THE PARABOLIC p-LAPLACIAN WITH FRACTIONAL DIFFERENTIABILITY

DOMINIC BREIT, LARS DIENING, JOHANNES STORN, AND JÖRN WICHMANN

ABSTRACT. We study the parabolic p-Laplacian system in a bounded domain. We deduce optimal convergence rates for the space-time discretization based on an implicit Euler scheme in time. Our estimates are expressed in terms of Nikolskii spaces and therefore cover situations when the (gradient of) the solution has only fractional derivatives in space and time. The main novelty is that, different to all previous results, we do not assume any coupling condition between the space and time resolution h and τ. The theoretical error analysis is complemented by numerical experiments.

Dedicated to the memory of John W. Barrett

1. INTRODUCTION

Let $\Omega \subset \mathbb{R}^n$ with Lipschitz boundary, $n \geq 2$, $N \geq 1$, $T > 0$ be finite and assume that $\mathbf{f} : Q \to \mathbb{R}^N$ and $\mathbf{u}_0 : \Omega \to \mathbb{R}^N$ are given and let $Q := I \times \Omega$ with $I := (0, T)$. We are interested in the parabolic p-Laplace system

$$
\begin{aligned}
\partial_t \mathbf{u} - \operatorname{div}\left((\kappa + |\nabla \mathbf{u}|)^{p-2} \nabla \mathbf{u} \right) &= \mathbf{f} && \text{in } Q, \\
\mathbf{u} &= 0 && \text{on } I \times \partial\Omega, \\
\mathbf{u}(0, \cdot) &= \mathbf{u}_0 && \text{in } \Omega.
\end{aligned}
\tag{1.1}
$$

with $\kappa \geq 0$ and $p \in (1, \infty)$. The existence of a unique weak solution to (1.1) in the function space

$$
C(\overline{I}; L^2(\Omega)) \cap L^p(I; W_0^{1,p}(\Omega))
$$

can be shown by standard monotonicity arguments under very weak assumptions on the data. We are concerned with its numerical approximation by finite elements. For this purpose we choose discrete subspace V_h of $W_0^{1,p}(\Omega)$, which consists of piecewise polynomials on a quasi-uniform triangulation of mesh size h. Furthermore, we use an implicit Euler scheme with step size $\tau = \frac{T}{M+1}$ for the time discretization. The discrete solution $\mathbf{u}_{m,h}$ is given at time points $t_m = m\tau$, $m = 0, \ldots, M$ of the time grid.

Many authors have studied the error of these discretization, e.g. [Wei92, BL94, EL05, DER07, BDN18]. A variety of quantities has been used to express the error and many error estimates have been deduced under different regularity assumptions on the solution $\mathbf{u}$. It turned out that the natural quantity to measure the error between the discrete and continuous solution is

$$
\max_{0 \leq m \leq M} \|\mathbf{u}(t_m) - \mathbf{u}_{m,h}\|_{L^2(\Omega)}^2 + \sum_{m=1}^{M} \|\mathbf{V}(\nabla \mathbf{u}(t_m)) - \mathbf{V}(\nabla \mathbf{u}_{m,h})\|_{L^2(\Omega)}^2,
\tag{1.2}
$$

where $\mathbf{V}(\boldsymbol{\xi}) = (\kappa + |\boldsymbol{\xi}|)^{\frac{p-2}{2}} \boldsymbol{\xi}$.

The term $\|\mathbf{V}(\nabla\mathbf{u}(t_m)) - \mathbf{V}(\nabla\mathbf{u}_{m,h})\|_{L^2(\Omega)}^2$ is natural to problems involving the p-Laplacian and captures the nonlinear character of the equation. It has been introduced by [BL94] for the numerical analysis of the stationary problem (p-Poisson problem)

$$(1.3) \qquad \begin{aligned} -\operatorname{div}\left((\kappa + |\nabla\mathbf{u}|)^{p-2}\nabla\mathbf{u}\right) &= \mathbf{f} && \text{in } \Omega, \\ \mathbf{u} &= 0 && \text{on } \partial\Omega \end{aligned}$$

in a slightly different but equivalent form under the name *quasi-norm*. Note that if $\kappa = 0$ (1.3) is the Euler-Lagrange equation of the energy

$$\mathcal{J}(\mathbf{v}) := \int\limits_{\Omega} \left(\frac{1}{p}|\nabla\mathbf{v}|^p - \mathbf{v}\cdot\mathbf{f}\right)\mathrm{d}x.$$

It has been observed in [DK08] that the quantity $\|\mathbf{V}(\nabla\mathbf{u}) - \mathbf{V}(\nabla\mathbf{u}_h)\|_{L^2(\Omega)}^2$ is equivalent to the energy error $\mathcal{J}(\mathbf{u}_h) - \mathcal{J}(\mathbf{u})$. This explains that the quantities in (1.2) are the natural way to express the error. The variational approach using $\mathcal{J}$ has been also used in [BDK12] to prove optimal convergence of the adaptive finite element method for the p-Poisson problem using Dörfler marking. It has been shown, starting with the seminal paper by Barrett and Liu [BL93] and with the subsequent extensions by Ebmeyer and Liu [EL05] and by Diening and Růžička [DR07], that solutions to (1.3) satisfy

$$\|\mathbf{V}(\nabla\mathbf{u}(t_m)) - \mathbf{V}(\nabla\mathbf{u}_{m,h})\|_{L^2(\Omega)} \lesssim h\,\|\nabla\mathbf{V}(\nabla\mathbf{u})\|_{L^2(\Omega)}.$$

The required regularity $\nabla\mathbf{V}(\nabla\mathbf{u}) \in L^2(\Omega)$ for the continuous solutions is well-known for problems involving the p-Laplacian. It arises naturally when testing the equation by $\Delta\mathbf{u}$. This test can be made rigorous by the method of difference-quotients under appropriate assumptions on the data (for instance for convex Ω or Ω with $C^{1,\alpha}$-boundary).

In the instationary setting the natural regularity using difference-quotients in time and space is

$$(1.4a) \qquad \mathbf{V}(\nabla\mathbf{u}) \in L^2(I; W^{1,2}(\Omega)) \cap W^{1,2}(I; L^2(\Omega)),$$

$$(1.4b) \qquad \mathbf{u} \in L^\infty(I; W^{1,2}(\Omega)) \cap \mathcal{C}^{0,1}(\overline{I}; L^2(\Omega)).$$

It is well-known that weak solutions to (1.1) enjoy the properties (1.4) provided the data is regular enough and Ω is either convex or has $C^{1,\alpha}$ boundary.

The expected optimal convergence result for linear elements under the regularity assumption (1.4) is

$$(1.5) \qquad \max_{0 \le m \le M} \|\mathbf{u}(t_m) - \mathbf{u}_{m,h}\|_{L^2(\Omega)}^2 + \tau\sum_{m=1}^{M} \|\mathbf{V}(\nabla\mathbf{u}(t_m)) - \mathbf{V}(\nabla\mathbf{u}_{m,h})\|_{L^2(\Omega)}^2 \\ \lesssim h^2 + \tau^2.$$

The analysis of implicit Euler schemes for (1.1) started with the work of Wei [Wei92], who considered the planar case for $p \ge 2$ and obtained sub-optimal estimates for the first part of the error only. In particular, he showed that $\max_m \|\mathbf{u}(t_m) - \mathbf{u}_{m,h}\|_{L^2(\Omega)}^2$ is of order $h^{\frac{1}{(p-1)}} + \tau$ provided that $\mathbf{u} \in C(\overline{I}, W^{2,p}(\Omega))$. Liu and Barett derived in [BL94] significantly better estimates for all $1 < p < \infty$, but still sub-optimal compared to (1.5). Instead of $h^2 + \tau^2$ in (1.5) they obtained $h^{\min\{p,2\}} + \tau$ for $\max_m \|\mathbf{u}(t_m) - \mathbf{u}_{m,h}\|_{L^2(\Omega)}^2$ under strong regularity assumptions of the solution.

The optimal rate (1.5) has been obtained by Diening, Ebmeyer and Růžička in [DER07] for piece-wise linear elements under the assumption $p > \frac{2n}{n+2}$.[1] However,

---

[1] The restriction $p > \frac{2n}{n+2}$ comes from the use of Gelfand triples, which requires $W^{1,p}(\Omega) \hookrightarrow L^2(\Omega)$; but could be avoided.

their analysis has the drawback that there is an unnatural coupling of the time-step size τ and h. In particular, for their optimal convergence result it is required that

$$\text{(1.6)} \qquad\qquad h^{\beta(p,n)} \lesssim \tau,$$

where $\beta(p,n) = 2 - n(1 - \frac{p}{2})$ if $p \in (\frac{n+2}{2n}, 2]$ and $\beta(p,n) = n + \frac{2(2-n)}{p}$ for $p \in [2, \infty)$. The bound for $p \leq 2$ has been recently improved in [BR20] to $\beta(p,n) = \frac{4}{p'}$, where $p' = \frac{p}{p-1}$. Note that different from the well-known CFL condition this is an upper bound for h in terms of τ. Nevertheless, this artificial condition is very much undesired. It is well known that for the linear case $p = 2$ such a condition is not needed. The main contribution of this paper is to remove such artificial restriction completely and to prove that the error estimate (1.5) holds for any choice of h and τ under the regularity assumption (1.4).

The reason for the coupling between h and τ in (1.6) is the use of the Scott-Zhang interpolation Π^1_{SZ} operator [SZ90] in the numerical analysis. This operator has very nice local properties which have been used in [DR07] to control the approximation error $\|\mathbf{V}(\nabla\mathbf{u}) - \mathbf{V}(\nabla\Pi^1_{\text{SZ}}\mathbf{u})\|^2_{L^2(\Omega)}$ in terms of $h^2\|\nabla\mathbf{V}(\nabla\mathbf{u})\|^2_{L^2(\Omega)}$. However, the operator Π^1_{SZ} is not self-adjoint and the treatment of the new term arising from $\partial_t\mathbf{u}$ in the instationary setting becomes harder to estimate. To overcome this problem we rather employ the L^2-projection Π_2 onto the finite element space V_h. This, however, requires a control of the new term $\|\mathbf{V}(\nabla\mathbf{u}) - \mathbf{V}(\nabla\Pi_2\mathbf{u})\|^2_{L^2(\Omega)}$. Since Π_2 is not a local operator, the latter control is rather delicate. We are able to overcome the arising problems by the use of sophisticated decay estimates for the L^2-projection due to Eriksson and Johnson [EJ95] and Boman [Bom06]. Our estimates for $\|\mathbf{V}(\nabla\mathbf{u}) - \mathbf{V}(\nabla\Pi_2\mathbf{u})\|^2_{L^2(\Omega)}$ are summarized in Theorem 7.

Our approach turns out to be flexible enough to even accommodate problems with fractional differentiability. If the data (initial datum, forcing term or boundary of the domain) is not regular enough, weak solutions fail to enjoy the properties (1.4). Consequently, an error estimate of the form (1.5) cannot be expected. In many cases, however, there is at least some fractional differentiability available and one has

$$\text{(1.7a)} \qquad\qquad \mathbf{V}(\nabla\mathbf{u}) \in L^2(I; N^{\alpha_x,2}(\Omega)) \cap N^{\alpha_t,2}(I; L^2(\Omega)),$$

$$\text{(1.7b)} \qquad\qquad \mathbf{u} \in L^\infty(I; N^{\alpha_x,2}(\Omega)) \cap \mathcal{C}^{0,\alpha_t}(\overline{I}; L^2(\Omega)),$$

for some $\alpha_x, \alpha_t \in (0, 1]$. Here $N^{\alpha,2}$ denotes the Nikolskiĭ space with differentiability $\alpha \in (0, 1]$, see Section 2 for details. The corresponding error estimate under these assumptions for $\alpha_t > \frac{1}{2}$ is

$$\text{(1.8)} \qquad \max_{0 \leq m \leq M} \|\mathbf{u}(t_m) - \mathbf{u}_{m,h}\|^2_{L^2(\Omega)} + \tau \sum_{m=1}^{M} \|\mathbf{V}(\nabla\mathbf{u}(t_m)) - \mathbf{V}(\nabla\mathbf{u}_{m,h})\|^2_{L^2(\Omega)}$$
$$\lesssim h^{2\alpha_x} + \tau^{2\alpha_t}.$$

The condition $\alpha_t > \frac{1}{2}$ is necessary for the point evaluation of $\mathbf{V}(\nabla\mathbf{u})$ using the embedding $N^{\alpha_t,2}(I; L^2(\Omega)) \hookrightarrow C(\overline{I}; L^2(\Omega))$. Indeed, such an error estimate has been shown by Breit and Mensah [BM19] in the more general situation of variable exponents $p = p(t, x)$ but again under some condition coupling h and τ. In particular, they require that $h \lesssim \tau^{\frac{1+2\alpha_t}{2\alpha_x}}$. They also require a very weak form of the CFL-condition, namely that $\tau^r \lesssim h$ for some arbitrary, large $r > 0$.

