}_0(\Omega)\cap H^2(\Omega)$ for some $\max\{n+2,4\} < p < \infty$.

Hence, Theorem 2.4 implies that (1.1) has a unique solution $u \in W^{2,1}_p(Q)$ and according to (1.7) there exists $\alpha \in (0,1)$ such that $u \in C^{\alpha}([0,T];C^{1,\alpha}(\bar{\Omega}))$. Finally, we assume local Hölder's regularity of a with respect to time.

Assumption 3.2 There exists $\beta \geq 1/2$ such that

$$\begin{aligned} &\forall M > 0\, \exists \Lambda_M : \forall x \in \Omega,\ \forall t_1, t_2 \in [0, T], \text{ and } \forall |u| \leq M \\ &|a(x, t_1, u) - a(x, t_2, u)| \leq \Lambda_M |t_2 - t_1|^\beta. \end{aligned} \tag{3.1}$$

This assumption is made to apply the results of [20]. Observe that if we define $b(x, t) = a(x, t, u(t, x))$, then using the the triangle inequality, (3.1) and (1.3) with $M = \|u\|_{L^\infty(Q)}$, we infer

$$\begin{aligned} &|b(x, t_1) - b(x, t_2)| \\ &\leq |a(x, t_1, u(x, t_1)) - a(x, t_2, u(x, t_1))| \\ &\quad + |a(x, t_2, u(x, t_1)) - a(x, t_2, u(x, t_2))| \\ &\leq \Lambda_M |t_2 - t_1|^\beta + L_M |u(x, t_1) - u(x, t_2)| \\ &\leq \Lambda_M |t_2 - t_1|^\beta + L_M |t_1 - t_2|^{\alpha_T} \leq \max\{\Lambda_M, L_M\} |t_1 - t_2|^{\beta_T}. \end{aligned} \tag{3.2}$$

where $\beta_T = \min\{\beta, \alpha_T\} \geq \frac{1}{2}$.

3.1 The fully-discrete scheme

First, a family of triangulations $\{\mathcal{K}_h\}_{h>0}$ of $\bar{\Omega}$, is defined in the standard way, e.g. in [8, Chapter 3.3]. With each element $K \in \mathcal{K}_h$, we associate two parameters h_K and ϱ_K, where h_K denotes the diameter of the set K and ϱ_K is the diameter of the largest ball contained in K. Define the size of the mesh by $h = \max_{K \in \mathcal{K}_h} h_K$. We also assume that the following regularity assumptions on the triangulation are satisfied.

(i) There exist two positive constants $\varrho_{\mathcal{K}}$ and $\delta_{\mathcal{K}}$ such that $\frac{h_K}{\varrho_K} \leq \varrho_{\mathcal{K}}$ and $\frac{h}{h_K} \leq \delta_{\mathcal{K}}$ $\forall K \in \mathcal{K}_h$ and $\forall h > 0$.

(ii) Define $\overline{\Omega}_h = \cup_{K \in \mathcal{K}_h} K$, and let Ω_h and Γ_h denote its interior and its boundary, respectively. We assume that the vertices of $\mathcal{K}_h$ placed on the boundary Γ_h are points of Γ.

Since Ω is convex, from the last assumption we have that Ω_h is also convex. Hereafter we assumed that

$$\exists M_\Gamma : \min_{x \in \Gamma} |x - \hat{x}| \leq M_\Gamma h^2 \quad \forall \hat{x} \in \Gamma_h, \tag{3.3}$$

Since Γ is of class C^2, this property holds in the case $n = 2$; see, for instance, [7] or [11].

From (3.3) we deduce the existence of some constant $C_\Gamma > 0$ independent of h such that

$$|\Omega \setminus \Omega_h| \leq Ch^2, \tag{3.4}$$

$$\|u\|_{C(\overline{\Omega \setminus \Omega_h} \times [0,T])} \leq Ch^2. \tag{3.5}$$

The last inequality follows from the fact that $u \in C([0, T]; C^1(\bar{\Omega}))$.

On the mesh $\mathcal{K}_h$ we consider the finite dimensional spaces $U_h \subset H_0^1(\Omega)$ formed by piecewise linear polynomials in Ω_h and vanishing in $\Omega \setminus \Omega_h$.

Throughout the rest of the paper $(\cdot, \cdot)_h$ will denote the scalar product in $L^2(\Omega_h)$:

$$(f_1, f_2)_h = \int_{\Omega_h} f_1(x) f_2(x)\, \mathrm{d}x.$$

For the discretization in time, we consider a grid of points $0 = t_0 < t_1 < \cdots < t_{N_\tau} = T$. We denote $\tau_k = t_k - t_{k-1}$. We make the following assumption

$$\exists \varrho_0 > 0 \text{ such that } \tau = \max_{1 \le k \le N_\tau} \tau_k < \varrho_0 \tau_k \ \ \forall 1 \le k \le N_\tau \text{ and } \forall \tau > 0. \tag{3.6}$$

Given a triangulation $\mathcal{K}_h$ of Ω and a grid of points $\{t_k\}_{k=0}^{N_\tau}$ of $[0, T]$, we set $\sigma = (h, \tau)$ and $Q_\sigma = \Omega_h \times (0, T)$. Finally, we consider the following spaces

$$\mathcal{U}_\sigma = \{u_\sigma \in L^2(0, T; U_h) : u_\sigma|_{(t_{k-1}, t_k]} \in U_h \ \text{ for } \ 1 \le k \le N_\tau\}.$$

We have that the functions of $\mathcal{U}_\sigma$ are piecewise constant in time. Any element of the above space takes the form,

$$u_\sigma = \sum_{k=1}^{N_\tau} u_{h,k} \chi_k = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} u_{j,k} \chi_k e_j \ \text{ with } \ u_{j,k} \in \mathbb{R},$$

where χ_k is the characteristic function of the interval $(t_{k-1}, t_k]$ and $\{e_j\}_{J=1}^{N_h}$ is the nodal basis associated to the interior nodes $\{x_j\}_{j=1}^{N_h}$ in Ω of the triangulation $\mathcal{K}_h$. Therefore, the dimension of $\mathcal{U}_\sigma$ is $N_\tau N_h$.

We briefly recall the definition of the standard projections associated to the dG time-stepping schemes.

Definition 3.3 We define the projection operator $P_h : L^2(\Omega) \longrightarrow U_h$ by

$$(P_h u, w_h)_h = (u, w_h)_h \quad \forall w_h \in U_h.$$

Then, define $P_\sigma : C([0, T], L^2(\Omega)) \longrightarrow \mathcal{U}_\sigma$ by $(P_\sigma u)_{h,k} = P_h u(t_k)$ for $1 \le k \le N_\tau$.

Lemma 3.4 *There exists a constant $C > 0$ independent of σ such that*

$$\begin{aligned} &\|v - P_\sigma v\|_{L^2(0,T;H^1(\Omega_h))} + \|v - P_\sigma v\|_{L^\infty(0,T;L^2(\Omega_h)} \\ &\quad \le C\left(\tau^{1/2} + h\right) \|v\|_{H^{2,1}(Q)} \quad \forall v \in H^{2,1}(Q). \end{aligned} \tag{3.7}$$

If in addition, $v \in W^{2,1}_q(Q)$ with $2 \le q < \infty$, then there exists a constant $C > 0$ independent of σ such that the following estimates hold

$$\begin{cases} \|v - P_\sigma v\|_{L^q(0,T;L^2(\Omega_h))} \le C\left(\tau + h^2\right)\|v\|_{W^{2,1}_q(Q)}, \\ \|v - P_\sigma v\|_{L^\infty(0,T;L^2(\Omega_h))} \le C\left(\tau^{-\frac{1}{q}}(\tau + h^2)\|v\|_{W^{2,1}_q(Q)}\right). \end{cases} \tag{3.8}$$

Proof For the first estimate (3.7), it suffices to prove the inequality

$$\|v - P_\sigma v\|_{L^2(0,T;H^1_0(\Omega_h))} \le C(\tau^{1/2} + h)\|v\|_{H^{2,1}(Q)} \;\; \forall v \in H^{2,1}(Q). \tag{3.9}$$

First of all we observe that for any function $v \in H^{2,1}(Q)$ there exists a sequence $\{v_k\}_{k=1}^\infty \subset H^{2,1}(Q) \cap C^\infty(\bar{Q})$ such that $v_k \to v$ strongly in $H^{2,1}(Q)$. Therefore, if (3.9) is proved for functions $v \in H^{2,1}(Q) \cap C^\infty(\bar{Q})$, then by density it follows that the inequality remains valid for arbitrary elements $v \in H^{2,1}(Q)$. Hence, we will assume that $v \in H^{2,1}(Q) \cap C^\infty(\bar{Q})$. We use the inequality

$$\begin{aligned} &\|v - P_\sigma v\|_{L^2(0,T;H^1_0(\Omega_h))} \\ &\quad \le \left(\sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k}\|v(t) - P_h v(t)\|^2_{H^1(\Omega_h)}\,dt + \sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k}\|P_h v(t) - P_h v(t_k)\|^2_{H^1(\Omega_h)}\,dt\right)^{1/2}. \end{aligned}$$

For the first term, we note that

$$\sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k}\|v(t) - P_h v(t)\|^2_{H^1(\Omega_h)}\,\mathrm{d}t \le Ch^2\sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k}\|v(t)\|^2_{H^2(\Omega_h)}\,\mathrm{d}t.$$

The delicate point is to estimate the second term. Using an inverse inequality it is immediate to check that P_h is stable with respect to the H^1 norm, hence we get

$$\begin{aligned} \int_0^T \|P_h v(t) - P_h v(t_k)\|^2_{H^1(\Omega_h)}\,dt &\le C\sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k}\|v(t) - v(t_k)\|^2_{H^1(\Omega)}\,dt \\ &\le C'\sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k}\|\nabla(v(t) - v(t_k))\|^2_{L^2(\Omega)}\,dt. \end{aligned}$$

To estimate this we proceed as follows:

$$\frac{1}{2}\frac{d}{dt}\|\nabla(v(t) - v(t_k))\|^2_{L^2(\Omega)} = (\nabla v_t(t), \nabla v(t) - \nabla v(t_k))_{L^2(\Omega)},$$

where v_t stands for $\frac{dv}{dt}$. Integrating between s and t_k, for $s \in (t_{k-1}, t_k)$ and using that $v_t = 0$ on Σ, we get

$$
\begin{aligned}
-\frac{1}{2}\|\nabla v(s) - \nabla v(t_k)\|^2_{L^2(\Omega)} &= \int_s^{t_k} (\nabla v_t(t), \nabla v(t) - \nabla v(t_k))_{L^2(\Omega)}\, dt \\
&= -(\nabla v(s), \nabla v(s) - \nabla v(t_k))_{L^2(\Omega)} - \int_s^{t_k} (\nabla v(t), \nabla v_t(t))_{L^2(\Omega)}\, dt \\
&= (\Delta v(s), v(s) - v(t_k))_{L^2(\Omega)} + \int_s^{t_k} (\Delta v(t), v_t(t))_{L^2(\Omega)}\, dt.
\end{aligned}
$$

From here we infer

$$
\begin{aligned}
&\|\nabla v(s) - \nabla v(t_k)\|^2_{L^2(\Omega)} \\
&\quad = -2(\Delta v(s), v(s) - v(t_k))_{L^2(\Omega)} - 2\int_s^{t_k} (\Delta v(t), v_t(t))_{L^2(\Omega)}\, dt \\
&\quad \le 2\Big(\|v(s)\|_{H^2(\Omega)}\|v(s) - v(t_k)\|_{L^2(\Omega)} + \int_{t_{k-1}}^{t_k} \|v(t)\|_{H^2(\Omega)}\|v_t(t)\|_{L^2(\Omega)}\, dt\Big) \\
&\quad \le 2\sqrt{\tau}\|v(s)\|_{H^2(\Omega)}\Big(\int_{t_{k-1}}^{t_k} \|v_t(\theta)\|^2_{L^2(\Omega)}\, d\theta\Big)^{1/2} \\
&\qquad + 2\int_{t_{k-1}}^{t_k} \|v(t)\|_{H^2(\Omega)}\|v_t(t)\|_{L^2(\Omega)}\, dt \\
&\quad \le \tau\|v(s)\|^2_{H^2(\Omega)} + \int_{t_{k-1}}^{t_k} \|v_t(\theta)\|^2_{L^2(\Omega)}\, d\theta \\
&\qquad + \int_{t_{k-1}}^{t_k} \|v(t)\|^2_{H^2(\Omega)}\, dt + \int_{t_{k-1}}^{t_k} \|v_t(t)\|^2_{L^2(\Omega)}\, dt.
\end{aligned}
$$

Now, integrating the above expression between t_{k-1} and t_k and summing in k we obtain

$$
\sum_{k=1}^{N_\tau} \int_{t_{k-1}}^{t_k} \|\nabla v(s) - \nabla v(t_k)\|^2_{L^2(\Omega)}\, ds \le 2\tau\Big(\|v\|^2_{L^2(0,T;H^2(\Omega))} + \|v_t\|^2_{L^2(0,T;L^2(\Omega))}\Big),
$$

which is the desired estimate. Let us prove the second estimate of (3.7). Given an arbitrary $t \in (t_{k-1}, t_k)$, we note that using the definition of the projection P_σ, adding and subtracting appropriate terms, approximation and stability properties of P_h, and Hölder's inequality, we deduce

$$
\begin{aligned}
&\|v(t) - P_\sigma v(t)\|_{L^2(\Omega_h)} \le \|v(t) - P_h v(t)\|_{L^2(\Omega_h)} + \|P_h v(t) - P_h v(t_k)\|_{L^2(\Omega_h)} \\
&\quad \le Ch\|v(t)\|_{H^1(\Omega)} + C\|v(t) - v(t_k)\|_{L^2(\Omega_h)}
\end{aligned}
$$

$$\leq Ch\|v(t)\|_{H^1(\Omega)} + C\|\int_t^{t_k} v_t(s)\ \mathrm{d}s\|_{L^2(\Omega_h)}$$
$$\leq Ch\|v\|_{L^\infty(0,T;H^1(\Omega))} + C\tau_k^{1/2}\|v_t\|_{L^2(0,T;L^2(\Omega))}.$$

In order to prove (3.8), we first note the embedding $W_q^{2,1}(Q) \subset C([0,T]; W^{1,q}(\Omega))$; see, for instance, [6, Theorem 3]. For the first estimate of (3.8) we observe that

$$\|v - P_\sigma v\|_{L^q(0,T;L^2(\Omega_h))} \leq \left(\sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k} \|v(t) - P_h v(t)\|^q_{L^2(\Omega_h)}\,\mathrm{d}t\right)^{1/q}$$
$$+\left(\sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k} \|P_h v(t) - P_h v(t_k)\|^q_{L^2(\Omega_h)} dt\right)^{1/q}$$
$$\leq |\Omega|^{\frac{q-2}{2q}}\left(\int_0^T \|v(t) - P_h v(t)\|^q_{L^q(\Omega_h)}\,dt\right)^{1/q} + \left(\sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k} \|v(t) - v(t_k)\|^q_{L^2(\Omega_h)} dt\right)^{1/q}$$
$$\leq Ch^2\|v\|_{L^q(0,T;W^{2,q}(\Omega))} + \tau^{\frac{q-1}{q}}\left(\sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t_k} \|v_t\|^q_{L^2(\Omega)}\,\mathrm{d}s\,\mathrm{d}t\right)^{1/q}$$
$$\leq Ch^2\|v\|_{W_q^{2,1}(Q)} + |\Omega|^{\frac{q-2}{2q}}\tau^{\frac{q-1}{q}}\left(\sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t_k} \|v_t\|^q_{L^q(\Omega)}\,\mathrm{d}s\,\mathrm{d}t\right)^{1/q}$$
$$\leq Ch^2\|v\|_{W_q^{2,1}(Q)} + |\Omega|^{\frac{q-2}{2q}}\tau\|v_t\|_{L^q(Q)} \leq C(\tau + h^2)\|v\|_{W_q^{2,1}(Q)},$$

where we have used that $\|P_h v\|_{L^2(\Omega)} \leq \|v\|_{L^2(\Omega)}$, the estimate $\|v - P_h v\|_{L^q(\Omega)} \leq Ch^2\|v\|_{W^{2,q}(\Omega)}$, and Hölder's inequality.

For the second estimate of (3.8), we take $t \in (0, T]$, $t \in (t_{k-1}, t_k]$ for some $k = 1, \ldots, N$, and we observe that

$$\|v(t) - P_\sigma v(t)\|_{L^2(\Omega_h)} \leq \|v(t) - v(t_k)\|_{L^2(\Omega)} + \|v(t_k) - P_h v(t_k)\|_{L^2(\Omega_h)}$$
$$\leq \left\|\int_{t_{k-1}}^{t_k} |v_t(s)|\,\mathrm{d}s\right\|_{L^2(\Omega)} + \tau_k^{-\frac{1}{q}}\|v(t_k) - P_h v(t_k)\|_{L^q(t_{k-1},t_k;L^2(\Omega_h))}$$
$$\leq |\Omega|^{\frac{q-2}{2q}}\int_{t_{k-1}}^{t_k} \|v_t\|_{L^q(\Omega)}\,\mathrm{d}t + C\tau^{-\frac{1}{q}}\|v(t_k) - P_h v(t_k)\|_{L^q(t_{k-1},t_k;L^2(\Omega_h))}.$$

For the first term we get with Hölder's inequality that

$$\int_{t_{k-1}}^{t_k} \|v_t\|_{L^q(\Omega)}\,\mathrm{d}t \leq \tau_k^{1-\frac{1}{q}}\|v_t\|_{L^q(t_{k-1},t_k;L^q(\Omega))} \leq \tau^{1-\frac{1}{q}}\|v_t\|_{L^q(Q)}.$$

For the second term we proceed as follows

$$\left(\sum_{k=1}^{N} \|v(t_k) - P_h v(t_k)\|^q_{L^q(t_{k-1},t_k;L^2(\Omega_h))}\right)^{1/q}$$

$$
\begin{aligned}
&\leq \left(\sum_{k=1}^{N} \|v(t_k) - v(t)\|^q_{L^q(t_{k-1},t_k;L^2(\Omega_h))} \right)^{1/q} \\
&\quad + \left(\sum_{k=1}^{N} \|v(t) - P_h v(t)\|^q_{L^q(t_{k-1},t_k;L^2(\Omega_h))} \right)^{1/q} \\
&\quad + \left(\sum_{k=1}^{N} \|P_h v(t) - P_h v(t_k)\|^q_{L^q(t_{k-1},t_k;L^2(\Omega_h))} \right)^{1/q} \\
&\leq 2 \left(\sum_{k=1}^{N} \|v(t_k) - v(t)\|^q_{L^q(t_{k-1},t_k;L^2(\Omega_h))} \right)^{1/q} \\
&\quad + \left(\sum_{k=1}^{N} \|v(t) - P_h v(t)\|^q_{L^q(t_{k-1},t_k;L^2(\Omega_h))} \right)^{1/q} \\
&\leq C(\tau + h^2)\|v\|_{W^{2,1}_q(Q)}
\end{aligned}
$$

where the last inequality was proved above. Combining the last two estimates we obtain the desired estimate.

Now, we proceed in the definition of the scheme. The discrete state equation is given by

$$
\begin{cases}
\text{For } k = 1, \ldots, N_\tau, \quad \text{and } \forall w_h \in U_h, \\
(u_{h,k} - u_{h,k-1}, w_h)_h + \displaystyle\int_{t_{k-1}}^{t_k} \Big[\big(a(x,t,u_{h,k})\nabla_x u_{h,k}, \nabla_x w_h\big)_h \\
\qquad\qquad + (a_0(x,t,u_{h,k}), w_h)_h \Big] \, \mathrm{d}t = (f_k, w_h)_{\tau_k,h}, \\
u_{h,0} = u_{0h},
\end{cases}
\tag{3.10}
$$

where $u_{0h} = P_h u_0$ and

$$
(f_k, w_h)_{\tau_k,h} = \int_{t_{k-1}}^{t_k} (f(t), w_h)_h \, \mathrm{d}t. \tag{3.11}
$$

The following estimate is well known:

$$
\|u_0 - u_{0h}\|_{L^2(\Omega_h)} \leq Ch^2 \|u_0\|_{H^2(\Omega)}. \tag{3.12}
$$

Note that when only the Laplacian term is involved, i.e., $a(x, t; .) =: 1$, the DG(0) formulation (3.10) is the standard implicit Euler scheme, with right hand side f_k computed by the integral $\int_{t_{k-1}}^{t_k} (f(t), w_h)_h \, \mathrm{d}t$.

The existence of a solution of (3.10) follows easily for every τ small enough by using Brower's fixed point theorem. The uniqueness is deduced in the standard way with a discrete Grönwall inequality. In addition, we have the following stability estimates.

Lemma 3.5 *There exists a constant C independent of σ, such that for τ small enough the fully-discrete solution u_σ satisfies*

$$\|u_\sigma\|_{L^\infty(0,T;L^2(\Omega))} + \|u_\sigma\|_{L^2(0,T;H_0^1(\Omega))} \le C.$$

Proof We note that setting $w_h = u_{h,k}$ into (3.10), and using (1.2), (1.4), and (3.21) we easily obtain

$$\begin{aligned}
&(1/2)\|u_{h,k}\|^2_{L^2(\Omega)} + \Lambda \int_{t_{k-1}}^{t_k} \|\nabla u_{h,k}\|^2_{L^2(\Omega)}\,\mathrm{d}t \\
&\quad \le (1/2)\|u_{h,k-1}\|^2_{L^2(\Omega)} + M\int_{t_{k-1}}^{t_k} \|u_{h,k}\|^2_{L^2(\Omega)}\,\mathrm{d}t + \int_{t_{k-1}}^{t_k} \|f(t)\|_{L^2(\Omega)}\|u_{h,k}\|_{L^2(\Omega)}\,\mathrm{d}t \\
&\quad \le (1/2)\|u_{h,k-1}\|^2_{L^2(\Omega)} + \left(M + \frac{1}{2}\right)\tau_k\|u_{h,k}\|^2_{L^2(\Omega)} + \frac{1}{2}\int_{t_{k-1}}^{t_k} \|f(t)\|^2_{L^2(\Omega)}\,\mathrm{d}t.
\end{aligned}$$

From here the statement of the lemma follows easily using the Discrete Grönwall Lemma.

3.2 An auxiliary projection with time-dependent coefficients

One of the main technical tools that allows us to handle the limited regularity of the involved variables, is an auxiliary parabolic equation with time-dependent coefficients in the elliptic part which plays the role of a global space-time projection onto $\mathcal{U}_\sigma$. This technique allows us to circumvent any differentiation of the nonlinear term. Hence, to derive the basic error estimate, we do not assume more regularity on the solution apart the one guaranteed from Theorem 2.1.

In the sequel associated to the solution $u \in W_p^{2,1}(Q)$ of (1.1), we define $u_\sigma^p \in \mathcal{U}_\sigma$ such that

$$\begin{cases}
\text{For } k = 1, \ldots, N_\tau, \quad \text{and } \forall w_h \in U_h, \\
\left(u^p_{h,k} - u^p_{h,k-1}, w_h\right)_h + \int_{t_{k-1}}^{t_k} \left(a(x,t,u(x,t))\nabla_x u^p_{h,k}, \nabla_x w_h\right)_h \mathrm{d}t \\
\quad = \int_{t_{k-1}}^{t_k} (F(t), w_h)_h\,\mathrm{d}t, \\
u^p_{h,0} = u_{0h},
\end{cases} \tag{3.13}$$

where $u_{0h} = P_h u_0$ was already introduced and

$$F(x,t) \equiv \frac{\partial u}{\partial t} - \mathrm{div}_x[a(x,t,u(x,t))\nabla_x u]. \tag{3.14}$$

We note that due to the boundedness of u and (1.1) we have that $F(x,t) \equiv f(x,t) - a_0(x,t,u(x,t)) \in L^{2p}(0,T;L^p(\Omega))$.

Theorem 3.6 *Suppose that $u_\sigma^p \in \mathcal{U}_\sigma$ is the solution of (3.13). Then, there exists a positive constant C independent of σ such that,*

$$\|u-u_\sigma^p\|_{L^\infty(0,T;L^2(\Omega_h))}+\|u-u_\sigma^p\|_{L^2(0,T;H^1(\Omega_h))}\leq C\left(\tau^{1/2}+h\right),$$
$$\|u-u_\sigma^p\|_{L^2(0,T;L^2(\Omega_h))}\leq C\left(\tau+h^2\right).$$

Proof *Step 1: Estimates at partition points and at the energy norm* We denote by $e_p=u_\sigma^p-u$. Recalling the definition F and integrating by parts we get

$$\begin{cases}\text{for } k=1,\ldots,N_\tau, \ \text{ and } \forall w_h\in U_h,\\ \left(e_p(t_k)-e_p(t_{k-1}),w_h\right)_h+\displaystyle\int_{t_{k-1}}^{t_k}\left(a(x,t,u(x,t))\nabla_x e_p,\nabla_x w_h\right)_h\,\mathrm{d}t=0,\\ e_p(0)=u_{0h}-u_0.\end{cases}\tag{3.15}$$

Now, we decompose the error $e_p=u_\sigma^p-u=(u_\sigma^p-P_\sigma u)+(P_\sigma u-u)\equiv r_\sigma+r_p$, where $P_\sigma u$ is introduced in Definition 3.3. Since we already have the estimates for r_p, we only need to establish the corresponding ones to r_σ. Working similar to [13, Section 2] we infer from the Definition 3.3 that $(r_p(t_k),w_h)_h=(r_p(t_{k-1}),w_h)_h=0$. Hence, (3.15) can be written as:

$$\begin{cases}\text{For } k=1,\ldots,N_\tau, \ \text{ and } \forall w_h\in U_h,\\ \left(r_{h,k}-r_{h,k-1},w_h\right)_h+\displaystyle\int_{t_{k-1}}^{t_k}\left(a(x,t,u(x,t))\nabla_x r_{h,k},\nabla_x w_h\right)_h\,\mathrm{d}t\\ \quad=-\displaystyle\int_{t_{k-1}}^{t_k}\left(a(x,t,u(x,t))\nabla_x r_p(t),\nabla_x w_h\right)_h\,\mathrm{d}t,\\ r_{h,0}=0.\end{cases}\tag{3.16}$$

Setting $w_h=r_{h,k}$ into (3.16) and using Assumption 1.1, we obtain with (3.7)

$$\begin{aligned}\|r_\sigma\|_{L^\infty(0,T;L^2(\Omega_h))}+\|r_\sigma\|_{L^2(0,T;H^1(\Omega_h))}&\leq C\|r_p\|_{L^2(0,T;H^1(\Omega_h))}\\ &\leq C\left(\tau^{1/2}+h\right)\|u\|_{H^{2,1}(Q)}.\end{aligned}$$

Now, using (3.7), Lemma 3.4 and the triangle inequality the first estimate of the theorem is obtained.