For general $\alpha_t \in (0,1]$ we switch to the following averaged version of the error estimate

$$(1.9) \quad \max_{1 \leq m \leq M} \|\langle \mathbf{u} \rangle_{J_m} - \mathbf{u}_{m,h}\|_{L^2(\Omega)}^2 + \sum_{m=1}^{M} \int_{t_{m-1}}^{t_{m+1}} \|\mathbf{V}(\nabla \mathbf{u}(s)) - \mathbf{V}(\nabla \mathbf{u}_{m,h})\|_{L^2(\Omega)}^2 \, ds$$
$$\lesssim h^{2\alpha_x} + \tau^{2\alpha_t},$$

where $\langle \mathbf{u} \rangle_{J_m}$ is a time average of $\mathbf{u}$ over the intervall $J_m = [t_{m-1}, t_{m+1}]$ for $m \geq 1$. This error estimate is the main result of this paper under the assumption

$$(1.10a) \qquad \mathbf{V}(\nabla \mathbf{u}) \in L^2(I; N^{\alpha_x, 2}(\Omega)) \cap N^{\alpha_t, 2}(I; L^2(\Omega)),$$

$$(1.10b) \qquad \mathbf{u} \in L^\infty(I; N^{\alpha_x, 2}(\Omega)).$$

All exponents $\alpha_x, \alpha_t \in (0,1]$ are allowed and a coupling between h and τ is not needed. The precise statement can be found in Theorem 9. If additionally $\mathbf{u} \in \mathcal{C}^{0,\alpha_t}(\overline{I}; L^2(\Omega))$, then we have also control on the pointwise error (see Remark 11)

$$(1.11) \qquad \max_{1 \leq m \leq M} \|\mathbf{u}(t_m) - \mathbf{u}_{m,h}\|_{L^2(\Omega)}^2 \lesssim h^{2\alpha_x} + \tau^{2\alpha_t}.$$

A main motivation for considering a low time-regularity of the solution comes from stochastic PDEs. In this case the equations are driven by a Wiener process which only belongs to the class $\mathcal{C}^{0,\alpha_t}(I)$ for all $\alpha_t < \frac{1}{2}$. Consequently, only a regularity of the form (1.7) with $\alpha_t < \frac{1}{2}$ is expected. Thus, point evaluations in time like $\mathbf{V}(\nabla \mathbf{u}(t_m))$ as they appear in (1.8) may not be possible in the stochastic case. This problem was circumvented in [BH19] by the use of randomly perturbed time grids. In expectation this corresponds to the time averages that we use in this paper.

The paper is organised as follows. In Section 2 we introduce the analytical setup for equation (1.1) followed by the discrete version in Section 3. Section 4 is devoted to the study of the L^2-projection Π_2 with respect to the approximability of $\mathbf{V}(\nabla \mathbf{u})$. The result can be found in Theorem 7. The main error analysis and the prove of the main result (1.9) without any h and τ coupling can be found in Section 5 in Theorem 9. Section 6 contains the results of a numerical simulation study concerning the discretisation error. In the appendix we recall some well-known results on Orlicz functions which are needed throughout the paper.

2. The continuous equation

In this section we introduce the analytical setup for equation (1.1) including the function spaces. Let $\Omega \subset \mathbb{R}^n$ for $n \geq 2$ be a bounded Lipschitz domain (further assumptions on Ω will be needed for the regularity of solutions and the numerical analysis respectively). For some given $T > 0$ we denote by $I = (0,T)$ the time interval and write $Q := I \times \Omega$ for the space time cylinder. We write $f \lesssim g$ for two non-negative quantities f and g if we f is bounded by g up to a multiplicative constant. The relations $\gtrsim$ and $\eqsim$ are defined accordingly. We denote by c a generic constant which can change its value from line to line.

As usual $L^q(\Omega)$ denotes the Lebesgue spaces and $W^{1,q}(\Omega)$ the Sobolev spaces, where $1 \leq q \leq \infty$. We denote by $W_0^{1,q}(\Omega)$ Sobolev spaces with zero boundary values. It is the closure $C_0^\infty(\Omega)$ (smooth functions with compact support) in $W^{1,q}(\Omega)$. We denote by $W^{-1,q'}(\Omega)$ the dual of $W_0^{1,q}(\Omega)$. In order to express higher regularity of the solutions we need the notation of Nikolskiĭ spaces. For $q \in [1, \infty)$ and

$\alpha \in (0, 1]$ we define the semi-norm and norm

$$[u]_{N^{\alpha,q}(\Omega)} := \sup_{h \in \mathbb{R}^n \setminus \{0\}} |h|^{-\alpha} \left(\int_{\Omega \cap (\Omega - h)} |u(x + h) - u(x)|^q \, dx \right)^{\frac{1}{q}},$$

$$\|u\|_{N^{\alpha,q}(\Omega)} := \|u\|_{L^q(\Omega)} + [u]_{N^{\alpha,q}(\Omega)}.$$

The Nikolskii space $N^{\alpha,q}(\Omega)$ is now defined as the subspace of $L^q(\Omega)$ consisting of functions having finite $\|\cdot\|_{N^{\alpha,q}(\Omega)}$-norm. We call $[\cdot]_{N^{\alpha,q}(\Omega)}$ the semi-norm of $N^{\alpha,q}(\Omega)$. Vector- and matrix-valued functions will usually be denoted in bold case, whereas normal case will be adopted for real-valued functions. We do not distinguish in the notation for the function spaces.

For a separable Banach space $(X, \|\cdot\|_X)$ let $L^q(I; X)$ be the Bochner space of (Bochner-) measureable functions $u : I \to X$ satisfying $t \mapsto \|u(t)\|_X \in L^q(I)$. Moreover, $C(\overline{I}; X)$ is the space of function $u : \overline{I} \to X$ which are continuous with respect to the norm-topology. We also use $\mathcal{C}^{0,\alpha}$ for the space of Hölder continuous functions and its generalization $C^{k,\alpha}$ for higher order derivatives. Similarly to the above, we can define fractional derivatives in time for functions $u : I \to X$, where $(X, \|\cdot\|_X)$ is a separable Banach space. We define for $q \in [1, \infty)$ and $\alpha \in (0, 1]$ the semi-norm and norm

$$[u]_{N^{\alpha,q}(I;X)} := \sup_{\tau \in I} |\tau|^{-\alpha} \left(\int_{I \cap (I - \tau)} \|u(\sigma + \tau) - u(\sigma)\|_X^q \, d\sigma \right)^{\frac{1}{q}},$$

$$\|u\|_{N^{\alpha,q}(I;X)} := \|u\|_{L^p(I;X)} + [u]_{N^{\alpha,q}(I;X)}.$$

The Nikolskii space $N^{\alpha,q}(I; X)$ is now defined as the subspace of the Bochner space $L^q(I; X)$ consisting of the functions having finite $\|\cdot\|_{N^{\alpha,q}(I;X)}$-norm.

For a given force $\mathbf{f} : Q \to \mathbb{R}^N$ and initial value $\mathbf{u}_0 : \Omega \to \mathbb{R}^N$ we are interested in the parabolic p-Laplace system

$$
(2.1) \quad
\begin{aligned}
\partial_t \mathbf{u} - \operatorname{div}\big(\mathbf{S}(\nabla \mathbf{u})\big) &= \mathbf{f} && \text{in } Q, \\
\mathbf{u} &= 0 && \text{on } I \times \partial\Omega, \\
\mathbf{u}(0, \cdot) &= \mathbf{u}_0 && \text{in } \Omega
\end{aligned}
$$

with $\kappa \geq 0$ and $p \in (1, \infty)$, where

$$(2.2) \qquad \mathbf{S}(\nabla \mathbf{u}) := (\kappa + |\nabla \mathbf{u}|)^{p-2} \nabla \mathbf{u}.$$

We will also later need

$$(2.3) \qquad \mathbf{V}(\nabla \mathbf{u}) := (\kappa + |\nabla \mathbf{u}|)^{\frac{p-2}{2}} \nabla \mathbf{u}.$$

It is easy to see that both $\mathbf{S}$ and $\mathbf{V}$ are monotone and invertible.

As usual we use the following notion of weak solutions.

Definition 1. *Assume that $\mathbf{f} \in L^1(Q)$ and $\mathbf{u}_0 \in L^1(\Omega)$. We call*

$$\mathbf{u} \in C(\overline{I}; L^2(\Omega)) \cap L^p(I; W_0^{1,p}(\Omega))$$

a weak solution to (2.1) *if*

$$(2.4) \qquad \int_\Omega \mathbf{u}(t) \cdot \boldsymbol{\xi} \, dx - \int_\Omega \mathbf{u}_0 \cdot \boldsymbol{\xi} \, dx + \int_0^t \!\!\! \int_\Omega \mathbf{S}(\nabla \mathbf{u}) : \nabla \boldsymbol{\xi} \, dx \, d\sigma = \int_0^t \!\!\! \int_\Omega \mathbf{f} \cdot \boldsymbol{\xi} \, dx \, d\sigma$$

for all $\boldsymbol{\xi} \in C_0^\infty(\Omega)$ and all $t \in I$.

The formulation in (2.4) is equivalent to

$$(2.5) \qquad -\int_Q \mathbf{u} \cdot \partial_t \boldsymbol{\xi} \, \mathrm{d}x \, \mathrm{d}t - \int_\Omega \mathbf{u}_0 \cdot \boldsymbol{\xi}(0) \, \mathrm{d}x + \int_Q \mathbf{S}(\nabla \mathbf{u}) : \nabla \boldsymbol{\xi} \, \mathrm{d}x \, \mathrm{d}t = \int_Q \mathbf{f} \cdot \boldsymbol{\xi} \, \mathrm{d}x \, \mathrm{d}t$$

for all $\boldsymbol{\xi} \in C_0^\infty([0,T] \times \Omega)$. It is well known that a weak solution exists provided $\mathbf{f} \in L^{p'}(I; W_0^{-1,p'}(\Omega))$ and $\mathbf{u}_0 \in L^2(\Omega)$.

The following regularity result is a special case of [BM19, Thm. 4.1] (note that the second inclusion in (2.9) is not explicitly stated in [BM19] but follows directly from the proof).

Theorem 2. *Let $\alpha_x, \alpha_t \in (0,1]$ be given and let Ω be a bounded C^{1,α_x}-domain. Let $\mathbf{u}$ be the unique weak solution to* (1.1) *in the sense of Definition 1 with*

$$(2.6) \qquad \mathbf{f} \in L^{p'}(I; N^{\alpha_x, p'}(\Omega)) \cap N^{\alpha_t, 2}(I; L^2(\Omega)),$$

$$(2.7) \qquad \mathbf{u}_0 \in N^{\alpha_x, 2}(\Omega), \quad \operatorname{div} \mathbf{S}(\nabla \mathbf{u}_0) \in L^2(\Omega).$$

Then we have

$$(2.8) \qquad \mathbf{V}(\nabla \mathbf{u}) \in L^2(I; N^{\alpha_x, 2}(\Omega)) \cap N^{\alpha_t, 2}(I; L^2(\Omega)),$$

$$(2.9) \qquad \mathbf{u} \in L^\infty(I; N^{\alpha_x, 2}(\Omega)) \cap \mathcal{C}^{0,\alpha_t}(\overline{I}; L^2(\Omega)).$$

In the case $\alpha_x = \alpha_t = 1$ the result from Theorem 2 is classical and follows formally by testing the equations with $\Delta \mathbf{u}$ and $\partial_t^2 \mathbf{u}$ (see, for instance, [DSS19] for an easy proof). See also [CM20] for sharp, regularity results from testing with the p-Laplacian. Results in a similar spirit concerning the fractional differentiability of nonlinear parabolic systems can be found in [DM05] and [DMS11]. Results concerning the fractional differentiability of related elliptic problems can be found in [AKM18], [DDH$^+$16], [EF01] and [Sav98].