Step 2: Estimate in $L^2(0,T;L^2(\Omega_h))$

For this estimate we employ a duality argument. Our goal is to derive an improved estimate: $\|r_\sigma\|_{L^2(Q)}\leq C(\tau+h^2)\|u\|_{H^{2,1}(Q)}$. Then the $L^2(Q)$ estimate for u_σ^p-u follows by triangle inequality and (3.8). For this purpose, we consider the following problem:

$$\begin{cases}-\dfrac{\partial\phi}{\partial t}-\mathrm{div}_x[a(x,t,u(x,t))\nabla_x\phi]=r_\sigma \ \text{ in } Q,\\ \phi(x,t)=0 \ \text{ on } \Sigma, \quad \phi(x,T)=0 \ \text{ in } \Omega,\end{cases}\tag{3.17}$$

We note that since $r_\sigma \in L^\infty(0,T;L^2(\Omega))$, we can apply Lemma 2.3, with $p=q=2$, $b(x,t)=a(x,t,u(x,t))$, and

$$f = r_\sigma - \nabla_x a(x,t,u)\nabla_x\phi - \frac{\partial a(x,t,u)}{\partial u}\nabla_x u \nabla_x \phi$$

to get $\phi \in H^{2,1}(Q)$ and the estimate $\|\phi\|_{H^{2,1}(Q)} \le C\|r_\sigma\|_{L^2(Q)}$. Indeed, first note that using (1.6) and (1.7) and Lemma 2.2, we deduce from (3.17) that $\phi \in W_2^{1,0}(Q)$. Now, observe that $\nabla_x a(x,t,u) - \frac{\partial a(x,t,u)}{\partial u}\nabla_x u$ belongs to $L^\infty(Q)$ due to Assumption 1.1 and the regularity $u \in W_p^{2,1}(Q)$ which implies that $u \in C^\alpha([0,T];C^{1,\alpha}(\bar\Omega))$.. Therefore, $f \in L^2(Q)$, which implies that $\phi \in H^{2,1}(Q)$. The fully-discrete version of (3.17), within the discontinuous Galerkin time-stepping framework takes the form: we seek $\phi_\sigma \in \mathcal{U}_\sigma$ such that

$$\begin{cases} \text{for } k = N_\tau,\dots,1, \quad \text{and } \forall w_h \in U_h, \\ (\phi_{h,k} - \phi_{h,k+1}, w_h)_h + \displaystyle\int_{t_{k-1}}^{t_k} (a(x,t,u(x,t))\nabla_x\phi_\sigma, \nabla_x w_h)_h \, \mathrm{d}t \\ \qquad\qquad = \displaystyle\int_{t_{k-1}}^{t_k} (r_{h,k}, w_h)_h \, \mathrm{d}t, \\ \phi_{h,N_\tau+1} = 0. \end{cases} \tag{3.18}$$

For the backwards in time problem we set $\phi_\sigma(t_{k-1}) = \phi_{h,k}$ for every $1 \le k \le N_\tau$. Now, using the estimate obtained for $\|\phi\|_{H^{2,1}(Q)}$ and arguing in the standard way we get

$$\|\phi - \phi_\sigma\|_{L^2(0,T;H^1(\Omega_h))} \le C\left(\tau^{1/2} + h\right)\|r_\sigma\|_{L^2(Q)}.$$

Setting $w_h = r_{h,k}$ in (3.18), we obtain

$$\begin{aligned} &(\phi_{h,k} - \phi_{k+1,h}, r_{h,k})_h + \int_{t_{k-1}}^{t_k} \big(a(x,t,u(x,t))\nabla\phi_{h,k}, \nabla_x r_{h,k}\big)_h \, \mathrm{d}t \\ &= \int_{t_{n-1}}^{t_n} \|r_{h,k}\|^2_{L^2(\Omega)} \, \mathrm{d}t. \end{aligned}$$

Putting $w_h = \phi_{h,k}$ into (3.16) it follows

$$\begin{cases} \text{for } k = 1,\dots,N_\tau, \quad \text{and } \forall w_h \in U_h, \\ \big(r_{h,k} - r_{h,k-1}, \phi_{h,k}\big)_h + \displaystyle\int_{t_{k-1}}^{t_k} \big(a(x,t,u(x,t))\nabla_x r_{h,k}, \nabla_x\phi_{h,k}\big)_h \, \mathrm{d}t \\ = -\displaystyle\int_{t_{k-1}}^{t_k} \big(a(x,t,u(x,t))\nabla_x r_p(t), \nabla_x\phi_{h,k}\big)_h \, \mathrm{d}t, \\ r_{h,0} = 0. \end{cases}$$

Subtracting the last two equalities, summing the resulting equalities, and using the final and initial conditions for ϕ_σ and r_σ we deduce

$$
\begin{aligned}
0 &= (r_{h,0}, \phi_{h,1})_h - (\phi_{h,N_\tau+1}, r_{h,N_\tau})_h \\
&= \int_0^T \|r_\sigma\|^2_{L^2(\Omega)}\, \mathrm{d}t + \int_0^T \big(a(x,t,u(x,t))\nabla_x r_p(t), \nabla_x \phi_\sigma\big)_h\, \mathrm{d}t \\
&= \int_0^T \|r_\sigma\|^2_{L^2(\Omega)}\, \mathrm{d}t + \int_0^T \big(a(x,t,u(x,t))\nabla_x r_p(t), \nabla_x \phi_\sigma - \nabla_x\phi(t)\big)_h\, \mathrm{d}t \\
&\quad + \int_0^T \big(a(x,t,u(x,t))\nabla_x r_p(t), \nabla_x \phi(t)\big)_h\, \mathrm{d}t.
\end{aligned}
$$

Hence using Assumption 1.1, and integrating by parts in space in the last term, the above equality leads to

$$
\begin{aligned}
&\|r_\sigma\|^2_{L^2(Q)} \le C\|r_p\|_{L^2(0,T;H^1(\Omega))}\|\phi - \phi_\sigma\|_{L^2(0,T;H^1(\Omega_h))} \\
&\quad + \left|\int_0^T \left[\int_{\Omega_h} r_p \mathrm{div}_x\, (a(x,t,u(x,t))\nabla_x\phi)\, \mathrm{d}x\, \mathrm{d}t + \int_{\Gamma_h} r_p a(x,t,u(x,t))\partial_{\nu_h}\phi\, \mathrm{d}x\right]\mathrm{d}t\right|.
\end{aligned}
$$

Now, recalling (3.8) and the estimate for $\phi - \phi_\sigma$, we infer from (3.17) and the above inequality

$$
\begin{aligned}
&\|r_\sigma\|^2_{L^2(Q)} \le C\|r_p\|_{L^2(0,T;H^1(\Omega_h))}\|\phi - \phi_\sigma\|_{L^2(0,T;H^1(\Omega_h))} \\
&\quad + \int_0^T \int_{\Omega_h} |r_p(t)\,(\phi_t(t) + r_\sigma)\,|\, \mathrm{d}x\, \mathrm{d}t + \int_{\Gamma_h} |r_p(x,t)||a(x,t,u(x,t))||\partial_{\nu_h}\phi|\, \mathrm{d}x\, \mathrm{d}t \\
&\le C\|r_p\|_{L^2(0,T;H^1(\Omega))}\|\phi - \phi_\sigma\|_{L^2(0,T;H^1(\Omega))} \\
&\quad + \|r_p\|_{L^2(Q)}\big(\|\phi_t\|_{L^2(Q)} + \|r_\sigma\|_{L^2(Q)}\big) + C\|u\|_{L^\infty(\Gamma_h\times(0,T))}\|\partial_{\nu_h}\phi\|_{L^1(\Gamma_h\times(0,T))} \\
&\le C\left((\tau^{1/2}+h)^2 + (\tau + h^2)\right)\|r_\sigma\|_{L^2(Q)}.
\end{aligned}
$$

In the last inequality, we have used the L^∞ stability of the L^2 projection P_h (see [19]), (3.5), and

$$
\|\partial_{\nu_h}\phi\|_{L^1(\Gamma_h\times(0,T))} \le C\|\phi_h\|_{L^1(0,T;H^2(\Omega_h))} \le C\|r_\sigma\|_{L^2(Q)}.
$$

Thus we have proved

$$
\|r_\sigma\|_{L^2(Q)} \le C(\tau + h^2)\|u\|_{H^{2,1}(Q)}. \tag{3.19}
$$

The desired estimate now follows by triangle inequality.

Next we want to prove an L^∞ estimate for $u - u^p_\sigma$.

Theorem 3.7 *Suppose that $u_\sigma^p \in \mathcal{U}_\sigma$ is the solution of (3.13), and let $u \in W_p^{2,1}(Q)$. Suppose also that the assumptions 3.1 and 3.2 hold. Then, there exists a positive constant C depending on the domain and the data, but independent of σ such that,*

$$\|u - u_\sigma^p\|_{L^p(0,T;L^2(\Omega_h))} \le C \ln\Big(\frac{T}{\tau}\Big)(\tau + h^2),$$
$$\|u - u_\sigma^p\|_{L^\infty(0,T;L^2(\Omega_h))} \le C \ln\Big(\frac{T}{\tau}\Big)\tau^{-\frac{1}{p}}(\tau + h^2),$$
$$\|u - u_\sigma^p\|_{L^\infty(0,T;L^\infty(\Omega_h))} \le C\Big(h^{\frac{3}{2}-\frac{n}{p}} + \ln\Big(\frac{T}{\tau}\Big)\tau^{-\frac{1}{p}}h^{-\frac{n}{2}}(\tau + h^2)\Big).$$

Proof We consider $b(x,t) \equiv a(x,t,u(x,t))$ and we define the problem,

$$\begin{cases} \dfrac{\partial v}{\partial t} - \mathrm{div}_x[b(x,t)\nabla_x v] = F & \text{in } Q,\\ v(x,t) = 0 \text{ on } \Sigma, \quad v(x,0) = v_0 & \text{in } \Omega, \end{cases}$$

where F is defined by (3.14). Obviously u is the unique solution of this problem, and u_σ^p is the DG(0) approximation of the above problem. From [20, Corollary 4.7], we deduce the first estimate for $u - u_\sigma^p$. It is easy to check the assumptions of [20] are fullfilled. In particular, assumption [20, (1.4)] follows from (3.2). The assumption $A_{k,m}u_0 \in L^2(\Omega)$ of [20, Theorem 3.12] which holds because $u_0 \in H^2(\Omega)$. In addition, [20, Assumption 1] requires that $b \in L^\infty(0,T;W^{1,\infty}(\Omega))$, which holds due to (1.6)-(1.7) and our Assumption (1.3).

For the remaining estimates, we write $u - u_\sigma^p = u - P_\sigma u + P_\sigma u - u_\sigma^p$. Note that by Lemma 3.4 we obtain

$$\|u - P_\sigma u\|_{L^\infty(0,T;L^2(\Omega_h))} \le C\tau^{-\frac{1}{p}}(\tau + h^2)\|u\|_{W_p^{2,1}(Q)}.$$

For the term $P_\sigma u - u_\sigma^p$ observe that an inverse estimate and the triangle inequality along with (3.8) and the proved estimate for $\|u - u_\sigma^p\|_{L^p(0,T;L^2(\Omega_h))}$ lead to

$$\begin{aligned}\|u_\sigma^p - P_\sigma u\|_{L^\infty(0,T;L^2(\Omega_h))} &\le \frac{C}{\tau^{1/p}}\|u_\sigma^p - P_\sigma u\|_{L^p(0,T;L^2(\Omega_h))}\\ &\le \frac{C}{\tau^{1/p}}\Big(\|u_\sigma^p - u\|_{L^p(0,T;L^2(\Omega_h))} + \|u - P_\sigma u\|_{L^p(0,T;L^2(\Omega_h))}\Big)\\ &\le C\ln\Big(\frac{T}{\tau}\Big)\frac{1}{\tau^{1/p}}(\tau + h^2). \end{aligned} \tag{3.20}$$

For the estimate $u - u_\sigma^p$ in $L^\infty(Q)$ we first prove an estimate for $\|u - P_\sigma u\|_{L^\infty(Q)}$. For this purpose note that, due to [6, Theorem 3], we deduce that $W_p^{2,1}(Q) \subset C(0,T;W^{3/2,p}(\Omega))$. Hence,

$$\begin{aligned}&\|u - P_\sigma u\|_{L^\infty(\Omega_h\times(0,T))}\\ &\quad\le \max_{t\in[0,T]}\Big(\|u(t) - P_h u(t)\|_{L^\infty(\Omega_h)} + \|P_h u(t) - P_\sigma u(t)\|_{L^\infty(\Omega_h)}\Big)\end{aligned}$$

$$
\begin{aligned}
&\leq Ch^{\frac{3}{2}-\frac{n}{p}}\|u\|_{L^\infty(0,T;W^{3/2,p}(\Omega))} + \frac{C}{h^{\frac{n}{2}}}\max_{t\in[0,T]}\|P_h u(t) - P_\sigma u(t)\|_{L^2(\Omega_h)} \\
&\leq Ch^{\frac{3}{2}-\frac{n}{p}}\|u\|_{W_p^{2,1}(Q)} \\
&\quad + \frac{C}{h^{\frac{n}{2}}}\left(\max_{t\in[0,T]}(\|P_h u(t) - u(t)\|_{L^2(\Omega_h)} + \|u(t) - P_\sigma u(t)\|_{L^2(\Omega_h)})\right) \\
&\leq Ch^{\frac{3}{2}-\frac{n}{p}}\|u\|_{W_p^{2,1}(Q)} + \frac{2C}{h^{\frac{n}{2}}}\|u(t) - P_\sigma u(t)\|_{L^\infty(0,T;L^2(\Omega_h))} \\
&\leq Ch^{\frac{3}{2}-\frac{n}{p}}\|u\|_{W_p^{2,1}(Q)} + \frac{2C}{h^{\frac{n}{2}}}\ln\left(\frac{T}{\tau}\right)\frac{1}{\tau^{1/p}}(\tau + h^2)\|u\|_{W_p^{2,1}(Q)}
\end{aligned}
$$

where at the last step we have used (3.20). For the term $P_\sigma u - u_\sigma^p$ we proceed analogously. Indeed,

$$
\begin{aligned}
\|P_\sigma u - u_\sigma^p\|_{L^\infty(\Omega_h\times(0,T))} &\leq \frac{C}{h^{\frac{n}{2}}}\|P_\sigma u - u_\sigma^p\|_{L^\infty(0,T;L^2(\Omega_h))} \\
&\leq \frac{2C}{h^{\frac{n}{2}}}\ln\left(\frac{T}{\tau}\right)\frac{1}{\tau^{1/p}}(\tau + h^2)\|u\|_{W_p^{2,1}(Q)}.
\end{aligned}
$$

3.3 Error estimates for the fully discrete scheme

Initially we make the following assumption regarding the nonlinear terms a and a_0. This assumption will be removed at the end of this section in the case $n \leq 2$, which is one of the goals of our work.

Assumption 3.8 The nonlinear terms satisfy for some constant M and all $x, x_1, x_2 \in \bar{\Omega}$, $t \in [0, T]$ and $u, u_1, u_2 \in \mathbb{R}$

$$
\begin{cases}
|a(x,t,u)| \leq M, \quad |a_0(x,t,u)| \leq M, \\
|a(x_1,t,u_1) - a(x_2,t,u_2)| \leq M(|x_1 - x_2| + |u_1 - u_2|), \\
|a_0(x_1,t,u_1) - a_0(x_2,t,u_2)| \leq M(|x_1 - x_2| + |u_1 - u_2|).
\end{cases} \tag{3.21}
$$

Now, we proceed with the first estimate.

Theorem 3.9 *Let u, and $u_\sigma \in \mathcal{U}_\sigma$ be the solutions of (1.1) and (3.10) respectively. Then, the following estimate holds, under the assumption 3.8:*

$$
\|u - u_\sigma\|_{L^\infty(0,T;L^2(\Omega_h))} + \|u - u_\sigma\|_{L^2(0,T;H^1(\Omega_h))} \leq C(\tau^{1/2} + h). \tag{3.22}
$$

Proof First, we derive the orthogonality condition by subtracting (1.1) from (3.10). Denoting by $e = u_\sigma - u$, the orthogonality condition can be written as follows:

$$\begin{cases} \text{For } k = 1, \dots, N_\tau, \quad \text{and } \forall w_h \in U_h, \\ (e(t_k) - e(t_{k-1}), w_h)_h \\ + \displaystyle\int_{t_{k-1}}^{t_k} \left[\big(a(x,t,u_{h,k})\nabla_x u_{h,k}, \nabla_x w_h\big)_h - (a(x,t,u(x,t))\nabla_x u, \nabla_x w_h)_h \right] \mathrm{d}t \\ + \displaystyle\int_{t_{k-1}}^{t_k} \left[(a_0(x,t,u_{h,k}), w_h)_h - (a_0(x,t,u(x,t)), w_h)_h \right] \mathrm{d}t = 0, \\ e(0) = u_{0h} - u_0. \end{cases} \tag{3.23}$$

Now, we split the error $e = (u_\sigma - u_\sigma^p) + (u_\sigma^p - u) \equiv d_\sigma + e_p$, where u_σ^p is defined by (3.13), with F defined by (3.14). Hence, inserting $e = d_\sigma + e_p$ into (3.23), adding and subtracting the terms $\Big(a(x,t,u_{h,k})\nabla_x u_{h,k}^p, \nabla_x w_h\Big)_h$ and $\Big(a(x,t,u(x,t))\nabla_x u_{h,k}^p, \nabla_x w_h\Big)_h$ we obtain,

$$\begin{cases} \text{for } k = 1, \dots, N_\tau, \quad \text{and } \forall w_h \in U_h, \\ \big(d_{h,k} - d_{h,k-1}, w_h\big)_h + \big(e_{k,p} - e_{k-1,p}, w_h\big)_h \\ + \displaystyle\int_{t_{k-1}}^{t_k} \left[\big(a(x,t,u_{h,k})\nabla_x u_{h,k}, \nabla_x w_h\big)_h - \Big(a(x,t,u_{h,k})\nabla_x u_{h,k}^p, \nabla_x w_h\Big)_h \right] \mathrm{d}t \\ + \displaystyle\int_{t_{k-1}}^{t_k} \left[\Big(a(x,t,u(x,t))\nabla_x u_{h,k}^p, \nabla_x w_h\Big)_h - (a(x,t,u(x,t)\nabla_x u, \nabla_x w_h)_h \right] \mathrm{d}t \\ + \displaystyle\int_{t_{k-1}}^{t_k} \left[\Big(a(x,t,u_{h,k})\nabla_x u_{h,k}^p, \nabla_x w_h\Big)_h - \Big(a(x,t,u(x,t))\nabla_x u_{h,k}^p, \nabla_x w_h\Big)_h \right] \mathrm{d}t \\ + \displaystyle\int_{t_{k-1}}^{t_k} \left[(a_0(x,t,u_{h,k}), w_h)_h - (a_0(x,t,u(x,t)), w_h)_h \right] \mathrm{d}t = 0, \\ d_{h,0} = 0, e_p(0) = u^p(0) - u(0). \end{cases}$$

Due to (3.15) the last equality is written as:

$$\begin{cases} \text{For } k = 1, \dots, N_\tau, \quad \text{and } \forall w_h \in U_h, \\ \big(d_{h,k} - d_{h,k-1}, w_h\big)_h \\ + \displaystyle\int_{t_{k-1}}^{t_k} \left[\big(a(x,t,u_{h,k})\nabla_x u_{h,k}, \nabla_x w_h\big)_h - \Big(a(x,t,u_{h,k})\nabla_x u_{h,k}^p, \nabla_x w_h\Big)_h \right] \mathrm{d}t \\ + \displaystyle\int_{t_{k-1}}^{t_k} \left[\Big(a(x,t,u_{h,k})\nabla_x u_{h,k}^p, \nabla_x w_h\Big)_h - \Big(a(x,t,u(x,t))\nabla_x u_{h,k}^p, \nabla_x w_h\Big)_h \right] \mathrm{d}t \\ + \displaystyle\int_{t_{k-1}}^{t_k} \left[(a_0(x,t,u_{h,k}), w_h)_h - (a_0(x,t,u(x,t)), w_h)_h \right] \mathrm{d}t = 0, \\ d_{h,0} = 0. \end{cases} \tag{3.24}$$

First, we need to set $w_h = d_{h,k}$. For the first term we note that,

$$
\begin{aligned}
&\int_{t_{k-1}}^{t_k} \Big[\Big(a(x,t,u_{h,k})\nabla_x u_{h,k}, \nabla_x d_{h,k}\Big)_h - \Big(a(x,t,u_{h,k})\nabla_x u^p_{h,k}, \nabla_x d_{h,k}\Big)_h\Big]\,\mathrm{d}t \\
&\quad = \int_{t_{k-1}}^{t_k} \big(a(x,t,u_{h,k})\nabla_x d_{h,k}, \nabla_x d_{h,k}\big)_h\,\mathrm{d}t \\
&\quad \geq \int_{t_{k-1}}^{t_k} \Lambda \|\nabla_x d_{h,k}\|^2_{L^2(\Omega_h)}\,\mathrm{d}t.
\end{aligned}
\tag{3.25}
$$

For the second term, adding and subtracting appropriate terms, using Assumption 3.8, Hölders' inequalities, the fact that $u \in L^\infty(0,T;W^{1,6}(\Omega))$, Young's inequalities, and the Gagliardo-Nirenberg inequality, we infer

$$
\begin{aligned}
&\int_{t_{k-1}}^{t_k} \Big[\Big(a(x,t,u_{h,k})\nabla_x u^p_{h,k}, \nabla_x d_{h,k}\Big)_h - \Big(a(x,t,u)\nabla_x u^p_{h,k}, \nabla_x d_{h,k}\Big)_h\Big]\,\mathrm{d}t \\
&= \int_{t_{k-1}}^{t_k} \Big(\big[a(x,t,u_{h,k}) - a(x,t,u)\big]\nabla_x(u^p_{h,k} - u), \nabla_x d_{h,k}\Big)_h\,\mathrm{d}t \\
&\quad + \int_{t_{k-1}}^{t_k} \big(\big[a(x,t,u_{h,k}) - a(x,t,u)\big]\nabla_x u, \nabla_x d_{h,k}\big)_h\,\mathrm{d}t \\
&\leq C\int_{t_{k-1}}^{t_k}\int_{\Omega_h} \Big[|\nabla_x(u^p_{h,k} - u)||\nabla_x d_{h,k}| + |u_{h,k} - u||\nabla_x u||\nabla_x d_{h,k}|\Big]\,\mathrm{d}x\,\mathrm{d}t \\
&\leq C\Big[\int_{t_{k-1}}^{t_k} \|\nabla_x e_p\|_{L^2(\Omega_h)}\|\nabla_x d_{h,k}\|_{L^2(\Omega_h)}\,\mathrm{d}t \\
&\quad + \int_{t_{k-1}}^{t_k} \Big(\|d_{h,k}\|_{L^3(\Omega_h)} + \|e_p\|_{L^3(\Omega_h)}\Big)\|\nabla_x u\|_{L^6(\Omega)}\|\nabla_x d_{h,k}\|_{L^2(\Omega_h)}\,\mathrm{d}t\Big] \\
&\leq C\Big[\int_{t_{k-1}}^{t_k} \|\nabla_x e_p\|_{L^2(\Omega_h)}\|\nabla_x d_{h,k}\|_{L^2(\Omega_h)}\,\mathrm{d}t \\
&\quad + \|u\|_{L^\infty(0,T;W^{1,6}(\Omega))}\int_{t_{k-1}}^{t_k} \Big(\|d_{h,k}\|_{L^3(\Omega_h)} + \|e_p\|_{L^3(\Omega_h)}\Big)\|\nabla_x d_{h,k}\|_{L^2(\Omega_h)}\,\mathrm{d}t\Big] \\
&\leq C\Big[\int_{t_{k-1}}^{t_k} \|\nabla_x e_p\|_{L^2(\Omega_h)}\|\nabla_x d_{h,k}\|_{L^2(\Omega_h)}\,\mathrm{d}t + \|u\|_{L^\infty(0,T;W^{1,6}(\Omega))}\times \\
&\quad \times \int_{t_{k-1}}^{t_k} \Big(\|\nabla_x d_{h,k}\|^{1/2}_{L^2(\Omega_h)}\|d_{h,k}\|^{1/2}_{L^2(\Omega_h)} + \|\nabla_x e_p\|^{1/2}_{L^2(\Omega_h)}\|e_p\|^{1/2}_{L^2(\Omega_h)}\Big)\|\nabla_x d_{h,k}\|_{L^2(\Omega_h)}\,\mathrm{d}t\Big] \\
&\leq \int_{t_{k-1}}^{t_k} \Big[C\Big[\|e_p\|^2_{H^1(\Omega_h)} + \|d_{h,k}\|^2_{L^2(\Omega_h)} + \|e_p\|^2_{L^2(\Omega_h)}\Big] + (\Lambda/2)\|d_{h,k}\|^2_{H^1(\Omega_h)}\Big]\,\mathrm{d}t \\
&\leq \int_{t_{k-1}}^{t_k} \Big[C\|e_p\|^2_{H^1(\Omega_h)} + (\Lambda/2)\|d_{h,k}\|^2_{H^1(\Omega_h)}\Big]\,\mathrm{d}t + C\tau_k\|d_{h,k}\|^2_{L^2(\Omega_h)}.
\end{aligned}
\tag{3.26}
$$

For the semilinear term, observe that

$$
\int_{t_{k-1}}^{t_k} \big[(a_0(x,t,u_{h,k}), d_{h,k})_h - (a_0(x,t,u(x,t)), d_{h,k})_h\big]\,\mathrm{d}t
$$

$$\leq \int_{t_{k-1}}^{t_k} \left[\|e_p + d_{h,k}\|_{L^2(\Omega_h)}\right] \|d_{h,k}\|_{L^2(\Omega_h)} \,\mathrm{d}t$$
$$\leq C\tau_k \|d_{h,k}\|^2_{L^2(\Omega_h)} + \int_{t_{k-1}}^{t_k} \|e_p\|^2_{L^2(\Omega_h)} \,\mathrm{d}t. \tag{3.27}$$

Substituting $w_h = d_{h,k}$ into (3.24) and using (3.25), (3.26), and (3.27) we obtain

$$\begin{cases} \text{for } k = 1, \ldots, N_\tau, \\ \|d_{h,k}\|^2_{L^2(\Omega_h)} - \|d_{h,k-1}\|^2_{L^2(\Omega_h)} + (\Lambda/2) \displaystyle\int_{t_{k-1}}^{t_k} \|\nabla_x d_{h,k}\|^2_{L^2(\Omega_h)} \,\mathrm{d}t \\ \leq C \displaystyle\int_{t_{k-1}}^{t_k} \|e_p\|^2_{H^1(\Omega_h)} \,\mathrm{d}t + \int_{t_{k-1}}^{t_k} \|e_p\|^2_{L^2(\Omega_h)} \,\mathrm{d}t + C\tau_k \|d_{h,k}\|^2_{L^2(\Omega_h)} \\ d_{h,0} = 0. \end{cases} \tag{3.28}$$

Here C is constant depending upon $\|u\|_{L^\infty(0,T;W^{1,6}(\Omega))}$. Hence, using a discrete Grönwall Lemma and the estimate of Theorem 3.6, we deduce that

$$\|d_\sigma\|_{L^\infty(0,T;L^2(\Omega_h))} + \|d_\sigma\|_{L^2(0,T;H^1(\Omega_h))} \leq C(\tau^{1/2} + h)\|u\|_{W^{2,1}_p(Q)}. \tag{3.29}$$

The estimate of $u - u_\sigma = d_\sigma + e_p$ now follows from the above inequality and Theorem 3.6.