3. The discrete equation

From now on let $\Omega \subset \mathbb{R}^n$ be a polyhedral domain. By $\mathcal{T}_h$ denote a regular partition (triangulation) of Ω (no hanging nodes), which consists of closed n-simplices called *elements*. For each element (n-simplex) $T \in \mathcal{T}_h$ we denote by h_T the diameter of T, and by ρ_T the supremum of the diameters of inscribed balls. By $|T|$ we denote the Lebesgue measure of T. By $\fint_T g \, \mathrm{d}x$ we denote the mean value integral over the set T. We also abbreviate $\langle g \rangle_T = \fint_T g \, \mathrm{d}x$ for the mean value.

We assume that $\mathcal{T}_h$ is *shape regular*, that is there exists a constant γ (the shape regularity constant or chunkiness constant) such that

$$(3.1) \qquad \max_{T \in \mathcal{T}_h} \frac{h_T}{\rho_T} \leq \gamma.$$

We define the maximal mesh-size by

$$h = \max_{T \in \mathcal{T}_h} h_T.$$

We assume further that our triangulation is *quasi-uniform*, i.e.

$$(3.2) \qquad h_T \eqsim h \qquad \text{for all } T \in \mathcal{T}_h.$$

For $T \in \mathcal{T}_h$ we define the set of neighbors ω_T, which consists of all elements $T' \in \mathcal{T}_h$ with $T \cap T' \neq \emptyset$. We define

$$\Omega(\omega_T) := \left(\bigcup \{ T \; : \; T \in \omega_T \} \right)^\circ.$$

We also assume that $\Omega(\omega_T)$ is a connected domain for each T. This only excludes some strange triangulations and is only a small technicality.

To simplify notations we will use ambiguously ω_T instead of $\Omega(\omega_T)$ for the domain for integrals.

It is easy to see that the shape regularity of $\mathcal{T}_h$ implies the following properties, where the constants are independent of h:

(a) $|\omega_T| \eqsim |T|$ for all $T \in \mathcal{T}_h$.
(b) There exists $m_1 \in \mathbb{N}$ such that $\#\omega_T \le m_1$ for all $T \in \mathcal{T}_h$.

For $\ell \in \mathbb{N}_0$ we denote by $\mathscr{P}_\ell(\Omega)$ the polynomials on Ω of degree less than or equal to ℓ. Moreover, we set $\mathscr{P}_{-1}(\Omega) := \{0\}$.

For fixed $r \in \mathbb{N}$ we define the the finite element space V_h as

$$(3.3) \qquad V_h := \{v \in (W_0^{1,1}(\Omega))^N \,:\, v|_T \in (\mathscr{P}_r(T))^N \ \forall T \in \mathcal{T}_h\}.$$

Remark 3. *For the numerical analysis it is only important that V_h contains all continuous, locally linear functions, and that the functions are locally polynomials of a fixed maximal degree. Thus it would for example also be possible to use velocity spaces like the MINI-element (locally linear functions enriched by bubble functions. This might be of interest if our results should be applied to the corresponding fluid system.*

Let $\{0 = t_0 < \cdots < t_M = T\}$ be a uniform partition of $[0, T]$ with mesh size $\tau = T/M$. For $m \ge 1$ define $I_m := [t_{m-1}, t_m]$ and $J_m := [t_{m-1}, t_{m+1}]$.

For a discrete sequence a_m we define the backwards-in-time discrete time derivative d_t by

$$d_t a_m := \tau^{-1}(a_m - a_{m-1}).$$

Then

$$(3.4) \qquad d_t a_m \cdot a_m = \tfrac{1}{2} d_t |a_m|^2 + \tfrac{\tau}{2} |d_t a_m|^2.$$

Let $\mathbf{u}_{0,h} := \Pi_2 \mathbf{u}_0$, where Π_2 is the L^2-projection to V_h. Now for some given time-discrete force $\mathbf{f}_m$, we define $\mathbf{u}_{m,h}$ as the solution of the implicit Euler scheme

$$(3.5) \qquad d_t \mathbf{u}_{m,h} - \operatorname{div}\big(\mathbf{S}(\nabla \mathbf{u}_{m,h})\big) = \mathbf{f}_m \qquad \text{in } V_h^*$$

in the discrete weak sense, i.e. for all $\boldsymbol{\xi}_h \in V_h$ and $m = 1, \ldots, M$ it holds

$$(3.6) \qquad \int_\Omega d_t \mathbf{u}_{m,h} \cdot \boldsymbol{\xi}_h \, dx + \int_\Omega \mathbf{S}(\nabla \mathbf{u}_{m,h}) : \nabla \boldsymbol{\xi}_h \, dx = \int_\Omega \mathbf{f}_m \cdot \boldsymbol{\xi}_h \, dx.$$

Notice that we discretize in space and time simultaneously and avoid an intermediate step with only time or only space discretization. This has the advantage that we do not need to derive regularity properties of additional intermediate problems.

Let us compare our discrete equation to the continuous one. We start with the time steps $m \ge 2$. We first take the average over $(s - \tau, s)$

$$\frac{u(s) - u(s - \tau)}{\tau} - \fint_{(s-\tau, s)} \operatorname{div}\big(\mathbf{S}(\nabla \mathbf{u})(\sigma)\big) \, d\sigma = \fint_{(s-\tau, s)} \mathbf{f}(\sigma) \, d\sigma.$$

Now, take the mean value over J_m with respect to s. Then

$$d_t \langle \mathbf{u} \rangle_{J_m} - \fint_{J_m} \fint_{(s-\tau, s)} \operatorname{div}\big(\mathbf{S}(\nabla \mathbf{u})(\sigma)\big) \, d\sigma \, ds = \fint_{J_m} \fint_{(s-\tau, s)} \mathbf{f}(\sigma) \, d\sigma \, ds.$$

We define $\theta_m : [0, \infty) \to [0, \infty)$ for $m = 2, \ldots, M$ by

(3.7)
$$\theta_m(\sigma) := \fint_\sigma^{\sigma+\tau} \frac{\mathbb{1}_{J_m}(s)}{2\tau} \, ds = \frac{1}{2\tau^2} \int_{\max\{\sigma, t_{m-1}\}}^{\min\{\sigma+\tau, t_{m+1}\}} ds$$
$$= \frac{(\sigma - t_{m-2})\mathbb{1}_{I_{m-1}}(\sigma) + \mathbb{1}_{I_m}(\sigma) + (t_{m+1} - \sigma)\mathbb{1}_{I_{m+1}}(\sigma)}{2\tau^2}.$$

Then θ_m is a weight with total mass one. Let us define weighted averages by

$$\langle \mathbf{g} \rangle_{\theta_m} := \int_{\mathbb{R}} \theta_m(s)\mathbf{g}(s) \, ds.$$

We obtain for $m \geq 2$

(3.8)
$$d_t\langle \mathbf{u} \rangle_{J_m} - \operatorname{div}\left(\langle \mathbf{S}(\nabla\mathbf{u}) \rangle_{\theta_m}\right) = \langle \mathbf{f} \rangle_{\theta_m}.$$

For $m = 1$ we have to proceed slightly differently. We start with our equation (2.1), take the integral over $(0, s)$ and divide by τ to obtain

$$\frac{\mathbf{u}(s) - \mathbf{u}_0}{\tau} - \tau^{-1} \int_{(0,s)} \operatorname{div}\left(\mathbf{S}(\nabla\mathbf{u})(\sigma)\right) d\sigma = \tau^{-1} \int_{(0,s)} \mathbf{f}(\sigma) \, d\sigma.$$

Now, we take the mean value over J_1 with respect to s and obtain

$$\frac{\langle \mathbf{u} \rangle_{J_1} - \mathbf{u}_0}{\tau} - \fint_{J_1} \tau^{-1} \int_0^s \operatorname{div}\left(\mathbf{S}(\nabla\mathbf{u})(\sigma)\right) d\sigma \, ds = \fint_{J_1} \tau^{-1} \int_0^s \mathbf{f}(\sigma) \, d\sigma \, ds,$$

Let us define the weight $\theta_1 : [0, \infty) \to [0, \infty)$ by

(3.9)
$$\theta_1(\sigma) := \frac{1}{2\tau^2} \int_\sigma^\infty \mathbb{1}_{J_1}(s) \, \mathrm{d}s = \frac{2\tau - \sigma}{2\tau^2} \mathbb{1}_{J_1}(\sigma).$$

Then θ_1 has total mass one and we may write

$$\frac{\langle \mathbf{u} \rangle_{J_1} - \mathbf{u}_0}{\tau} - \operatorname{div}\left(\langle \mathbf{S}(\nabla\mathbf{u}) \rangle_{\theta_1}\right) = \langle \mathbf{f} \rangle_{\theta_1}.$$

With $\langle \mathbf{u} \rangle_{J_0} := \mathbf{u}_0$ the initial equation is now given by

(3.10)
$$d_t\langle \mathbf{u} \rangle_{J_1} - \operatorname{div}\left(\langle \mathbf{S}(\nabla\mathbf{u}) \rangle_{\theta_1}\right) = \langle \mathbf{f} \rangle_{\theta_1}.$$

4. PROJECTIONS OPERATORS

In this section we consider projections onto the finite element space V_h introduced in the previous section. In particular, we recall some known properties of the Scott-Zhang interpolation operator and prove a gradient estimate for the error of the L^2-orthogonal projection Π_2 in Theorem 7. The latter one is crucial for the error analysis in the subsequent section.

Let $\Pi_{\mathrm{SZ}}^1 : W_0^{1,1}(\Omega) \to V_h$ denote the standard Scott-Zhang interpolation operator [SZ90] that preserves zero boundary values, where the values of $\Pi_{\mathrm{SZ}}^1\mathbf{v}$ on $\partial\Omega$ are obtained by averaging over edges in $\partial\Omega$. Then this operator is stable in $W^{1,1}$ but unfortunately not in L^1 or L^2. Therefore, we use slight variant of the Scott-Zhang operator: Given a function $\mathbf{v} \in L^1(\Omega)$ we extend it by zero outside of Ω on an additional layer of triangles. Now, we take the Scott-Zhang operator that averages only over n-simplices. At the boundary the n-simplices, where the average is calculated, are chosen to lie outside of Ω. In such a way we obtain zero boundary values and preserve the L^1-stability, see the remark after (4.6) in [SZ90]. Note that Π_{SZ}^0 does not preserve general polynomial boundary data.

These operators have the following nice properties:

(a) (**Projection**) Π_{SZ}^0 and Π_{SZ}^1 are linear projections onto V_h.

(b) (**Local Stability**) There holds uniformly in $T \in \mathcal{T}_h$

(4.1)
$$\fint_T |\Pi_{\text{SZ}}^0 \mathbf{v}| \, dx \lesssim \fint_{\omega_T} |\mathbf{v}| \, dx \qquad \text{for } \mathbf{v} \in L^1(\Omega),$$

$$\fint_T |\Pi_{\text{SZ}}^1 \mathbf{v}| \, dx \lesssim \fint_{\omega_T} |\mathbf{v}| \, dx + \fint_{\omega_T} h_T |\nabla \mathbf{v}| \, dx \qquad \text{for } \mathbf{v} \in W_0^{1,1}(\Omega).$$

It is well-known that these properties imply the following L^p stability results for $1 \le p < \infty$, e.g. [DR07],

(4.2)
$$\left(\fint_T |\Pi_{\text{SZ}}^0 \mathbf{v}|^p \, dx \right)^{\frac{1}{p}} \lesssim \left(\fint_{\omega_T} |\mathbf{v}|^p \, dx \right)^{\frac{1}{p}}$$

$$\left(\fint_T |\Pi_{\text{SZ}}^1 \mathbf{v}|^p \, dx \right)^{\frac{1}{p}} \lesssim \left(\fint_{\omega_T} |\mathbf{v}|^p \, dx \right)^{\frac{1}{p}} + \left(\fint_{\omega_T} h_T^p |\nabla \mathbf{v}|^p \, dx \right)^{\frac{1}{p}}.$$

For $p = \infty$ the mean value integrals have to be exchange by maxima.

The following local estimate has been shown by Diening and Růžička in [DR07, Thm. 5.7].