Remark 3.10 Checking the proof of Theorem 3.9 we observe that we have used only the regularity $u \in H^{2,1}(Q) \cap L^\infty(0, T; W^{1,6}(\Omega))$. Additionally we have used the first inequality of Theorem 3.6 which holds for $u \in H^{2,1}(Q) \cap L^\infty(Q)$. The rates obtained in Theorem 3.9 are optimal under the above regularity assumptions on u, and there is no need to impose any restriction between the spacial and temporal discretization parameters. However, the Theorem was proved under the Assumption 3.8. In order to remove this Assumption we are going to prove $u_\sigma \to u$ in $L^\infty(0, T; L^\infty(\Omega_h))$.

3.4 Improved estimates in $L^2(Q)$

The scope of this section is the derivation of an improved estimate in the $L^2(Q)$ norm using an appropriate duality argument. Recall the notation introduced in the proof of Theorem 3.9: $u_\sigma - u = (u_\sigma - u^p_\sigma) + (u^p_\sigma - u) \equiv d_\sigma + e_p$.

We define $\psi \in H^{2,1}(Q)$ (due to Theorem 2.1), as the solution of,

$$\begin{cases} -\dfrac{\partial \psi}{\partial t} - \mathrm{div}_x[a(x,t,u)\nabla_x \psi] + \theta(x,t,u_\sigma,u^p_\sigma)\nabla_x u \nabla_x \psi \\ \quad + \eta(x,t,u_\sigma,u^p_\sigma)\psi = d_\sigma \ \text{ in } Q = \Omega \times (0,T), \\ \psi(x,t) = 0 \ \text{ on } \Sigma = \Gamma \times (0,T), \quad \psi(x,T) = 0 \ \text{ in } \Omega, \end{cases} \tag{3.30}$$

$$\eta(x,t,r,s) = \begin{cases} \dfrac{a_0(x,t,r) - a_0(x,t,s)}{r - s} & \text{if } r \neq s, \\ 0 \quad \text{otherwise,} \end{cases}$$

and

$$\theta(x,t,r,s)=\begin{cases}\dfrac{a(x,t,r)-a(x,t,s)}{r-s} & \text{if } r\neq s,\\ 0 \quad \text{otherwise.}\end{cases}$$

Note that due to the Assumption 3.8, the functions η and θ are bounded. Taking into account the regularity of $d_\sigma \in L^4(0,T;L^4(\Omega))\cap L^\infty(0,T;L^2(\Omega))$ and the boundedness of $a(.,.,.)$ and $a_0(.,.,.)$ and the fact that $\nabla_x u\in L^\infty(Q)$ we have that the problem (3.30) is well defined and $\psi \in W_2^{1,0}(Q)$. Its weak formulation takes the form:

$$\begin{cases}-(\psi_t,w)+(a(x,t,u)\nabla_x\psi,\nabla_x w)+(\theta(x,t,u_\sigma,u_\sigma^p)\nabla_x u\nabla_x\psi,w)\\ \qquad +(\eta(x,t,u_\sigma,u_\sigma^p)\psi,w)=(d_\sigma,w)\quad \forall w\in H_0^1(\Omega),\\ \psi(T)=0.\end{cases} \tag{3.31}$$

A boot-strap argument implies (via Lemma 2.2) that $\psi\in H^{2,1}(Q)$. The discontinuous Galerkin time-stepping scheme is to seek $\psi_\sigma \in \mathcal{U}_\sigma$, such that for $k=N_\tau,\dots,1$ and for all $w_h\in U_h$,

$$\begin{cases}(\psi_{h,k}-\psi_{k+1,h},w_h)_h+\displaystyle\int_{t_{k-1}}^{t_k}\big(a(x,t,u)\nabla_x\psi_{h,k},\nabla_x w_h\big)_h\,\mathrm{d}t\\ +\displaystyle\int_{t_{k-1}}^{t_k}(\theta(x,t,u_\sigma,u_\sigma^p)\nabla_x u\nabla_x\psi_{h,k},w_h)_h\,\mathrm{d}t\\ +\displaystyle\int_{t_{k-1}}^{t_k}\big(\eta(x,t,u_\sigma,u_\sigma^p)\psi_{h,k},w_h\big)_h\,\mathrm{d}t=\int_{t_{k-1}}^{t_k}(d_{n,h},w_h)_h\,\mathrm{d}t,\\ \psi_{h,N_\tau+1}=0.\end{cases} \tag{3.32}$$

For the backward in time problem we set $\psi_\sigma(t_{k-1})=\psi_{h,k}$ for every $1\le k\le N_\tau$. Recall that, $\|\psi\|_{H^{2,1}(Q)}\le C\|d_\sigma\|_{L^2(Q)}$, and working similarly to Theorem 3.6 (see also [13, Section 2]) we deduce

$$\|\psi-\psi_\sigma\|_{L^\infty(0,T;L^2(\Omega))}+\|\psi-\psi_\sigma\|_{L^2(0,T;H^1(\Omega))}\le C\left(\tau^{1/2}+h\right)\|d_\sigma\|_{L^2(Q)}. \tag{3.33}$$

Now, we are ready to state the main result.

Theorem 3.11 *Let $n=2$ and the assumptions 3.1 and 3.2 hold. Suppose that u, ψ, and $u_\sigma,\psi_\sigma\in\mathcal{U}_\sigma$ are the solutions of (1.1), (3.31), (3.10), and (3.32) respectively. Then, there exists a positive constant C independent of σ, depending on $\|u\|_{W_p^{2,1}(Q)}$ such that*

$$\|u-u_\sigma\|_{L^2(0,T;L^2(\Omega_h))}\le C(\tau+h^2). \tag{3.34}$$

Proof Setting $w_h=d_{h,k}=d_\sigma(t_k)$ in (3.32), we obtain,

$$(\psi_{h,k}-\psi_{k+1,h},d_{h,k})_h+\int_{t_{k-1}}^{t_k}\big(a(x,t,u)\nabla\psi_{h,k},\nabla_x d_{h,k}\big)_h\,\mathrm{d}t$$
$$+\int_{t_{k-1}}^{t_k}(\theta(x,t,u_\sigma,u_\sigma^p)\nabla_x u\nabla_x\psi_{h,k},d_{h,k})_h\,\mathrm{d}t$$

$$+\int_{t_{k-1}}^{t_k} (\eta(x,t,u_\sigma,u_\sigma^p)\psi_{h,k}, d_{h,k})_h \,\mathrm{d}t = \int_{t_{k-1}}^{t_k} \|d_{h,k}\|^2_{L^2(\Omega_h)} \,\mathrm{d}t. \tag{3.35}$$

Setting $w_h = \psi_{h,k}$ into (3.24) we obtain: For $k = 1, \ldots, N_\tau$

$$\begin{cases} \left(d_{h,k} - d_{h,k-1}, \psi_{h,k}\right)_h + \int_{t_{k-1}}^{t_k} \left(a(x,t,u_{h,k})\nabla_x d_{h,k}, \nabla_x \psi_{h,k}\right)_h \,\mathrm{d}t \\ + \int_{t_{k-1}}^{t_k} \Big[\Big(a(x,t,u_{h,k})\nabla_x u_{h,k}^p, \nabla_x \psi_{h,k}\Big)_h \\ \quad - \Big(a(x,t,u(x,t))\nabla_x u_{h,k}^p, \nabla_x \psi_{h,k}\Big)_h \Big] \mathrm{d}t \\ + \int_{t_{k-1}}^{t_k} \Big[(a_0(x,t,u_{h,k}), \psi_{h,k})_h - (a_0(x,t,u(x,t)), \psi_{h,k})_h \Big] \mathrm{d}t = 0, \\ d_{h,0} = 0. \end{cases} \tag{3.36}$$

Subtracting (3.36) from (3.35), we obtain: For $k = 1, \ldots, N_\tau$,

$$\begin{aligned} &\left(d_{h,k} - d_{h,k-1}, \psi_{h,k}\right)_h - (\psi_{h,k} - \psi_{k+1,h}, d_{h,k})_h \\ &\quad + \int_{t_{k-1}}^{t_k} \Big[\left(a(x,t,u)\nabla \psi_{h,k}, \nabla_x d_{h,k}\right)_h - \left(a(x,t,u_{h,k})\nabla_x d_{h,k}, \nabla_x \psi_{h,k}\right)_h \Big] \mathrm{d}t \\ &\quad + \int_{t_{k-1}}^{t_k} \Big[\Big(a(x,t,u_{h,k})\nabla_x u_{h,k}^p, \nabla_x \psi_{h,k}\Big)_h - \Big(a(x,t,u(x,t))\nabla_x u_{h,k}^p, \nabla_x \psi_{h,k}\Big)_h \Big] \mathrm{d}t \\ &\quad - \int_{t_{k-1}}^{t_k} (\theta(x,t,u_\sigma,u_\sigma^p)\nabla_x u \nabla_x \psi_{h,k}, d_{h,k})_h \,\mathrm{d}t \\ &\quad + \int_{t_{k-1}}^{t_k} \big[(a_0(x,t,u_{h,k}), \psi_{h,k}) - (a_0(x,t,u(x,t)), \psi_{h,k})_h \big] \,\mathrm{d}t \\ &\quad - \int_{t_{k-1}}^{t_k} (\eta(x,t,u_\sigma,u_\sigma^p)\psi_{h,k}, d_{h,k})_h \,\mathrm{d}t = - \int_{t_{k-1}}^{t_k} \|d_{h,k}\|^2_{L^2(\Omega_h)} \,\mathrm{d}t. \end{aligned} \tag{3.37}$$

Note that $\psi_{N_\tau+1} = 0 = d_{h,0} = 0$, hence using the definition of $\theta(.,.,.,.)$ and $\eta(.,.,.,.)$, and adding and subtracting appropriate terms we obtain upon summation,

$$\begin{aligned} &- \int_0^T \|d_\sigma\|^2_{L^2(\Omega_h)} \,\mathrm{d}t \\ &= \int_0^T \big[(a(x,t,u)\nabla \psi_\sigma, \nabla_x d_\sigma)_h - (a(x,t,u_\sigma)\nabla_x d_\sigma, \nabla_x \psi_\sigma)_h \big] \,\mathrm{d}t \\ &\quad + \int_0^T \left((a(x,t,u_\sigma) - a(x,t,u(x,t)))\, \nabla_x u_\sigma^p, \nabla_x \psi_\sigma \right)_h \,\mathrm{d}t \\ &\quad - \int_0^T \left(\left(a(x,t,u_\sigma) - a(x,t,u_\sigma^p)\right) \nabla_x u, \nabla_x \psi_\sigma \right)_h \,\mathrm{d}t \\ &\quad + \int_0^T (a_0(x,t,u_\sigma^p), \psi_\sigma)_h - (a_0(x,t,u(x,t)), \psi_\sigma)_h \,\mathrm{d}t \end{aligned}$$

$$
\begin{aligned}
&= \int_0^T \left[(a(x,t,u)\nabla\psi_\sigma, \nabla_x d_\sigma)_h - (a(x,t,u_\sigma)\nabla_x d_\sigma, \nabla_x \psi_\sigma)_h \right] \mathrm{d}t \\
&\quad + \int_0^T \left(\left(a(x,t,u_\sigma) - a(x,t,u_\sigma^p) \right) \nabla_x (u_\sigma^p - u), \nabla_x \psi_\sigma \right)_h \mathrm{d}t \\
&\quad + \int_0^T (a(x,t,u_\sigma^p) - (a(x,t,u(x,t))\nabla_x u_\sigma^p, \nabla_x \psi_\sigma)_h \, \mathrm{d}t \\
&\quad + \int_0^T [(a_0(x,t,u_\sigma^p), \psi_\sigma)_h - (a_0(x,t,u(x,t)), \psi_\sigma)_h] \, \mathrm{d}t \equiv \sum_{i=1}^4 I_i. \qquad (3.38)
\end{aligned}
$$