Proposition 4. *For all $\mathbf{v} \in W^{1,p}(\Omega)$ and all $T \in \mathcal{T}_h$ it holds that*

(4.3)
$$\fint_T \left| \mathbf{V}(\nabla \mathbf{v}) - \mathbf{V}(\nabla \Pi_{\text{SZ}}^1 \mathbf{v}) \right|^2 dx \lesssim \inf_{\mathbf{Q} \in \mathbb{R}^{N \times n}} \fint_{\omega_T} \left| \mathbf{V}(\nabla \mathbf{v}) - \mathbf{V}(\mathbf{Q}) \right|^2 dx$$

$$= \fint_{\omega_T} \left| \mathbf{V}(\nabla \mathbf{v}) - \langle \mathbf{V}(\nabla \mathbf{v}) \rangle_{\omega_T} \right|^2 dx.$$

The implicit constant only depends on p and the shape regularity constant γ.

If follows by a simple application of Poincaré's inequality that

(4.4)
$$\fint_T \left| \mathbf{V}(\nabla \mathbf{v}) - \mathbf{V}(\nabla \Pi_{\text{SZ}}^1 \mathbf{v}) \right|^2 dx \lesssim h_T^2 \fint_{\omega_T} \left| \nabla \mathbf{V}(\nabla \mathbf{v}) \right|^2 dx$$

and by summation over all T

(4.5)
$$\left\| \mathbf{V}(\nabla \mathbf{v}) - \mathbf{V}(\nabla \Pi_{\text{SZ}}^1 \mathbf{v}) \right\|_{L^2(\Omega)} \lesssim h \left\| \nabla \mathbf{V}(\nabla \mathbf{v}) \right\|_{L^2(\Omega)}.$$

Let us make a short remark on local estimates in Nikolskii spaces. For all $g \in N^{\alpha,q}(T)$ with $\alpha \in (0,1]$ and $q \in [1,\infty)$ it follows by Jensen's inequality and the definition of $N^{\alpha,q}(T)$ that

(4.6)
$$\left(\fint_T |g - \langle g \rangle_T|^q \, dx \right)^{\frac{1}{q}} \le \left(\frac{1}{|T|^2} \int_{|z| \le h_T} \int_{T \cap (T-z)} |g(y+z) - g(y)|^q \, dy \, dz \right)^{\frac{1}{q}}$$

(4.7)
$$\lesssim h_T^\alpha \frac{[g]_{N^{\alpha,q}(T)}}{|T|^{\frac{1}{q}}}.$$

It is possible to replace T by ω_T.

The next theorem extends (4.4) and (4.5) to the case of Nikolskii spaces.

Theorem 5. *Let $\alpha \in (0,1]$. For all $T \in \mathcal{T}_h$ it holds that*

(4.8)
$$\left(\fint_T \left| \mathbf{V}(\nabla \mathbf{v}) - \mathbf{V}(\nabla \Pi_{\text{SZ}}^1 \mathbf{v}) \right|^2 dx \right)^{\frac{1}{2}} \lesssim h_T^\alpha \frac{[\mathbf{V}(\nabla \mathbf{v})]_{N^{\alpha,2}(\omega_T)}}{|\omega_T|^{\frac{1}{2}}}.$$

Assume additionally that $\mathcal{T}_h$ is quasi-uniform. Then we have

$$(4.9) \qquad \left\| \mathbf{V}(\nabla\mathbf{v}) - \mathbf{V}(\nabla\Pi_{\mathrm{SZ}}^1\mathbf{v}) \right\|_{L^2(\Omega)} \lesssim h^\alpha [\mathbf{V}(\nabla\mathbf{v})]_{N^{\alpha,2}(\Omega)}.$$

Proof. Estimate (4.8) follows directly from Proposition 4 and (4.7). Similarly with (4.6) we obtain

$$\left\| \mathbf{V}(\nabla\mathbf{v}) - \mathbf{V}(\nabla\Pi_{\mathrm{SZ}}^1\mathbf{v}) \right\|_{L^2(\Omega)}^2$$
$$\lesssim \sum_{T\in\mathcal{T}_h} \int_{\omega_T} \left| \mathbf{V}(\nabla\mathbf{v}) - \langle\mathbf{V}(\nabla\mathbf{v})\rangle_{\omega_T} \right|^2 \, \mathrm{d}x$$
$$\lesssim \frac{1}{h^n} \int_{|z|\leq h} \int_{\Omega\cap(\Omega-z)} \left| \mathbf{V}(\nabla\mathbf{v})(y+z) - \mathbf{V}(\nabla\mathbf{v})(y) \right|^2 \, \mathrm{d}y \, \mathrm{d}z$$
$$\lesssim \left(h^\alpha [\mathbf{V}(\nabla\mathbf{v})]_{N^{\alpha,2}(\Omega)} \right)^2.$$

This proves (4.9). $\qquad\square$

Lemma 6. *Let $\alpha \in (0,1]$. For all $T \in \mathcal{T}_h$ it holds that*

$$\|\mathbf{v} - \Pi_{\mathrm{SZ}}^0\mathbf{v}\|_{L^2(\Omega)} \lesssim h^\alpha [\mathbf{v}]_{N^{\alpha,2}(\Omega)}.$$

Proof. Arguing similarly as in the proof of (4.9) we have for all $\mathbf{v} \in N^{\alpha,p}(\Omega)$

$$\int_\Omega \left| \mathbf{v} - \Pi_{\mathrm{SZ}}^0\mathbf{v} \right|^2 \, \mathrm{d}x \lesssim \sum_{T\in\mathcal{T}_h} \int_T \left| \mathbf{v} - \langle\mathbf{v}\rangle_{\omega_T} \right|^2 \, \mathrm{d}x + \sum_{T\in\mathcal{T}_h} \int_T \left| \Pi_{\mathrm{SZ}}^0\big(\mathbf{v} - \langle\mathbf{v}\rangle_{\omega_T}\big) \right|^2 \, \mathrm{d}x$$
$$\lesssim \sum_{T\in\mathcal{T}_h} \int_{\omega_T} \left| \mathbf{v} - \langle\mathbf{v}\rangle_{\omega_T} \right|^2 \, \mathrm{d}x \lesssim h^{2\alpha} [\mathbf{v}]_{N^{\alpha,2}(\Omega)}^2$$

using the projection property and the local stability estimate (4.2). $\qquad\square$

Although the Scott-Zhang operator has wonderful properties it is not always the best choice for parabolic problems. In particular, the lack of self-adjointness makes serious problems with the discretization of term $\partial_t\mathbf{u}$. For the latter one it is much better to use the L^2-projection $\Pi_2 : L^2(\Omega) \to V_h$.

In fact, we will later use the following identity for the error $\mathbf{e}_m$ (see Section 5 for the exact definition of the error $\mathbf{e}_m$)

$$(4.10) \qquad \begin{aligned} \int_\Omega d_t\mathbf{e}_m \cdot \Pi_2\mathbf{e}_m \, \mathrm{d}x &= \int_\Omega d_t\Pi_2\mathbf{e}_m \cdot \Pi_2\mathbf{e}_m \, \mathrm{d}x \\ &= \frac{1}{2}d_t\|\Pi_2\mathbf{e}_m\|_{L^2(\Omega)}^2 + \frac{\tau}{2}\|d_t\Pi_2\mathbf{e}_m\|_{L^2(\Omega)}^2. \end{aligned}$$

This important identity relies strongly on the self-adjointness of Π_2, which is not available for Π_{SZ}^1. This was the reason for the h and τ coupling in previous papers.

In the following we will extend (4.5) to the L^2-projection Π_2. In particular, we want to prove the following theorem.

Theorem 7. *Let $\mathcal{T}_h$ be quasi-uniform and $\alpha \in (0,1]$. Then*

$$\|\mathbf{V}(\nabla\mathbf{v}) - \mathbf{V}(\nabla\Pi_2\mathbf{v}))\|_{L^2(\Omega)} \lesssim h^\alpha [\mathbf{V}(\nabla\mathbf{v})]_{N^{\alpha,2}(\Omega)}.$$

Before we get to the proof of the theorem let us make a few remarks. The case $p = 2$ reduces to

$$(4.11) \qquad \|\nabla\mathbf{v} - \nabla\Pi_2\mathbf{v}\|_{L^2(\Omega)} \lesssim h^\alpha [\nabla\mathbf{v}]_{N^{\alpha,2}(\Omega)}.$$

Since $\mathcal{T}_h$ is quasi-uniform, this special case can be easily shown with the help of the Scott-Zhang operator. Indeed, using $\Pi_2 \Pi_{\mathrm{SZ}}^1 = \Pi_{\mathrm{SZ}}^1$, inverse estimates and the approximation properties of Π_{SZ}^1, in particular (4.9), we can estimate

$$
\begin{aligned}
\|\nabla \mathbf{v} - \nabla \Pi_2 \mathbf{v}\|_{L^2(\Omega)} &\leq \|\nabla(\mathbf{v} - \Pi_{\mathrm{SZ}}^1 \mathbf{v})\|_{L^2(\Omega)} + \|\nabla \Pi_2(\mathbf{v} - \Pi_{\mathrm{SZ}}^1 \mathbf{v})\|_{L^2(\Omega)} \\
&\lesssim \|\nabla(\mathbf{v} - \Pi_{\mathrm{SZ}}^1 \mathbf{v})\|_{L^2(\Omega)} + h^{-1}\|\Pi_2(\mathbf{v} - \Pi_{\mathrm{SZ}}^1 \mathbf{v})\|_{L^2(\Omega)} \\
&\lesssim \|\nabla(\mathbf{v} - \Pi_{\mathrm{SZ}}^1 \mathbf{v})\|_{L^2(\Omega)} + h^{-1}\|\mathbf{v} - \Pi_{\mathrm{SZ}}^1 \mathbf{v}\|_{L^2(\Omega)} \\
&\lesssim \|\nabla(\mathbf{v} - \Pi_{\mathrm{SZ}}^1 \mathbf{v})\|_{L^2(\Omega)} \\
&\lesssim h^\alpha [\nabla \mathbf{v}]_{N^{\alpha,2}(\Omega)}.
\end{aligned}
$$

However, in the non-linear case $p \neq 2$, where $\nabla \mathbf{v}$ has to be replaced by $\mathbf{V}(\nabla \mathbf{v})$, such a simple trick is not possible. To overcome this problem we will use sophisticated decay estimates of the L^2-projection which are due to Eriksson and Johnson [EJ95] and refined by Boman [Bom06]. In the following we will derive from their results decay estimates of the L^2-projection for our simple situation of quasi-uniform meshes.

Let us define the mollifier

$$
\eta(x) := c_\nu \exp(-\nu|x|),
$$

with c_ν such that $\|\eta\|_{L^1(\mathbb{R}^n)} = 1$. Then $\eta_h(x) := h^{-n}\eta(x/h)$ satisfies $\|\eta_h\|_{L^1(\mathbb{R}^n)} = 1$ as well.

Lemma 8 (Decay estimates of the L^2-projection). *Let $\mathcal{T}_h$ be quasi-uniform. Then for every $x \in T$ and all $\mathbf{v} \in L^1(\Omega)$ and all $\mathbf{w} \in W^{1,1}(\Omega)$ we have*

$$
(4.12) \qquad |(\Pi_2 \mathbf{v})(x)| \lesssim \big(\eta_h * (\mathbb{1}_\Omega |\mathbf{v}|)\big)(x),
$$

$$
(4.13) \qquad |(\nabla \Pi_2 \mathbf{w})(x)| \lesssim \big(\eta_h * (\mathbb{1}_\Omega |\nabla \mathbf{w}|)\big)(x).
$$

The implicit constants only depend on n and the shape regularity γ.

Proof. We begin with the proof of (4.12)

$$
\begin{aligned}
|(\Pi_2 \mathbf{v})(x)| &\lesssim \fint_T |\Pi_2 \mathbf{v}|\, \mathrm{d}y \\
&\lesssim \int_\Omega \eta_h(x-y)|(\Pi_2 \mathbf{v})(y)|\, \mathrm{d}y = \|\eta_h(x-\cdot)\Pi_2\mathbf{v}(\cdot)\|_{L^1(\Omega)}.
\end{aligned}
$$

Since our triangulation is quasi-uniform we may choose a constant regularized mesh function $h(x) := h$ in order to apply the results of [EJ95] and [Bom06]. In particular, by Lemma 2.3 of [Bom06] (applied to the case $p = 1$) it follows that

$$
|(\Pi_2 \mathbf{v})(x)| \lesssim \|\eta_h(x-\cdot)\mathbf{v}(\cdot)\|_{L^1(\Omega)} \lesssim \big(\eta_h * (\mathbb{1}_\Omega |\mathbf{v}|)\big)(x).
$$

This proves (4.12). Let us remark that the results of Boman are unfortunately not properly displayed. In particular, they define

$$
\delta_{\mathcal{T}_h} := \max_{T \in \mathcal{T}_h} \max_{T' \in \omega_T} |1 - h_T^2/h_{T'}^2|.
$$

Thus, only a uniform mesh gives $\delta_{\mathcal{T}_h} = 0$. It would have been better to use

$$
\delta_{\mathcal{T}_h} := \max_{T \in \mathcal{T}_h} \max_{x \in T} \max_{y \in \omega_T} |1 - h(x)^2/h(y)^2|,
$$

which is zero for all quasi-uniform meshes with constant regularized mesh function. In the paper of Eriksson and Johnson [EJ95] this was done properly and the case of quasi-uniform meshes is included. A careful inspection of the proofs by Boman shows that it is enough to use the alternative definition of $\delta_{\mathcal{T}_h}$, so that quasi-uniform meshes are included.

The proof of (4.13) is analogously using Lemma 2.5 of Boman [Bom06], i.e.

$$\|\eta_h(x - \cdot)\nabla\Pi_2\mathbf{w}(\cdot)\|_{L^1(\Omega)} \lesssim \|\eta_h(x - \cdot)\nabla\mathbf{w}(\cdot)\|_{L^1(\Omega)}.$$

This proves the claim. □

Closely related to (2.2) and (2.3) is the shifted Orlicz function φ_a defined by

$$(4.14) \qquad \varphi_a(t) := \int_0^t (\kappa + a + s)^{p-2} s \, ds$$

for $a \geq 0$ (see also the appendix). We are now prepared for the proof of Theorem 7.

Proof of Theorem 7. We estimate using Lemma 13

$$\mathrm{I} := \int_\Omega |\mathbf{V}(\nabla\mathbf{v}) - \mathbf{V}(\nabla\Pi_2\mathbf{v})|^2 \, dx$$

$$\eqsim \int_\Omega \varphi_{|\nabla\mathbf{v}|}(|\nabla\mathbf{v} - \nabla\Pi_2\mathbf{v}|) \, dx$$

$$\lesssim \int_\Omega \varphi_{|\nabla\mathbf{v}|}(|\nabla\mathbf{v} - \nabla\Pi^1_{\mathrm{SZ}}\mathbf{v}|) \, dx + \int_\Omega \varphi_{|\nabla\mathbf{v}|}(|\nabla\Pi_2(\mathbf{v} - \Pi^1_{\mathrm{SZ}}\mathbf{v})|) \, dx$$

$$=: \mathrm{II} + \mathrm{III}.$$

Now, by Lemma 13 and Theorem 5

$$\mathrm{II} \eqsim \int_\Omega |\mathbf{V}(\nabla\mathbf{v}) - \mathbf{V}(\nabla\Pi^1_{\mathrm{SZ}}\mathbf{v})|^2 \, dy \lesssim h^{2\alpha} [\mathbf{V}(\nabla\mathbf{v})]^2_{N^{\alpha,2}(\Omega)}.$$

Moreover, by Lemma 8, Jensen's inequality

$$\mathrm{III} \lesssim \int_\Omega \varphi_{|\nabla\mathbf{v}(x)|}\big(\eta_h * (\mathbb{1}_\Omega|\nabla(\mathbf{v} - \Pi^1_{\mathrm{SZ}}\mathbf{v})|)\big)(x) \, dx$$

$$= \int_\Omega \varphi_{|\nabla\mathbf{v}(x)|}\left(\int_\Omega \eta_h(x - y)|\nabla(\mathbf{v} - \Pi^1_{\mathrm{SZ}}\mathbf{v})(y)| \, dy\right) dx$$

$$\lesssim \int_\Omega \int_\Omega \eta_h(x - y)\varphi_{|\nabla\mathbf{v}(x)|}(|\nabla(\mathbf{v} - \Pi^1_{\mathrm{SZ}}\mathbf{v})(y)|) \, dy \, dx.$$

Now, by the shift-change Lemma 15 and Lemma 13

$$\lesssim \int_\Omega \int_\Omega \eta_h(x - y)\varphi_{|\nabla\mathbf{v}(y)|}(|\nabla(\mathbf{v} - \Pi^1_{\mathrm{SZ}}\mathbf{v})(y)|) \, dy \, dx$$

$$+ \int_\Omega \int_\Omega \eta_h(x - y)|\mathbf{V}(\nabla\mathbf{v})(x) - \mathbf{V}(\nabla\mathbf{v})(y)|^2 \, dy \, dx$$

$$\lesssim \int_\Omega |\mathbf{V}(\nabla\mathbf{v}) - \mathbf{V}(\nabla\Pi^1_{\mathrm{SZ}}\mathbf{v})|^2 \, dy$$

$$+ \int_\Omega \int_\Omega \eta_h(x - y)|\mathbf{V}(\nabla\mathbf{v})(x) - \mathbf{V}(\nabla\mathbf{v})(y)|^2 \, dy \, dx$$

$$=: \mathrm{III}_1 + \mathrm{III}_2.$$

Again by Theorem 5

$$\mathrm{III}_1 \lesssim h^{2\alpha} [\mathbf{V}(\nabla\mathbf{v})]^2_{N^{\alpha,2}(\Omega)}.$$

We estimate further

$$
\begin{aligned}
\text{III}_2 &= \int_\Omega \int_\Omega \eta_h(x-y)|\mathbf{V}(\nabla\mathbf{v})(x) - \mathbf{V}(\nabla\mathbf{v})(y)|^2 \, \mathrm{d}y\,\mathrm{d}x \\
&= \int_{\mathbb{R}^n} \int_{\Omega\cap(\Omega-z)} \eta_h(z)|\mathbf{V}(\nabla\mathbf{v}(y+z)) - \mathbf{V}(\nabla\mathbf{v}(y))|^2 \, \mathrm{d}y\,\mathrm{d}z \\
&\lesssim \int_{\mathbb{R}^n} \eta_h(z)|z|^{2\alpha}\,\mathrm{d}z\,[\mathbf{V}(\nabla\mathbf{v})]^2_{N^{\alpha,2}(\Omega)} \\
&= h^{2\alpha} \int_{\mathbb{R}^n} \eta(z)|z|^{2\alpha}\,\mathrm{d}z\,[\mathbf{V}(\nabla\mathbf{v})]^2_{N^{\alpha,2}(\Omega)} \\
&\lesssim h^{2\alpha}\,[\mathbf{V}(\nabla\mathbf{v})]^2_{N^{\alpha,2}(\Omega)} \,,
\end{aligned}
$$

which proves the claim. $\qquad\square$

5. Error analysis

Our aim now is to establish the convergence rate of the difference between the solution to the continuous problem solving (2.4), and that of the discrete problem (3.6). To do this, we first collect the following assumptions on the data. Throughout the rest of this section we assume that

$$(5.1) \qquad \mathbf{u}_0 \in L^2(\Omega), \quad \mathbf{f} \in L^{p'}(0,T;W^{-1,p'}(\Omega)).$$

Recall the definition of weighted averages

$$
\langle \mathbf{g} \rangle_{\theta_m} := \int_{\mathbb{R}} \theta_m(\sigma)\mathbf{g}(\sigma)\,\mathrm{d}\sigma
$$

with θ_m given by (3.7) resp. (3.9).

In the following we state the main result of the paper.

Theorem 9. *Suppose that* (5.1) *holds. Let $\mathbf{u}$ be the unique weak solution to* (1.1) *in the sense of Definition 1. Assume that*

$$(5.2) \qquad \begin{aligned} \mathbf{V}(\nabla\mathbf{u}) &\in L^2(I; N^{\alpha_x,2}(\Omega)) \cap N^{\alpha_t,2}(I; L^2(\Omega)), \\ \mathbf{u} &\in L^\infty(I; N^{\alpha_x,2}(\Omega)) \end{aligned}$$

for some $\alpha_x, \alpha_t \in (0,1]$. Let $\mathcal{T}_h$ be quasi-uniform. Then we have uniformly in τ and h

$$
\max_{1\le m\le M}\|\langle\mathbf{u}\rangle_{J_m} - \mathbf{u}_{m,h}\|^2_{L^2(\Omega)} + \sum_{m=1}^M \int_{J_m} \|\mathbf{V}(\nabla\mathbf{u}(s)) - \mathbf{V}(\nabla\mathbf{u}_{m,h})\|^2_{L^2(\Omega)}\,\mathrm{d}s
$$

$$
\lesssim h^{2\alpha_x}\left(\sup_{s\in[0,T]}[\mathbf{u}(s)]^2_{N^{\alpha_x,2}(\Omega)} + \int_0^T [\mathbf{V}(\nabla\mathbf{u}(s))]^2_{N^{\alpha_x,2}(\Omega)}\,\mathrm{d}s \right)
$$

$$
+ \tau^{2\alpha_t}\,[\mathbf{V}(\nabla\mathbf{u})]^2_{N^{\alpha_t,2}(I;L^2(\Omega))} \,,
$$

where $\mathbf{u}_{m,h}$ is the solution to (3.5) *with $\mathbf{f}_m = \langle\mathbf{f}\rangle_{\theta_m}$ and $\mathbf{u}_{0,h} := \Pi_2\mathbf{u}_0$, where Π_2 is the L^2-projection to V_h. The hidden constant is independent of T.*

Proof. Define the averaged error by $\mathbf{e}_m := \langle\mathbf{u}\rangle_{J_m} - \mathbf{u}_{m,h}$.

Recall the solution $\mathbf{u}$ satisfies (3.8) and (3.10), whereas the discrete solution $\mathbf{u}_{m,h}$ satisfies (3.5). Therefore it holds for $m = 1, \ldots, M$ as an equation in V_h^*

$$(5.3) \qquad d_t\mathbf{e}_m - \mathrm{div}\left(\langle\mathbf{S}(\nabla\mathbf{u})\rangle_{\theta_m} - \mathbf{S}(\nabla\mathbf{u}_{m,h})\right) = \langle\mathbf{f}\rangle_{\theta_m} - \mathbf{f}_m = 0.$$

Let us also recall that the L^2-projection $\Pi_2 : L^2(\Omega) \to V_h$ is defined by

$$(5.4) \qquad\qquad (\Pi_2\mathbf{v} - \mathbf{v}, \boldsymbol{\xi}_h) = 0$$

for all $\boldsymbol{\xi}_h \in V_h$. In particular for $\boldsymbol{\xi}_h = \Pi_2\mathbf{v}$ it holds

$$(5.5) \qquad\qquad (\Pi_2\mathbf{v}, \Pi_2\mathbf{v}) = (\mathbf{v}, \Pi_2\mathbf{v}).$$

Let $m \in \{1, \ldots, M\}$ and choose $\boldsymbol{\xi}_h = \Pi_2\mathbf{e}_m$ in (5.3) and get

$$
\begin{aligned}
\mathrm{I} + \mathrm{II} &:= \int_\Omega d_t\mathbf{e}_m \cdot \Pi_2\mathbf{e}_m \,\mathrm{d}x + \int_\mathbb{R} \int_\Omega \theta_m(\sigma)(\mathbf{S}(\nabla\mathbf{u}(\sigma)) - \mathbf{S}(\nabla\mathbf{u}_{m,h})) : \nabla\mathbf{e}_m \,\mathrm{d}x\,\mathrm{d}\sigma \\
&= \int_\mathbb{R} \int_\Omega \theta_m(\sigma)(\mathbf{S}(\nabla\mathbf{u}(\sigma)) - \mathbf{S}(\nabla\mathbf{u}_{m,h})) : \nabla(\mathbf{e}_m - \Pi_2\mathbf{e}_m) \,\mathrm{d}x\,\mathrm{d}\sigma \\
&=: \mathrm{III}.
\end{aligned}
$$

Keep in mind that the weights θ_1 and θ_m are supported in $I_1 \cup I_2$ and $I_{m-1} \cup I_m \cup I_{m+1}$ respectively, see (3.7) and (3.9). Consequently, the integrals over $\mathbb{R}$ are well-defined.