It remains to bound the last four integrals. Starting from I_1, adding and subtracting appropriate terms, using Assumption 3.8 and Hölders inequalities, and the error estimates (3.33) of $\psi - \psi_\sigma$ and (3.29) of d_σ, and setting again $e = u_\sigma - u$, we infer

$$
\begin{aligned}
|I_1| &= \left| \int_0^T \int_{\Omega_h} (a(x,t,u) - a(x,t,u_\sigma)) \nabla_x \psi_\sigma \nabla_x d_\sigma \, \mathrm{d}x \, \mathrm{d}t \right| \\
&\le \int_0^T \int_{\Omega_h} |(a(x,t,u) - a(x,t,u_\sigma)) \nabla_x d_\sigma \nabla_x (\psi_\sigma - \psi)| \, \mathrm{d}x \, \mathrm{d}t \\
&\quad + \int_0^T \int_{\Omega_h} |(a(x,t,u) - a(x,t,u_\sigma) \nabla_x d_\sigma \nabla_x \psi| \, \mathrm{d}x \, \mathrm{d}t \\
&\le C \|d_\sigma\|_{L^2(0,T;H^1(\Omega_h))} (\tau^{1/2} + h) \|\psi\|_{H^{2,1}(Q)} \\
&\quad + C \|e\|_{L^4(0,T;L^4(\Omega_h))} \|d_\sigma\|_{L^2(0,T;H^1(\Omega_h))} \|\psi\|_{L^4(0,T;W^{1,4}(\Omega))} \\
&\le C(\tau^{1/2} + h)^2 \|\psi\|_{H^{2,1}(Q)} \le C(\tau + h^2) \|d_\sigma\|_{L^2(0,T;L^2(\Omega_h)},
\end{aligned}
$$

where C depends on $\|u\|_{W_p^{2,1}(Q)}$. At the last two steps we have also used the stability bound $\|\psi\|_{H^{2,1}(Q)} \le C\|d_\sigma\|_{L^2(Q)}$, and the error estimate of Theorem 3.7 which implies with Gagliardo inequality that

$$
\|e\|_{L^4(0,T;L^4(\Omega_h))} \le C \|e\|^{1/2}_{L^\infty(0,T;L^2(\Omega_h))} \|e\|^{1/2}_{L^2(0,T;H^1(\Omega_h))} \le C(\tau^{1/2} + h).
$$

For I_2, adding and subtracting appropriate terms, and working similarly to I_1,

$$
\begin{aligned}
|I_2| &= \left| \int_0^T \left(\left(a(x,t,u_\sigma) - a(x,t,u_\sigma^p) \right) \nabla_x (u_\sigma^p - u), \nabla_x \psi_\sigma \right)_h \mathrm{d}t \right| \\
&\le \int_0^T \left| \left(\left(a(x,t,u_\sigma) - a(x,t,u_\sigma^p) \right) \nabla_x (u_\sigma^p - u), \nabla_x (\psi_\sigma - \psi) \right)_h \right| \mathrm{d}t \\
&\quad + \int_0^T \left| \left(\left(a(x,t,u_\sigma) - a(x,t,u_\sigma^p) \right) \nabla_x (u_\sigma^p - u), \nabla_x \psi \right)_h \right| \mathrm{d}t \\
&\le C \|e_p\|_{L^2(0,T;H^1(\Omega_h))} \|\psi_\sigma - \psi\|_{L^2(0,T;H^1(\Omega)_h)} \\
&\quad + \|d_\sigma\|_{L^4(0,T;L^4(\Omega_h))} \|e_p\|_{L^2(0,T;H^1(\Omega_h))} \|\psi\|_{L^4(0,T;W^{1,4}(\Omega))}
\end{aligned}
$$

$$
\begin{aligned}
&\leq C(\tau^{1/2}+h)^2\|\psi\|_{H^{2,1}(Q)} + (\tau^{1/2}+h)^2\|\psi\|_{L^4(0,T;W^{1,4}(\Omega))} \\
&\leq C(\tau+h^2)\|d_\sigma\|_{L^2(0,T;L^2(\Omega_h))}.
\end{aligned}
$$

where C depends on $\|u\|_{W_p^{2,1}(Q)}$. The estimate for d_σ in $L^4(0,T;L^4(\Omega))$ is deduced by (3.29) and the Gagliardo inequality as we did before for e. For integral I_3, we proceed as follows: First, we split the integral into two parts, adding and subtracting appropriate terms

$$
\begin{aligned}
|I_3| &\leq \left|\int_0^T \big((a(x,t,u_\sigma^p) - a(x,t,u(x,t)))\nabla_x u_\sigma^p, \nabla_x(\psi_\sigma-\psi)\big)_h \,\mathrm{d}t\right| \\
&\quad + \left|\int_0^T \big((a(x,t,u_\sigma^p) - a(x,t,u(x,t)))\nabla_x u_\sigma^p, \nabla_x\psi\big)_h \,\mathrm{d}t\right| \\
&\leq \int_0^T \left|\big((a(x,t,u_\sigma^p) - a(x,t,u(x,t)))\nabla_x (u_\sigma^p-u), \nabla_x(\psi_\sigma-\psi)\big)_h\right| \mathrm{d}t \\
&\quad + \int_0^T \left|\big((a(x,t,u_\sigma^p) - a(x,t,u(x,t)))\nabla_x u, \nabla_x(\psi_\sigma-\psi)\big)_h\right| \mathrm{d}t \\
&\quad + \int_0^T \left|\big((a(x,t,u_\sigma^p) - a(x,t,u(x,t)))\nabla_x (u_\sigma^p-u), \nabla_x\psi\big)_h\right| \mathrm{d}t \\
&\quad + \int_0^T \left|\big((a(x,t,u_\sigma^p) - a(x,t,u(x,t)))\nabla_x u, \nabla_x\psi\big)_h\right| \mathrm{d}t \\
&\equiv I_{3,1} + I_{3,2} + I_{3,3} + I_{3,4}.
\end{aligned}
$$

For the integral $I_{3,1}$, using the boundedness of $a(\cdot,\cdot,\cdot)$, according to Assumption 3.8,

$$
\begin{aligned}
I_{3,1} &\leq \int_0^T \|u_\sigma^p - u\|_{H^1(\Omega_h)}\|\psi_\sigma-\psi\|_{H^1(\Omega_h)}\,\mathrm{d}t \\
&\leq C\|e_p\|_{L^2(0,T;H^1(\Omega_h))}\|\psi_\sigma-\psi\|_{L^2(0,T;H^1(\Omega_h))} \\
&\leq C(\tau^{1/2}+h)^2\|\psi\|_{H^{2,1}(Q)} \leq C(\tau+h)^2\|d_\sigma\|_{L^2(0,T;L^2(\Omega_h))}.
\end{aligned}
$$

For the integral $I_{3,2}$, using Hölder's inequality the stability bound on $\|u\|_{L^\infty(0,T;W^{1,\infty}(\Omega))} \leq C$, we obtain:

$$
\begin{aligned}
I_{3,2} &\leq \int_0^T \|u_\sigma^p - u\|_{L^2(\Omega_h)}\|u\|_{W^{1,\infty}(\Omega)}\|\psi_\sigma-\psi\|_{H^1(\Omega_h)}\,\mathrm{d}t \\
&\leq \|u\|_{L^\infty(0,T;W^{1,\infty}(\Omega))}\|e_p\|_{L^2(0,T;L^2(\Omega_h))}\|\psi_\sigma-\psi\|_{L^2(0,T;H^1(\Omega_h))} \\
&\leq C(\tau+h^2)\|d_\sigma\|_{L^2(0,T;L^2(\Omega_h))}.
\end{aligned}
$$

The integral $I_{3,3}$ can be bounded using Theorem 3.6 and Gagliardo inequality, $\|e_p\|_{L^4(0,T;L^2(\Omega_h))} \leq \|e_p\|^{1/2}_{L^\infty(0,T;L^2(\Omega_h))}\|e_p\|^{1/2}_{L^\infty(0,T;H^1(\Omega_h))}$, the estimate $\|\psi\|_{L^4(0,T;W^{1,4}(\Omega))} \leq C\|\psi\|_{H^{2,1}(Q)}$,

$$
\begin{aligned}
I_{3,3} &\leq \|u_\sigma^p - u\|_{L^4(0,T;L^2(\Omega_h))}\|\nabla_x(u_\sigma^p - u)\|_{L^2(0,T;H^1(\Omega_h))}\|\psi\|_{L^4(0,T;W^{1,4}(\Omega))}\\
&\leq C(\tau + h^2)\|d_\sigma\|_{L^2(0,T;L^2(\Omega_h))}.
\end{aligned}
$$

For the integral $I_{3,4}$ observe that,

$$
\begin{aligned}
I_{3,4} &\leq \|u_\sigma^p - u\|_{L^2(0,T;L^2(\Omega_h))}\|u\|_{L^\infty(0,T;W^{1,\infty}(\Omega))}\|\psi\|_{L^2(0,T;H^1(\Omega))}\\
&\leq C(\tau + h^2)\|d_\sigma\|_{L^2(0,T;L^2(\Omega_h))}.
\end{aligned}
$$

Finally he last integral I_4 can be estimated with Theorem 3.6 as follows:

$$
\begin{aligned}
|I_4| &= \left|\int_0^T [(a_0(x,t,u_\sigma^p), \psi_\sigma)_h - (a_0(x,t,u(x,t)), \psi_\sigma)_h]\,\mathrm{d}t\right|\\
&\leq C\int_0^T \|u_\sigma^p - u\|_{L^2(\Omega_h)}\|\psi_\sigma\|_{L^2(\Omega_h)}\,\mathrm{d}t\\
&\leq C\|u_p^\sigma - u\|_{L^2(0,T;L^2(\Omega_h))}\|d_\sigma\|_{L^2(0,T;L^2(\Omega_h))} \leq C(\tau + h^2)\|d_\sigma\|_{L^2(0,T;L^2(\Omega_h))}.
\end{aligned}
$$

Inserting the bounds of I_1, I_2, I_3 and I_4 into (3.38) we obtain the estimate

$$
\|d_\sigma\|_{L^2(0,T;L^2(\Omega_h))} \leq C(\tau + h^2). \tag{3.39}
$$

The estimate (3.34) follows from the above inequality, the fact that $u_\sigma - u = d_\sigma + e_p$ and Theorem 3.6.

3.5 Removing the Assumption 3.8 in the case $n < 3$

The goal of this section is to remove the restrictive Assumption 3.8 in the case $n = 1$ or 2. To this end, we first prove the convergence $u_\sigma \to u$ in $L^\infty(Q)$.

Theorem 3.12 *Suppose that $n = 2$. Under the assumptions of Theorems 3.9 and 3.11, the following statements hold: There exists a constant C depending on $\|u\|_{W_p^{2,1}(Q)}$ such that,*

$$
\begin{aligned}
&\|u_\sigma - u\|_{L^\infty(0,T;L^2(\Omega_h))} \leq C\ln\left(\frac{T}{\tau}\right)h^{\kappa(2-\frac{1}{p})-2}, \ \textit{if } \tau \approx h^\kappa, \textit{ with } \kappa \leq 2,\\
&\lim_{\sigma\to 0}\|u_\sigma - u\|_{L^\infty(0,T;L^\infty(\Omega_h))} = 0, \qquad \textit{when } \tau \approx h^\kappa \textit{ with } \frac{3}{2-\frac{1}{p}} < \kappa \leq 2.
\end{aligned}
$$

Proof Similar to Theorem 3.9, we employ an appropriate duality argument and a bootstrap argument. Once again we split $u - u_\sigma = (u - u_\sigma^p) + (u_\sigma^p - u_\sigma) = e_p + d_\sigma$. Due to the fact that d_σ is piecewise constant in time, it is enough to obtain an estimate for the terminal point. If the superconvergent estimate is obtained at the terminal point, then we can employ a identical procedure via a duality argument to get an estimate at $d_{h,N_\tau-1}$ etc. Now, we define $\psi \in W(0,T)$ as the solution of,

$$\begin{cases} -\dfrac{\partial \psi}{\partial t} - \mathrm{div}_x[a(x,t,u_\sigma)\nabla_x \psi] + \theta(x,t,u_\sigma,u_\sigma^p)\nabla_x u \nabla_x \psi \\ \qquad + \eta(x,t,u_\sigma,u_\sigma^p)\psi = 0 \ \text{ in } Q = \Omega \times (0,T), \\ \psi(x,t) = 0 \ \text{ on } \Sigma = \Gamma \times (0,T), \quad \psi(x,T) = d_{h,N_\tau} \ \text{ in } \Omega, \end{cases} \tag{3.40}$$

where the auxiliary functions $\eta(x,t,\cdot,\cdot)$ and $\theta(x,t,\cdot,\cdot)$ are defined as in Theorem 3.11. The weak formulation associated to the above problem takes the form:

$$\begin{cases} -(\psi_t, w) + (a(x,t,u_\sigma)\nabla_x \psi, \nabla_x w) + (\theta(x,t,u_\sigma,u_\sigma^p)\nabla_x u \nabla_x \psi, w) \\ \qquad + (\eta(x,t,u_\sigma,u_\sigma^p)\psi, w) = 0 \quad \forall w \in H_0^1(\Omega), \\ \psi(T) = d_{h,N_\tau}. \end{cases} \tag{3.41}$$