Due to (5.5) and (3.4) the first term can be written as

$$\mathrm{I} = \int_\Omega d_t\Pi_2\mathbf{e}_m \cdot \Pi_2\mathbf{e}_m \,\mathrm{d}x = \frac{1}{2}d_t\|\Pi_2\mathbf{e}_m\|_{L^2(\Omega)}^2 + \frac{\tau}{2}\|d_t\Pi_2\mathbf{e}_m\|_{L^2(\Omega)}^2.$$

We split the second term into two parts and use Lemma 13

$$
\begin{aligned}
\mathrm{II} &= \int_\mathbb{R} \fint_{J_m} \int_\Omega \theta_m(\sigma)(\mathbf{S}(\nabla\mathbf{u}(\sigma)) - \mathbf{S}(\nabla\mathbf{u}_{m,h})) : \nabla(\mathbf{u}(s) - \mathbf{u}_{m,h}) \,\mathrm{d}x\,\mathrm{d}s\,\mathrm{d}\sigma \\
&\approx \fint_{J_m} \int_\Omega |\mathbf{V}(\nabla\mathbf{u}(s)) - \mathbf{V}(\nabla\mathbf{u}_{m,h})|^2 \,\mathrm{d}x\,\mathrm{d}s \\
&\quad + \int_\mathbb{R} \fint_{J_m} \int_\Omega \theta_m(\sigma)(\mathbf{S}(\nabla\mathbf{u}(\sigma)) - \mathbf{S}(\nabla\mathbf{u}(s))) : \nabla(\mathbf{u}(s) - \mathbf{u}_{m,h}) \,\mathrm{d}x\,\mathrm{d}s\,\mathrm{d}\sigma \\
&=: \mathrm{II}_1 + \mathrm{II}_2.
\end{aligned}
$$

The latter is estimated using Lemma 14

$$
\begin{aligned}
\mathrm{II}_2 &\le \delta \fint_{J_m} \int_\Omega |\mathbf{V}(\nabla\mathbf{u}(s)) - \mathbf{V}(\nabla\mathbf{u}_{m,h})|^2 \,\mathrm{d}x\,\mathrm{d}\sigma \\
&\quad + c_\delta \int_\mathbb{R} \fint_{J_m} \int_\Omega \theta_m(\sigma)|\mathbf{V}(\nabla\mathbf{u}(s)) - \mathbf{V}(\nabla\mathbf{u}(\sigma))|^2 \,\mathrm{d}x\,\mathrm{d}s\,\mathrm{d}\sigma.
\end{aligned}
$$

The next step is to decompose III in an analogous way. Due to Lemma 13 and 14 it holds

$$\text{III} = \fint_{J_m} \int_\Omega (\mathbf{S}(\nabla\mathbf{u}(s)) - \mathbf{S}(\nabla\mathbf{u}_{m,h})) : \nabla(\mathbf{u}(s) - \Pi_2\mathbf{u}(s))\, dx\, ds$$

$$+ \int_\mathbb{R} \fint_{J_m} \int_\Omega \theta_m(\sigma)(\mathbf{S}(\nabla\mathbf{u}(\sigma)) - \mathbf{S}(\nabla\mathbf{u}(s))) : \nabla(\mathbf{u}(s) - \Pi_2\mathbf{u}(s))\, dx\, ds\, d\sigma$$

$$\leq \delta \fint_{J_m} \int_\Omega |\mathbf{V}(\nabla\mathbf{u}(s)) - \mathbf{V}(\nabla\mathbf{u}_{m,h})|^2\, dx\, ds$$

$$+ c_\delta \fint_{J_m} \int_\Omega |\mathbf{V}(\nabla\mathbf{u}(s)) - \mathbf{V}(\nabla\Pi_2\mathbf{u}(s))|^2\, dx\, ds$$

$$+ c_\delta \int_\mathbb{R} \fint_{J_m} \int_\Omega \theta_m(\sigma)|\mathbf{V}(\nabla\mathbf{u}(\sigma)) - \mathbf{V}(\nabla\mathbf{u}(s))|^2\, dx\, ds\, d\sigma$$

$$=: \text{III}_1 + \text{III}_2 + \text{III}_3.$$

Recall the definition of the weights (3.7) and (3.9) and estimate

(5.6)
$$\text{III}_2 \leq \int_\mathbb{R} \fint_{J_m} \int_\Omega \theta_m(\sigma)|\mathbf{V}(\nabla\mathbf{u}(s)) - \mathbf{V}(\nabla\mathbf{u}(\sigma))|^2\, dx\, ds\, d\sigma$$

$$\lesssim \frac{1}{\tau} \int_{|z|\leq\tau} \int_{\operatorname{supp}\theta_m} \|\mathbf{V}(\nabla\mathbf{u}(\sigma + z)) - \mathbf{V}(\nabla\mathbf{u}(\sigma))\|^2_{L^2(\Omega)}\, d\sigma\, dz.$$

Now, sum over m and multiply by τ,

$$\|\Pi_2\mathbf{e}_m\|^2_{L^2(\Omega)} + \tau \sum_{l=1}^m \tau\|d_t\Pi_2\mathbf{e}_l\|^2_{L^2(\Omega)}$$

$$+ \tau \sum_{l=1}^m \fint_{J_l} \int_\Omega |\mathbf{V}(\nabla\mathbf{u}(s)) - \mathbf{V}(\nabla\mathbf{u}_{l,h})|^2\, dx\, ds$$

$$\lesssim \tau \sum_{l=1}^m \fint_{J_l} \int_\Omega |\mathbf{V}(\nabla\mathbf{u}(s)) - \mathbf{V}(\nabla\Pi_2\mathbf{u}(s))|^2\, dx\, ds$$

$$+ \tau \sum_{l=1}^m \int_\mathbb{R} \fint_{J_l} \int_\Omega \theta_l(\sigma)|\mathbf{V}(\nabla\mathbf{u}(\sigma)) - \mathbf{V}(\nabla\mathbf{u}(s))|^2\, dx\, ds\, d\sigma$$

$$=: K_1 + K_2.$$

The non-linear stability Theorem 7 for the L^2-projection yields

$$K_1 \lesssim h^{2\alpha_x} \int_0^T [\mathbf{V}(\nabla\mathbf{u}(s))]^2_{N^{\alpha_x,2}(\Omega)}\, ds.$$

The second term is bounded using (5.6) by

$$K_2 \lesssim \sum_{l=1}^m \int_{|z|\leq\tau} \int_{\operatorname{supp}\theta_l} \|\mathbf{V}(\nabla\mathbf{u}(\sigma + z)) - \mathbf{V}(\nabla\mathbf{u}(\sigma))\|^2_{L^2(\Omega)}\, d\sigma\, dz$$

$$\lesssim \tau^{2\alpha_t} [\mathbf{V}(\nabla\mathbf{u})]^2_{N^{\alpha_t,2}(I;L^2(\Omega))}.$$

We have obtained uniform estimates for the projected averaged error $\Pi_2 \mathbf{e}_m$

$$\|\Pi_2 \mathbf{e}_m\|^2_{L^2(\Omega)} + \tau \sum_{l=1}^{m} \fint_{J_l} \int_{\Omega} |\mathbf{V}(\nabla \mathbf{u}(s)) - \mathbf{V}(\nabla \mathbf{u}_{l,h})|^2 \, \mathrm{d}x \, \mathrm{d}s$$

$$\lesssim h^{2\alpha_x} \int_0^T [\mathbf{V}(\nabla \mathbf{u}(s))]^2_{N^{\alpha_x,2}(\Omega)} \, \mathrm{d}s + \tau^{2\alpha_t} [\mathbf{V}(\nabla \mathbf{u})]^2_{N^{\alpha_t,2}(I;L^2(\Omega))}.$$

However, we are interested in the error $\mathbf{e}_m := \langle \mathbf{u} \rangle_{J_m} - \mathbf{u}_{m,h}$. Thus, we estimate

$$\|\mathbf{e}_m\|^2_{L^2(\Omega)} \lesssim \|\langle \mathbf{u} \rangle_{J_m} - \Pi_2 \langle \mathbf{u} \rangle_{J_m}\|^2_{L^2(\Omega)} + \|\Pi_2 \mathbf{e}_m\|^2_{L^2(\Omega)}$$
$$=: K_3 + K_4.$$

The first term can be bounded using Jensen's inequality and Lemma 6

$$K_3 \lesssim \fint_{J_m} \|\mathbf{u}(s) - \Pi_2 \mathbf{u}(s)\|^2_{L^2(\Omega)} \, \mathrm{d}s \leq \fint_{J_m} \|\mathbf{u}(s) - \Pi^0_{\mathrm{SZ}} \mathbf{u}(s)\|^2_{L^2(\Omega)} \, \mathrm{d}s$$

$$\lesssim h^{2\alpha_x} \sup_{s \in J_m} [\mathbf{u}(s)]^2_{N^{\alpha_x,2}(\Omega)}.$$

The bound for K_4 has already been established.

Collecting all terms and taking the maximum over m, we arrive at the desired estimate

$$\max_{1 \leq m \leq M} \|\langle \mathbf{u} \rangle_{J_m} - \mathbf{u}_{m,h}\|^2_{L^2(\Omega)} + \sum_{m=1}^{M} \int_{J_m} \|\mathbf{V}(\nabla \mathbf{u}(s)) - \mathbf{V}(\nabla \mathbf{u}_{m,h})\|^2_{L^2(\Omega)} \, \mathrm{d}s$$

$$\lesssim h^{2\alpha_x} \left(\sup_{s \in [0,T]} [\mathbf{u}(s)]^2_{N^{\alpha_x,2}(\Omega)} + \int_0^T [\mathbf{V}(\nabla \mathbf{u}(s))]^2_{N^{\alpha_x,2}(\Omega)} \, \mathrm{d}s \right)$$

$$+ \tau^{2\alpha_t} [\mathbf{V}(\nabla \mathbf{u})]^2_{N^{\alpha_t,2}(I;L^2(\Omega))}.$$

This proves the claim. $\qquad \square$

It is not that important that we choose $\mathbf{f}_m = \langle \mathbf{f} \rangle_{\theta_m}$. Indeed, we can allow for a certain class of discrete forces $\mathbf{f}_m$ in our numerical scheme and still have convergence of order α_t.

Corollary 10. *Let the assumption of Theorem 9 be satisfied. Additionally, assume that there exists $c_{\mathbf{f}} \geq 0$ independent of τ such that*

$$(5.7) \qquad \tau \sum_{m=1}^{M} \|\langle \mathbf{f} \rangle_{\theta_m} - \mathbf{f}_m\|^2_{L^2(\Omega)} \leq c_{\mathbf{f}} \tau^{2\alpha_t}.$$

Then we have uniformly in τ and h

$$\max_{1 \leq m \leq M} \|\langle \mathbf{u} \rangle_{J_m} - \mathbf{u}_{m,h}\|^2_{L^2(\Omega)} + \sum_{m=1}^{M} \int_{J_m} \|\mathbf{V}(\nabla \mathbf{u}(s)) - \mathbf{V}(\nabla \mathbf{u}_{m,h})\|^2_{L^2(\Omega)} \, \mathrm{d}s$$

$$\lesssim h^{2\alpha_x} \left(\sup_{s \in [0,T]} [\mathbf{u}(s)]^2_{N^{\alpha_x,2}(\Omega)} + \int_0^T [\mathbf{V}(\nabla \mathbf{u}(s))]^2_{N^{\alpha_x,2}(\Omega)} \, \mathrm{d}s \right)$$

$$+ \tau^{2\alpha_t} \left([\mathbf{V}(\nabla \mathbf{u})]^2_{N^{\alpha_t,2}(I;L^2(\Omega))} + c_{\mathbf{f}} \right).$$

In particular, we can use $\mathbf{f}_m := \mathbf{f}(t_m)$ for $\mathbf{f} \in \mathcal{C}^{0,\alpha_t}(0,T;L^2(\Omega))$. The hidden constant may depend exponentially on T.

Proof. The proof is essentially the same as for Theorem 9. But we do not have the the cancellation of $\langle \mathbf{f} \rangle_{\theta_m}$ and $\mathbf{f}_m$ as in (5.3). Instead it holds for all $m \in \{1, \ldots, M\}$ and $\boldsymbol{\xi}_h \in V_h$

(5.8)
$$\int\limits_{\Omega} d_t \mathbf{e}_m \cdot \boldsymbol{\xi}_h \, \mathrm{d}x + \int\limits_{\mathbb{R}} \int\limits_{\Omega} \theta_m(\sigma)(\mathbf{S}(\nabla \mathbf{u}(\sigma)) - \mathbf{S}(\nabla \mathbf{u}_{m,h})) \cdot \nabla \boldsymbol{\xi}_h \, \mathrm{d}x \, \mathrm{d}\sigma$$
$$= \int\limits_{\mathbb{R}} \int\limits_{\Omega} \theta_m(\sigma) \left(\mathbf{f}(\sigma) - \mathbf{f}_m \right) \cdot \boldsymbol{\xi}_h \, \mathrm{d}x \, \mathrm{d}\sigma.$$

We choose $\xi_h = \Pi_2 \mathbf{e}_m$ and proceed as in the proof of Theorem 9. The additional term involving the difference in $\mathbf{f}$ is bounded using Hölder's and Young's inequality by

$$\int\limits_{\Omega} \int\limits_{\mathbb{R}} \theta_m(\sigma) \left(\mathbf{f}(\sigma) - \mathbf{f}_m \right) \mathrm{d}\sigma \cdot \Pi_2 \mathbf{e}_m \, \mathrm{d}x \leq \|\langle \mathbf{f} \rangle_{\theta_m} - \mathbf{f}_m\|_{L^2(\Omega)} \|\Pi_2 \mathbf{e}_m\|_{L^2(\Omega)}$$

$$\leq \|\langle \mathbf{f} \rangle_{\theta_m} - \mathbf{f}_m\|^2_{L^2(\Omega)} + \|\Pi_2 \mathbf{e}_m\|^2_{L^2(\Omega)}.$$

The summation over m and the multiplication by τ yield

$$\|\Pi_2 \mathbf{e}_m\|^2_{L^2(\Omega)} + \tau \sum_{l=1}^{m} \tau \|d_t \Pi_2 \mathbf{e}_l\|^2_{L^2(\Omega)}$$
$$+ \tau \sum_{l=1}^{m} \fint_{J_l} \int\limits_{\Omega} |\mathbf{V}(\nabla \mathbf{u}(s)) - \mathbf{V}(\nabla \mathbf{u}_{l,h})|^2 \, \mathrm{d}x \, \mathrm{d}s$$
$$\lesssim \tau \sum_{l=1}^{m} \fint_{J_l} \int\limits_{\Omega} |\mathbf{V}(\nabla \mathbf{u}(s)) - \mathbf{V}(\nabla \Pi_2 \mathbf{u}(s))|^2 \, \mathrm{d}x \, \mathrm{d}s$$
$$+ \tau \sum_{l=1}^{m} \int\limits_{\mathbb{R}} \fint_{J_l} \int\limits_{\Omega} \theta_l(\sigma) |\mathbf{V}(\nabla \mathbf{u}(\sigma)) - \mathbf{V}(\nabla \mathbf{u}(s))|^2 \, \mathrm{d}x \, \mathrm{d}s \, \mathrm{d}\sigma$$
$$+ \tau \sum_{l=1}^{m} \|\langle \mathbf{f} \rangle_{\theta_l} - \mathbf{f}_l\|^2_2 + \tau \sum_{l=1}^{m} \|\Pi_2 \mathbf{e}_l\|^2_{L^2(\Omega)}.$$

The same arguments as in Theorem 9, an application of the discrete Gronwall inequality (with constants that depend exponentially on the time horizon T) and (5.7) allow to close the argument. $\square$

Remark 11. *If additionally* $\mathbf{u} \in \mathcal{C}^{0,\alpha_t}(\overline{I}; L^2(\Omega))$, *then we can replace*

$$\max_{1 \leq m \leq M} \|\langle \mathbf{u} \rangle_{J_m} - \mathbf{u}_{m,h}\|^2_{L^2(\Omega)}$$

in the error estimate of Theorem 9 and Corollary 10 by the point-wise error

$$\max_{1 \leq m \leq M} \|\mathbf{u}(t_m) - \mathbf{u}_{m,h}\|^2_{L^2(\Omega)}.$$

This follows immediately from

$$\|\mathbf{u}(t_m) - \mathbf{u}_{m,h}\|_{L^2(\Omega)} \lesssim \|\langle \mathbf{u} \rangle_{J_m} - \mathbf{u}_{m,h}\|_{L^2(\Omega)} + \tau^{\alpha_t} \, [\mathbf{u}]_{\mathcal{C}^{0,\alpha_t}(\overline{I}; L^2(\Omega))} \, .$$

6. NUMERICAL EXPERIMENTS

This section underlines the theoretical results of this paper and exploits the practical behaviour of the proposed numerical scheme. The experiments base on the open source tool for solving partial differential equations FEniCS [LMW12]. The supplementary material of this paper contains the implementations of the following three experiments.

6.1. **Constant force on slit domain.** The first experiment approximates the solution to the p-heat equation (1.1) on the slit domain $\Omega = (-1, 1)^2 \setminus (-1, 0] \times \{0\}$ with constant right-hand side $\mathbf{f} \equiv 2$. The experiment applies the numerical scheme of this paper with the discrete spaces V_h from (3.3) with $r = 1$ (piece-wise affine polynomials) and $r = 2$ (piece-wise quadratic polynomials) for a sequence of uniformly refined meshes $\mathcal{T}_h$ and halved time steps τ. For $r = 1$, the smallest space V_h is of dimension $\dim V_h = \mathrm{ndof} = 12$ and the computation utilizes $M = 4$ time steps, the largest space V_h is of dimension $\dim V_h = \mathrm{ndof} = 82689$ and the computation utilizes $M = 512$ time steps ($33 \le \dim V_h \le 82689$ and $4 \le M \le 256$ for $r = 2$). Since the exact solution to this problem is unknown, we compare the solutions $\mathbf{u}_{m,h} \in V_h$ to a reference solution $\mathbf{u}_{\mathrm{ref}} \in L^2(I; V_h^{\mathrm{ref}})$. The space V_h^{ref} is of polynomial degree $r + 1$ and dimension $\dim V_h^{\mathrm{ref}} = 1337107$ for $r = 1$ and $\dim V_h = 875467$ for $r = 2$. The underlying triangulation of V_h^{ref} results from an adaptive finite element loop (see [CFPP14, Sec. 5.1]) for the Poisson model problem $-\Delta \mathbf{u} \equiv 1$ in $H^{-1}(\Omega)$. The reference solution $\mathbf{u}_{\mathrm{ref}}$ results from this paper's scheme with $M = 1024$ ($r = 1$) and $M = 512$ ($r = 2$) time steps. We expect from regularity theory that

$$
\begin{aligned}
(6.1) \qquad \mathbf{V}(\nabla \mathbf{u}) &\in L^2(I; N^{\frac{1}{2}, 2}(\Omega)) \cap N^{1,2}(I; L^2(\Omega)), \\
\mathbf{u} &\in L^\infty(I; N^{1,2}(\Omega)) \cap \mathcal{C}^{0,1}(\overline{I}; L^2(\Omega)).
\end{aligned}
$$

Thus, the solution is smooth in time ($\alpha_t = 1$) but rough in space ($\alpha_x = 1/2$). The convergence history plots in Figure 1 displays the convergence of the error contributions

$$
\mathrm{err}_{L^\infty L^2} := \max_{1 \le m \le M} \|\mathbf{u}_{h,m} - \langle \mathbf{u}_{\mathrm{ref}} \rangle_{J_m}\|_{L^2(\Omega)}^2,
$$

$$
\begin{aligned}
(6.2) \qquad \mathrm{err}_{L^2 \mathbf{V}} &:= \tau \sum_{m=1}^{M} \|\mathbf{V}(\nabla \mathbf{u}_{h,m}) - \mathbf{V}(\nabla \langle \mathbf{u}_{\mathrm{ref}} \rangle_{J_m})\|_{L^2(\Omega)}^2 \\
&\approx \tfrac{1}{2} \sum_{m=1}^{M} \int_{J_m} \|\mathbf{V}(\nabla \mathbf{u}_{h,m}) - \mathbf{V}(\nabla \mathbf{u}_{\mathrm{ref}})\|_{L^2(\Omega)}^2 \, \mathrm{d}\sigma.
\end{aligned}
$$

The resulting rate is in agreement with the a priori estimate in Theorem 9, that is, the error $\mathrm{err}_{L^2 \mathbf{V}} \eqsim h \eqsim \tau$. The fast convergence of the error $\mathrm{err}_{L^\infty L^2} \eqsim h^2 \eqsim \tau^2$ might result from the higher regularity of $\mathbf{u}$ in (6.1).

6.2. **Rough in time.** This experiment solves the p-heat equation (1.1) on the unit square domain $\Omega = (0, 1)^2$ in the time interval $I = (-0.1, 0.1)$ with right hand side $\mathbf{f}(t) \equiv \mathrm{sgn}(t)|t|^{-\beta}$ for fixed parameter $\beta \in \{0.1, 0.5, 0.9\}$ and all $t \in I$. The discrete space V_h is of polynomial degree $r = 1$. The experiment compares the discrete solutions to a reference solution $\mathbf{u}_{\mathrm{ref}}$ that results from a computation with $M = 1024$ time steps and a discrete space V_h^{ref} of higher polynomial degree $r + 1 = 2$ with dimension $\dim V_h^{\mathrm{ref}} = 263169$. Besides the squared errors from (6.2), the convergence history plots in Figure 2 display the squared error

$$
\mathrm{err}_{L^{p'} \mathbf{S}} := \left(\tau \sum_{m=1}^{M} \|\mathbf{S}(\nabla \mathbf{u}_{h,m}) - \mathbf{S}(\nabla \langle \mathbf{u}_{\mathrm{ref}} \rangle_{J_m})\|_{L^{p'}(\Omega)}^{p'} \right)^{2/p'}.
$$

We expect that $\mathbf{u}$ is smooth in space ($\alpha_x = 1$) but rough in time. In particular, $\mathbf{u} \in C^{1-\beta}(\overline{I}; L^2(\Omega))$. Thus, Remark 11 would apply with $\alpha_t = 1 - \beta$. However, if we only look at averaged errors, Theorem 9 shows that the restriction for α_t comes from $\mathbf{V}(\nabla \mathbf{u}) \in N^{\alpha_t, 2}(I; L^2(\Omega))$. Inspired by the regularity of [CM20], we expect that $\partial_t \mathbf{u}$ behaves similar to $\mathbf{f}$. Hence, heuristic calculations suggest for $p = 1.5$ that $\mathbf{V}(\nabla \mathbf{u}) \in N^{\alpha_t, 2}(I; L^2(\Omega))$ with $\alpha_t = 0.575$ for $\beta = 0.9$ and $\alpha_t = 0.875$ for
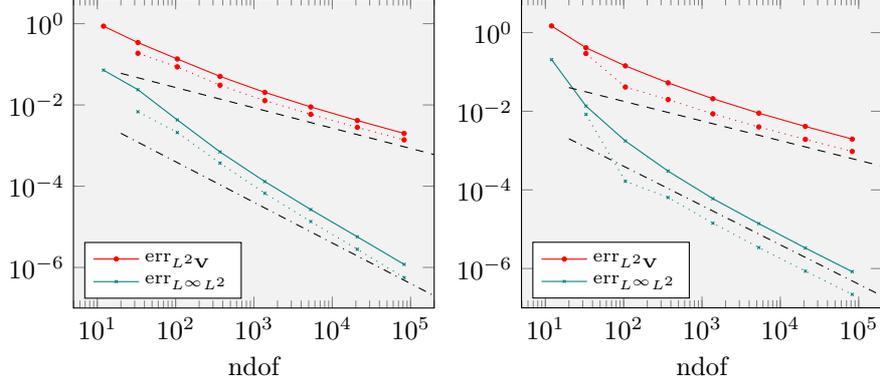
FIGURE 1. Convergence history plots for the experiment in Section 6.1 with $p = 1.5$ (left) and $p = 3$ (right) as well as polynomial degrees $r = 1$ (solid) and $r = 2$ (dotted). The dashed line (- - -) indicates rate $\text{err}_{\text{total}} \eqsim \text{ndof}^{-1/2} \eqsim h$ and the dash-dotted line (-·-·-) the rate $\text{ndof}^{-1} \eqsim h^2$.
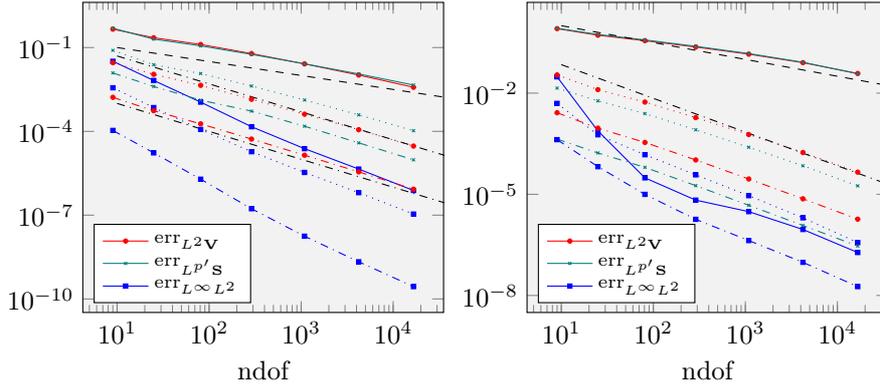
FIGURE 2. Convergence history plots of the experiment from Section 6.2 with $\beta = 0.9$ (solid), $\beta = 0.5$ (dotted), and $\beta = 0.1$ (dash-dotted) for $p = 1.5$ (left) and $p = 3$ (right). The dashed line (- - -) indicates the rate $\text{ndof}^{-1/2} \eqsim h \eqsim \tau$ and the dash-dotted lines (-·-·-) indicate the rate $\text{ndof}^{-1} \eqsim h^2 \eqsim \tau^2$.