The discontinuous Galerkin time-stepping scheme is to seek $\psi_\sigma \in \mathcal{U}_\sigma$, such that

$$\begin{cases} \text{for } k = N_\tau, \ldots, 1, \ \text{ and } \forall w_h \in U_h, \\ (\psi_{h,k} - \psi_{k+1,h}, w_h)_h + \displaystyle\int_{t_{k-1}}^{t_k} \big(a(x,t,u_{h,k})\nabla_x \psi_{h,k}, \nabla_x w_h\big)_h \, \mathrm{d}t \\ + \displaystyle\int_{t_{k-1}}^{t_k} (\theta(x,t,u_\sigma^p,u_\sigma)\nabla_x u \nabla_x \psi_{h,k}, w_h)_h \, \mathrm{d}t \\ + \displaystyle\int_{t_{k-1}}^{t_k} \big(\eta(x,t,u_\sigma^p,u_\sigma)\psi_{h,k}, w_h\big)_h \, \mathrm{d}t = 0 \\ \psi_{h,N_\tau+1} = d_{h,N_\tau}. \end{cases} \tag{3.42}$$

Setting $w_h = d_\sigma$ in (3.42), we obtain

$$\begin{aligned} &(\psi_{h,k} - \psi_{h,k+1}, d_{h,k})_h + \int_{t_{k-1}}^{t_k} \big(a(x,t,u_{h,k})\nabla \psi_{h,k}, \nabla_x d_{h,k}\big)_h \, \mathrm{d}t \\ &\quad + \int_{t_{k-1}}^{t_k} (\theta(x,t,u_\sigma,u_\sigma^p)\nabla_x u \nabla_x \psi_{h,k}, d_{h,k})_h \, \mathrm{d}t \\ &\quad + \int_{t_{k-1}}^{t_k} (\eta(x,t,u_\sigma,u_\sigma^p)\psi_{h,k}, d_{h,k})_h \, \mathrm{d}t = 0. \end{aligned} \tag{3.43}$$

Returning back to (3.24) and setting $w_h = \psi_{h,k}$, we obtain:

$$\begin{cases} \text{For } k = 1, \ldots, N_\tau, \\ \big(d_{h,k} - d_{h,k-1}, \psi_{h,k}\big)_h + \displaystyle\int_{t_{k-1}}^{t_k} \big(a(x,t,u_{h,k})\nabla_x d_{h,k}, \nabla_x \psi_{h,k}\big)_h \, \mathrm{d}t \\ + \displaystyle\int_{t_{k-1}}^{t_k} \Big[\Big(a(x,t,u_{h,k})\nabla_x u_{h,k}^p, \nabla_x \psi_{h,k}\Big)_h \\ \qquad - \Big(a(x,t,u(x,t))\nabla_x u_{h,k}^p, \nabla_x \psi_{h,k}\Big)_h \Big] \mathrm{d}t \\ + \displaystyle\int_{t_{k-1}}^{t_k} [(a_0(x,t,u_{h,k}), \psi_{h,k})_h - (a_0(x,t,u(x,t)), \psi_{h,k})_h] \, \mathrm{d}t = 0, \\ d_{h,0} = 0. \end{cases} \tag{3.44}$$

Subtracting (3.44) from (3.43), we obtain

$$
\begin{cases}
\text{For } k = 1, \dots, N_\tau, \\
\left(d_{h,k} - d_{h,k-1}, \psi_{h,k}\right)_h - (\psi_{h,k} - \psi_{k+1,h}, d_{h,k})_h \\
+ \displaystyle\int_{t_{k-1}}^{t_k} \Big[\left(a(x,t,u_{h,k}) \nabla_x \psi_{h,k}, \nabla_x d_{h,k}\right)_h \\
\quad - \left(a(x,t,u_{h,k}) \nabla_x d_{h,k}, \nabla \psi_{h,k}\right)_h \Big] \, \mathrm{d}t \\
+ \displaystyle\int_{t_{k-1}}^{t_k} \Big[\Big(a(x,t,u_{h,k}) \nabla_x u^p_{h,k}, \nabla_x \psi_{h,k}\Big)_h \\
\quad - \Big(a(x,t,u(x,t)) \nabla_x u^p_{h,k}, \nabla_x \psi_{h,k}\Big)_h \Big] \, \mathrm{d}t \\
- \displaystyle\int_{t_{k-1}}^{t_k} (\theta(x,t,u_\sigma,u^p_\sigma) \nabla_x u \nabla_x \psi_{n,h}, d_{n,h})_h \, \mathrm{d}t \\
+ \displaystyle\int_{t_{k-1}}^{t_k} [(a_0(x,t,u_{h,k}), \psi_{h,k})_h - (a_0(x,t,u(x,t)), \psi_{h,k})_h] \, \mathrm{d}t \\
- \displaystyle\int_{t_{k-1}}^{t_k} (\eta(x,t,u_\sigma,u^p_\sigma) \psi_{h,k}, d_{h,k})_h \, \mathrm{d}t = 0.
\end{cases}
\tag{3.45}
$$

Summing the above inqualities, using the fact that $\psi_{N_\tau+1} = d_{N,h}$, $d_{h,0} = 0$, and working identically to Theorem 3.11 we obtain,

$$
\begin{aligned}
\|d_{N_\tau,h}\|^2_{L^2(\Omega_h)} &= \int_0^T \left(\left(a(x,t,u_\sigma) - a(x,t,u^p_\sigma)\right) \nabla_x (u^p_\sigma - u), \nabla_x \psi_\sigma\right)_h \, \mathrm{d}t \\
&+ \int_0^T ((a(x,t,u^p_\sigma) - a(x,t,u(x,t))) \nabla_x u^p_\sigma, \nabla_x \psi_\sigma)_h \, \mathrm{d}t \\
&+ \int_0^T [(a_0(x,t,u^p_\sigma), \psi_\sigma)_h - (a_0(x,t,u(x,t)), \psi_\sigma)_h] \, \mathrm{d}t \equiv \sum_{i=1}^3 I_i.
\end{aligned}
\tag{3.46}
$$

For the first integral, adding and subtracting $P_\sigma u$ we get

$$
\begin{aligned}
|I_1| &\le \int_0^T \left| \left(\left[a(x,t,u_\sigma) - a(x,t,u^p_\sigma)\right] \nabla_x (u^p_\sigma - u), \nabla_x \psi_\sigma \right)_h \right| \, \mathrm{d}t \\
&\le \int_0^T \left| \left(\left[a(x,t,u_\sigma) - a(x,t,u^p_\sigma)\right] \nabla_x (u^p_\sigma - P_\sigma u), \nabla_x \psi_\sigma \right)_h \right| \, \mathrm{d}t \\
&\quad + \int_0^T \left| \left(\left[a(x,t,u_\sigma) - a(x,t,u^p_\sigma)\right] \nabla_x (P_\sigma u - u), \nabla_x \psi_\sigma \right)_h \right| \, \mathrm{d}t \\
&= I_{1,1} + I_{1.2}.
\end{aligned}
$$

For $I_{1,1}$ we proceeed as follows:

$$
I_{1,1} \le \int_0^T \|u_\sigma - u^p_\sigma\|_{L^{\frac{2p}{p-2}}(\Omega)} \|\nabla_x (u^p_\sigma - P_\sigma u)\|_{L^p(\Omega_h)} \|\nabla_x \psi_\sigma\|_{L^2(\Omega_h)}
$$

$$
\begin{aligned}
&\le C\|d_\sigma\|_{L^2(0,T;L^{\frac{2p}{p-2}}(\Omega_h))}\|u_\sigma^p - P_\sigma u\|_{L^\infty(0,T;W^{1,p}(\Omega_h))}\|\psi_\sigma\|_{L^2(0,T;H^1(\Omega_h))}\\
&\le C\frac{1}{h^{\frac{2}{p}}}\|d_\sigma\|_{L^2(0,T;L^2(\Omega_h))}\frac{1}{h^{2-\frac{2}{p}}}\|u_\sigma^p - P_\sigma u\|_{L^\infty(0,T;L^2(\Omega_h))}\|d_{h,N_\tau}\|_{L^2(\Omega_h)}\\
&\le C\ln\left(\frac{T}{\tau}\right)h^{-2}\tau^{-1/p}(\tau+h^2)^2\|d_{h,N_\tau}\|_{L^2(\Omega_h)}.
\end{aligned}
$$

Here we have used the inverse estimates $\|\cdot\|_{L^{\frac{2p}{p-2}}(\Omega_h)} \le Ch^{-\frac{2}{p}}\|\cdot\|_{L^2(\Omega_h)}$, $\|\cdot\|_{W^{1,p}(\Omega_h)} \le Ch^{-2+\frac{2}{p}}\|\cdot\|_{L^2(\Omega_h)}$, Theorem 3.7 to bound $\|u_\sigma^p - P_\sigma u\|_{L^\infty(0,T;L^2(\Omega_h)}$ and (3.39) to estimate $d_\sigma = u_\sigma - u_\sigma^p$.

For $I_{1,2}$ we first note that due to the stability of P_h in $W^{1,q}(\Omega)$ for any $1 \le q \le \infty$ we obtain

$$
\begin{aligned}
\|P_\sigma u\|_{L^\infty(0,T;W^{1,q}(\Omega_h))} &= \max_{k=1,..,N_\tau}\|P_h u(t_k)\|_{W^{1,q}(\Omega_h)}\\
&\le C\max_{k=1,..,N_\tau}\|u(t_k)\|_{W^{1,q}(\Omega_h)} \le C\|u\|_{L^\infty(0,T;W^{1,q}(\Omega))}.
\end{aligned}
\tag{3.47}
$$

Hence, using the above stability bound, the inverse estimate $\|\cdot\|_{L^{\frac{2p}{p-2}}(\Omega_h)} \le Ch^{-\frac{2}{p}}\|\cdot\|_{L^2(\Omega_h)}$, and (3.39) to estimate d_σ, we conclude,

$$
\begin{aligned}
I_{1,2} &= \int_0^T \left|\Big(\big(a(x,t,u_\sigma) - a(x,t,u_\sigma^p)\big)\nabla_x(P_\sigma u - u), \nabla_x\psi_\sigma\Big)_h\right| \mathrm{d}t\\
&\le C\int_0^T \|d_\sigma\|_{L^{\frac{2p}{p-2}}(\Omega_h)}\|P_\sigma u - u\|_{W^{1,p}(\Omega_h)}\|\psi_\sigma\|_{H^1(\Omega_h)}\,\mathrm{d}t\\
&\le C\|P_\sigma u - u\|_{L^\infty(0,T;W^{1,p}(\Omega_h))}\|d_\sigma\|_{L^2(0,T;L^{\frac{2p}{p-2}}(\Omega_h))}\|\psi_\sigma\|_{L^2(0,T;H^1(\Omega_h))}\\
&\le C\frac{1}{h^{\frac{2}{p}}}\|d_\sigma\|_{L^2(0,T;L^2(\Omega_h))}\|d_{h,N_\tau}\|_{L^2(\Omega_h)}\\
&\le C\frac{1}{h^{\frac{2}{p}}}(\tau+h^2)\|d_{h,N_\tau}\|_{L^2(\Omega_h)}.
\end{aligned}
$$

For integral I_2 of (3.46), we add and subtract $P_\sigma u$ to get

$$
\begin{aligned}
|I_2| &\le \int_0^T \left|((a(x,t,u_\sigma^p) - a(x,t,u(x,t))\nabla_x u_\sigma^p, \nabla_x\psi_\sigma)_h\right| \mathrm{d}t\\
&\le \int_0^T \left|((a(x,t,u_\sigma^p) - a(x,t,u(x,t))\nabla_x(u_\sigma^p - P_\sigma u), \nabla_x\psi_\sigma)_h\right| \mathrm{d}t\\
&\quad + \int_0^T \left|((a(x,t,u_\sigma^p) - a(x,t,u(x,t))\nabla_x P_\sigma u, \nabla_x\psi_\sigma)_h\right| \mathrm{d}t\\
&= I_{2,1} + I_{2,2}.
\end{aligned}
$$

Now, for $I_{2,1}$, using the inverse estimate $\|\cdot\|_{L^\infty(\Omega_h)} \leq Ch^{-1}\|\cdot\|_{L^2(\Omega_h)}$, (3.20) and Theorem 3.6 to estimate $u_\sigma^p - P_\sigma u$ and $e_p = u_\sigma^p - u$ respectively, we deduce

$$\begin{aligned}
I_{2,1} &\leq \int_0^T \|u_\sigma^p - u\|_{L^2(\Omega_h)} \|\nabla_x(u_\sigma^p - P_\sigma u)\|_{L^\infty(\Omega_h)} \|\nabla_x \psi_\sigma\|_{L^2(\Omega_h)} \\
&\leq C\frac{1}{h}\|u_\sigma^p - P_\sigma u\|_{L^\infty(0,T;L^2(\Omega_h))} \|u_\sigma^p - u\|_{L^2(0,T;L^2(\Omega_h))} \|\psi_\sigma\|_{L^2(0,T;H^1(\Omega_h))} \\
&\leq C\ln\left(\frac{T}{\tau}\right)h^{-1}\tau^{-1/p}(\tau+h^2)^2\|d_{h,N_\tau}\|_{L^2(\Omega_h)}.
\end{aligned}$$

Finally for integral $I_{2,2}$, since $u \in L^\infty(0,T;W^{1,\infty}(\Omega))$, using (3.47) with $q = \infty$ we obtain

$$\begin{aligned}
|I_{2,2}| &\leq \int_0^T \|u_\sigma^p - u\|_{L^2(\Omega_h)} \|\nabla_x P_\sigma u\|_{L^\infty(\Omega_h)} \|\nabla_x \psi_\sigma\|_{L^2(\Omega_h)} \\
&\leq C\|P_\sigma u\|_{L^\infty(0,T;W^{1,\infty}(\Omega_h))} \|u_\sigma^p - u\|_{L^2(0,T;L^2(\Omega_h))} \|\psi_\sigma\|_{L^2(0,T;H^1(\Omega_h))} \\
&\leq C(\tau+h^2)\|d_{h,N_\tau}\|_{L^2(\Omega_h)}.
\end{aligned}$$

The last integral I_3 of (3.46) can be treated similarly and more easily, using the stability bound $\|\psi_\sigma\|_{L^\infty(0,T;L^2(\Omega_h))} \leq C\|d_{h,N_\tau}\|_{L^2(\Omega_h)}$.

$$\begin{aligned}
I_3 &= \int_0^T [(a_0(x,t,u_\sigma^p), \psi_\sigma)_h - (a_0(x,t,u(x,t)), \psi_\sigma)_h]\,\mathrm{d}t \\
&\leq \int_0^T \|u_\sigma^p - u\|_{L^2(\Omega_h)} \|\psi_\sigma\|_{L^2(\Omega_h)}\,\mathrm{d}t \leq C\|u_\sigma^p - u\|_{L^2(0,T;L^2(\Omega_h))} \|d_{h,N_\tau}\|_{L^2(\Omega_h))} \\
&\leq C(\tau+h^2)\|d_{h,N_\tau}\|_{L^2(\Omega_h)}.
\end{aligned}$$

Inserting the bounds of I_i, $i = 1,2,3$, into (3.46) and selecting $\tau \approx h^\kappa$ with $\kappa \leq 2$ we obtain

$$\|d_{h,N_\tau}\|_{L^2(\Omega_h)} \leq C\ln\left(\frac{T}{\tau}\right)h^{\kappa(2-\frac{1}{p})-2}, \quad \text{if} \quad \tau \approx h^\kappa, \text{ with } \kappa \leq 2.$$

In the same way we obtain the same estimate for every $d_{h,k}$, $k = 1,\ldots,N_\tau$.
which completes our argument. Hence

$$\begin{aligned}
\|d_\sigma\|_{L^\infty(0,T;L^2(\Omega_h))} &= \max_{k=1,\ldots,N_\tau} \|d_{h,k}\|_{L^2(\Omega_h)} \\
&\leq C\ln\left(\frac{T}{\tau}\right)h^{\kappa(2-\frac{1}{p})-2}, \quad \text{if } \tau \approx h^\kappa, \text{ with } \kappa \leq 2.
\end{aligned} \tag{3.48}$$

Combining this estimate with Theorem 3.7, and using the fact that $u_\sigma - u = d_\sigma + e_p$ we infer the desired estimate.

Finally to prove the convergence of $u - u_\sigma$ in $L^\infty(0, T; L^\infty(\Omega_h))$ we use (3.48) and an inverse estimate to obtain

$$\begin{aligned}\|d_\sigma\|_{L^\infty(0,T;L^\infty(\Omega_h))} &\le \frac{C}{h}\|d_\sigma\|_{L^\infty(0,T;L^2(\Omega_h))}\\ &\le C\ln\left(\frac{T}{\tau}\right)h^{\kappa(2-\frac{1}{p})-3}, \quad \text{if} \quad \tau \approx h^\kappa, \text{ with } \kappa \le 2.\end{aligned}$$

In addition from Theorem 3.7 we deduce

$$\|e_p\|_{L^\infty(0,T;L^2(\Omega_h))} \le C\ln\left(\frac{T}{\tau}\right)h^{\kappa(1-\frac{1}{p})-1}, \quad \text{if} \quad \tau \approx h^\kappa, \text{ with } \kappa \le 2.$$

From the last two estimates we infer that

$$\|u - u_\sigma\|_{L^\infty(0,T;L^\infty(\Omega_h))} \le C\ln\left(\frac{T}{\tau}\right)h^{\kappa(2-\frac{1}{p})-3}, \quad \text{if} \quad \tau \approx h^\kappa, \text{ with } \kappa \le 2.$$

Since κ is assumed to satisfy $k > \frac{3}{2-\frac{1}{p}}$ the above estimate proves the convergence in $L^\infty(0, T; L^\infty(\Omega_h))$.

Remark 3.13 Note that $\tau \approx h^2$ leads to an estimate of order $h^{2-\frac{2}{p}}$ in $L^\infty(0, T; L^2(\Omega_h))$ and $h^{1-\frac{2}{p}}$ in $L^\infty(0, T; L^\infty(\Omega_h))$ up to a logarithmic factor.

Corollary 3.14 *Let $n = 2$. Suppose that $u \in W^{2,1}_p(Q)$, $p > 4$ is the solution of (1.1) and u_σ its discontinuous Galerkin approximation as defined in (3.10). Under Assumptions 1.1, 1.2, 3.1 and 3.2, and supposing that $\tau \approx h^\kappa$ with $\frac{3}{2-\frac{1}{p}} < \kappa \le 2$, then there exists a constant independent of σ such that*

$$\|u - u_\sigma\|_{L^\infty(0,T;L^2(\Omega_h))} \le C\ln\left(\frac{T}{\tau}\right)h^{\kappa(2-\frac{1}{p})-2} \tag{3.49}$$

$$\|u - u_\sigma\|_{L^2(0,T;L^2(\Omega_h))} \le Ch^\kappa, \tag{3.50}$$

$$\|u - u_\sigma\|_{L^2(0,T;H^1(\Omega_h)} \le Ch^{\frac{\kappa}{2}}. \tag{3.51}$$

Here C depends on $\|u\|_{W^{2,1}_p(Q)}$.

Proof Take $M = \|u\|_{L^\infty(0,T;L^\infty(\Omega))} + 1$ and define

$$a_M(x, t, s) = \begin{cases} a(x, t, s) & \text{if } |s| \le M,\\ a(x, t, +M) & \text{if } s > M,\\ a(x, t, -M) & \text{if } s < -M.\end{cases}$$

Analogously we define $a_{0,M}$. Now we consider the problem,

$$\begin{cases} \dfrac{\partial u}{\partial t} - \mathrm{div}_x[a_M(x, t, u(x, t))\nabla_x u] + a_{0,M}(x, t, u(x, t)) = f \;\text{ in } Q,\\ u(x, t) = 0 \;\text{ on } \Sigma, \quad u(x, 0) = u_0(x) \;\text{ in } \Omega,\end{cases}$$

Obviously, Assumptions 1.1, 1.2 and 3.8 are satisfied. Moreover u is the unique solution of the above problem. Consider also $u_\sigma^M \in \mathcal{U}_\sigma$ the solution of the associated discrete problem (3.10) replacing $a(\cdot,\cdot,\cdot)$ by $a_M(\cdot,\cdot,\cdot)$ and $a_0(\cdot,\cdot,\cdot)$ by $a_{0,M}(\cdot,\cdot,\cdot)$. Hence, as a consequence of Theorems 3.12, 3.11, and 3.9 and taking into account that $\tau \approx h^\kappa$, $u_\sigma^M - u$ satisfies estimates (3.49), (3.50) and (3.51).

Moreover, from the previous theorem we get that $\|u_\sigma^M - u\|_{L^\infty(0,T;L^\infty(\Omega_h))} \to 0$, which implies that there exists $\sigma_0 > 0$ such that $\|u_M^\sigma\|_{L^\infty(0,T;L^\infty(\Omega_h))} \le M$ for every $\tau + h \le \sigma_0$. Consequently, the identities $a_M(x,t,u_\sigma^M(x,t)) = a(x,t,u_\sigma^M(x,t))$ and $a_{0,M}(x,t,u_\sigma^M(x,t) = a_0(x,t,u_\sigma^M(x,t))$ hold. This implies that $u_\sigma^M = u_\sigma$ for $\tau + h \le \sigma_0$ and inequalities (3.49), (3.50) and (3.51) are satisfied by $u - u_\sigma$.

Remark 3.15 The results of Theorem 3.12 and Corollary 3.14 hold when $n = 1$. Indeed, modifying the proof of Theorem 3.12, we easily deduce that there exists a constant C depending on $\|u\|_{W_p^{2,1}(Q)}$ such that,

$$\|u_\sigma - u\|_{L^\infty(0,T;L^2(\Omega_h))} \le C \ln\left(\frac{T}{\tau}\right) h^{\kappa(2-\frac{1}{p})-\frac{3}{2}}, \text{ if } \tau \approx h^\kappa, \text{ with } \kappa \le 2,$$

$$\lim_{\sigma\to 0} \|u_\sigma - u\|_{L^\infty(0,T;L^\infty(\Omega_h))} = 0, \qquad \text{when} \quad \tau \approx h^\kappa \quad \text{with} \ \frac{2}{2-\frac{1}{p}} < \kappa \le 2.$$

References

1. Akrivis, G., Crouzeix, M., Makridakis, Ch.: Implicit–explicit multistep methods for quasilinear parabolic equations. Numer. Math. **82**, 521–541 (1999)
2. Akrivis, G., Li, B., Lubich, C.: Combining maximal regularity and energy estimates for time discretizations of quasilinear parabolic equations. Math. Comp. **86**(306), 1527–1552 (2017)
3. Akrivis, G., Makridakis, C.: Galerkin time-stepping methods for nonlinear parabolic equations. ESAIM: M2AN **38**, 261–289 (2004)
4. Amann, H.: Linear and Quasilinear Parabolic Problems, vol. 1. Birkhäuser, Boston (1995)
5. Amann, H.: Compact embeddings of vector-valued Sobolev and Besov spaces. Glas. Mat. **35**(55), 161–177 (2000)
6. Amann, H.: Linear parabolic problems involving measures. Rev. R. Acad. Cienc. Serie A. Mat. **95**(1), 85–119 (2001)
7. Bramble, J.H., King, J.T.: A robust finite element method for nonhomogeneous Dirichlet problems in domains with curved boundaries. Math. Comp. **63**(207), 1–17 (1994)
8. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods. Texts in Applied Mathematics, vol. 15, 3rd edn. Springer, New York (2008)
9. Casas, E., Chrysafinos, K.: Analysis and optimal control of some quasilinear parabolic equations. Math. Control Relat. Fields **8**(3&4), 607–623 (2018)
10. Casas, E., Dhamo, V.: Error estimates for the numerical approximation of a quasilinear Neumann problem under minimal regularity of the data. Numer. Math. **117**, 115–145 (2011)
11. Casas, E., Sokolowski, J.: Approximation of boundary control problems on curved domains. SIAM J. Control Optim. **48**(6), 3746–3780 (2010)
12. Casas, E., Tröltzsch, F.: Numerical analysis of some optimal control problems governed by a class of quasilinear elliptic equations. ESAIM:COCV **17**, 771–800 (2010)
13. Chrysafinos, K., Walkington, N.J.: Error estimates for the discontinuous Galerkin methods for parabolic equations. SIAM J. Numer. Anal. **44**(1), 349–366 (2006)
14. Dendy Jr., J.E.: Galerkin's method for some highly nonlinear problems. SIAM J. Numer. Anal. **14**(2), 327–347 (1977)

15. Dobrowolski, M.: L^∞-convergence of linear finite element approximation to quasilinear initial boundary value problems. RAIRO Anal. Numer. **12**, 247–266 (1978)
16. Dobrowolski, M.: L^∞-convergence of linear finite element approximation to nonlinear initial boundary value problems. SIAM J. Numer. Anal. **17**, 663–674 (1980)
17. Douglas Jr., J., Dupont, T.: Galerkin methods for parabolic equations. SIAM J. Numer. Anal. **7**, 575–626 (1970)
18. Douglas Jr., J., Dupont, T.: A Galerkin method for a nonlinear Dirichlet problem. Math. Comp. **29**(131), 689–696 (1975)
19. Douglas Jr., J., Dupont, T., Wahlbin, L.: The stability in L^q of the L^2 projection into finite element function spaces. Numer. Math. **23**, 193–197 (1975)
20. Leykekham, D., Vexler, B.: Discrete maximal parabolic regularity for Galerkin finite element methods for non-autonomous parabolic problems. SIAM J. Numer. Anal. **56**(4), 2178–2202 (2018)
21. Liu, L., Křížek, M., Neittaanmäki, P.: Higher order finite element approximation of a quasilinear elliptic boundary value problem of a non-monotone type. Appl. Math. **41**(6), 467–478 (1996)
22. Luskin, M.: A Galerkin method for nonlinear parabolic parabolic equations with nonlinear boundary conditions. SIAM J. Numer. Anal. **16**, 284–299 (1979)
23. Prüss, J., Schnaubelt, R.: Solvability and maximal regularity of parabolic evolution equations with coefficients continuous in time. J. Math. Anal. Appl. **256**, 405–430 (2001)
24. Rachford Jr., H.H.: Two-level discrete-time Galerkin approximations for second order nonlinear parabolic partial differential equations. SIAM J. Numer. Anal. **10**, 1010–1026 (1973)
25. Wheeler, M.: A priori L^2 error estimates for Galerkin approximations to parabolic partial differential equations. SIAM J. Numer. Anal. **10**, 723–759 (1973)
26. Zlamal, M.: Finite element methods for nonlinear parabolic equations. RAIRO Anal. Numer. **11**, 93–107 (1977)