$\beta = 0.5$ and $\alpha_t = 1$ for $\beta = 0.1$. For $p = 3$, we expect $\alpha_t = 1$ for $\beta = 0.5, 0.1$ and $\alpha_t = 0.65$ for $\beta = 0.9$. Indeed, the convergence history plot in Figure 2 indicates rates of convergence far better than $1 - \beta$. For $p = 1.5$, the convergence of the error $\text{err}_{L^2\mathbf{V}}$ agrees with our heuristic calculations and Theorem 9. For $p = 3$, the rate for $\beta = 0.9$ and $\beta = 0.5$ seems to be slightly worse than our heuristic predicts. Either the heuristic is inexact, or we observe a pre-asymptotic effect. Such a pre-asymptotic effect might also lead to the slightly worse convergence rate in the experiment of Figure 3 with $p = 3$. The rate of the errors $\text{err}_{L^{p'}\mathbf{S}} \eqsim \text{err}_{L^2\mathbf{V}}$ is similar in all computations. The error $\text{err}_{L^\infty L^2}$ converges with a better rate than the error $\text{err}_{L^2\mathbf{V}}$.

6.3. **Known solution.** This experiment designs the right-hand side $\mathbf{f}$ and the inhomogeneous Dirichlet boundary conditions such that the solution reads

$$
\begin{aligned}
\mathbf{u}(x,t) &= p'|t|^{1/2}|x|^{1/p'} && (\text{with } 1/p + 1/p' = 1), \\
\partial_t \mathbf{u}(x,t) &= p'/2 \operatorname{sgn}(t)|t|^{-1/2}|x|^{1/p'}, \\
\mathbf{V}(\nabla \mathbf{u}(x,t)) &= |t|^{p/4}|x|^{-1/2}\frac{x}{|x|}, \\
\mathbf{S}(\nabla \mathbf{u}(x,t)) &= |t|^{(p-1)/2}|x|^{-1/p'}\frac{x}{|x|}.
\end{aligned}
\tag{6.3}
$$

The time interval $I = (-1,1)$ and the domain is either the centred square domain $\Omega_1 = (-1,1)^2$ or the shifted square domain $\Omega_2 = (1,3) \times (-1,1)$. The function $\mathbf{u}$ has singularities in space (at $x = 0$) and time (at $t = 0$), which cause the reduced regularity $\alpha_t = 1/2$ and $\alpha_x = 1/2$ (for the domain Ω_1) and $\alpha_x = 1$ (for Ω_2). We include the inhomogeneous Dirichlet boundary conditions by averaging the nodal interpolation: Let $\Pi_h : W^{1,p} \to V_h$ be the nodal interpolation operator onto the space of piece-wise affine polynomials V_h from (3.3), then we set at the boundary

$$
\mathbf{u}_{m,h}|_{\partial\Omega} := \langle \Pi_h \mathbf{u} \rangle_{J_m}|_{\partial\Omega} \qquad \text{for all } m = 1, \ldots, M.
$$

The convergence history plots in Figure 3 display the squared errors

$$
\operatorname{err}_{L^\infty L^2} := \max_{1 \le m \le M} \|\mathbf{u}_{h,m} - \langle \mathbf{u}_{\mathrm{ref}} \rangle_{J_m}\|_{L^2(\Omega)}^2,
$$

$$
\operatorname{err}_{L^2 \langle \mathbf{V} \rangle_{J_m}} := \tau \sum_{m=1}^{M} \|\mathbf{V}(\nabla \mathbf{u}_{h,m}) - \langle \mathbf{V}(\nabla \mathbf{u}) \rangle_{J_m}\|_{L^2(\Omega)}^2,
$$

$$
\operatorname{err}_{L^2 \mathbf{V}} := \sum_{m=1}^{M} \int_{J_m} \|\mathbf{V}(\nabla \mathbf{u}_{h,m}) - \mathbf{V}(\nabla \mathbf{u})\|_{L^2(\Omega)}^2 \, d\sigma
$$

$$
= \operatorname{err}_{L^2 \langle \mathbf{V} \rangle_{J_m}} + \sum_{m=1}^{M} \int_{J_m} \|\mathbf{V}(\nabla \mathbf{u}) - \langle \mathbf{V}(\nabla \mathbf{u}) \rangle_{J_m}\|_{L^2(\Omega)}^2 \, d\sigma,
$$

$$
\operatorname{err}_{L^{p'} \langle \mathbf{S} \rangle_{J_m}} := \left(\tau \sum_{m=1}^{M} \|\mathbf{S}(\nabla \mathbf{u}_{h,m}) - \langle \mathbf{S}(\nabla \mathbf{u}) \rangle_{J_m}\|_{L^{p'}(\Omega)}^{p'} \right)^{2/p'}.
$$

As in the previous experiments, the error $\operatorname{err}_{L^2 \mathbf{V}}$ dominates the error $\operatorname{err}_{L^\infty L^2}$. The solution to the problem on Ω_1 (solid lines) converges with the expected rate $h \eqsim \tau$ for $p = 1.5$. Since $\mathbf{S}(\nabla u) \in L^{p'}(I; N^{\frac{1}{3},p'}(\Omega))$, the rate of $\operatorname{err}_{L^{p'} \langle \mathbf{S} \rangle_{J_m}}$ is worse but agrees with the possible approximation rate. For $p = 3$ and Ω_1 the rate of the error $\operatorname{err}_{L^2 \mathbf{V}}$ is slightly worse than τ. This might be some pre-asymptotic effect or it is caused by quadrature errors due to the highly singular right-hand side $\mathbf{f} := \partial_t \mathbf{u} - \operatorname{div} \mathbf{S}(\nabla \mathbf{u})$. The averaged errors in the computation on Ω_2 overcome, as stated in Theorem 9, the reduced regularity of $\mathbf{u}$ in $t = 0$: the squared error $\operatorname{err}_{L^\infty L^2} + \operatorname{err}_{L^2 \mathbf{V}}$ converges with the optimal rate $h^2 \eqsim \tau^2$. As in the previous experiments, the error $\operatorname{err}_{L^\infty L^2}$ is much smaller than the error $\operatorname{err}_{L^2 \mathbf{V}}$. In addition, the error $\operatorname{err}_{L^\infty L^2}$ converges faster than $\operatorname{err}_{L^\infty \mathbf{V}}$ for $p = 1.5$ and Ω_1. For $p = 3$ and Ω_1, the rates of $\operatorname{err}_{L^\infty L^2}$ and $\operatorname{err}_{L^\infty \mathbf{V}}$ are similar.

Appendix A. Orlicz spaces

The following definitions and results are standard in the theory of Orlicz spaces. A continuous, convex and strictly increasing function $\varphi : [0,\infty) \to [0,\infty)$ satisfying

$$
\lim_{t \to 0} \frac{\varphi(t)}{t} = \lim_{t \to \infty} \frac{t}{\varphi(t)} = 0
$$
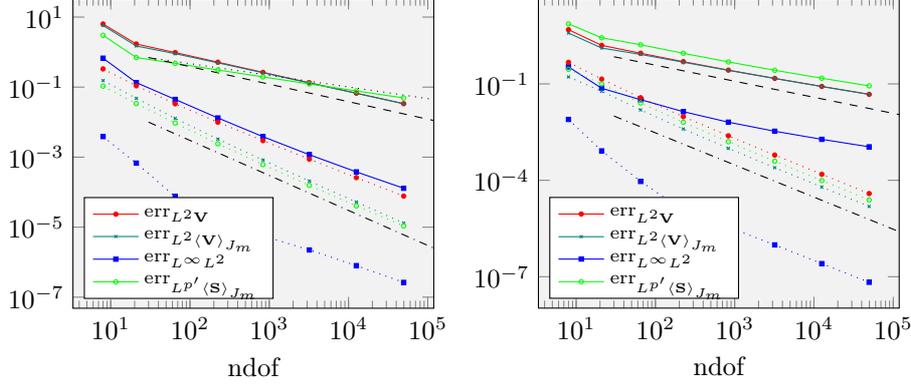
FIGURE 3. Convergence history plots for the experiment in Section 6.3 with $p = 1.5$ (left), $p = 3$ (right) and domains Ω_1 (solid), Ω_2 (dotted). The dotted line (······) indicates the rate $\mathrm{ndof}^{-1/3} \eqsim h^{2/3}$, the dashed line (– – –) indicates the rate $\mathrm{ndof}^{-1/2} \eqsim h$ and the dash-dotted line (–·–·–) indicates the rate $\mathrm{ndof}^{-1} \eqsim h^2$.

is called an N-function.

We say that φ satisfies the Δ_2–condition, if there exists $c > 0$ such that for all $t \geq 0$ holds $\varphi(2t) \leq c\,\varphi(t)$. By $\Delta_2(\varphi)$ we denote the smallest such constant. Since $\varphi(t) \leq \varphi(2t)$ the Δ_2-condition is equivalent to $\varphi(2t) \eqsim \varphi(t)$ uniformly in t. Note that if $\Delta_2(\varphi) < \infty$ then $\varphi(t) \eqsim \varphi(c\,t)$ uniformly in $t \geq 0$ for any fixed $c > 0$. For a family φ_λ of N-functions we define $\Delta_2(\{\varphi_\lambda\}) := \sup_\lambda \Delta_2(\varphi_\lambda)$. By L^φ and $W^{k,\varphi}$, $k \in \mathbb{N}_0$, we denote the classical Orlicz and Orlicz-Sobolev spaces, i.e. $f \in L^\varphi$ iff $\int \varphi(|f|)\,\mathrm{d}x < \infty$ and $f \in W^{k,\varphi}$ iff $\nabla^j f \in L^\varphi$, $0 \leq j \leq k$.

By φ^* we denote the conjugate N-function of φ, which is given by $\varphi^*(t) = \sup_{s \geq 0}(st - \varphi(s))$. Then $\varphi^{**} = \varphi$.

The following definitions and results are summarized from [DR07, BDK12, DFTW20].

Definition 12. *Let φ be an N-function. We say that φ is* uniformly convex, *if φ is C^1 on $[0,\infty)$ and C^2 on $(0,\infty)$ and assume that*

$$(A.1) \qquad\qquad \varphi'(t) \eqsim t\,\varphi''(t)$$

uniformly in $t > 0$. The constants hidden in $\eqsim$ are called the characteristics *of φ.*

Note that (A.1) is stronger than $\Delta_2(\varphi, \varphi^*) < \infty$. In fact, the Δ_2-constants can be estimated in terms of the characteristics of φ.

Associated to an uniformly convex N-function φ we define the tensors

$$\mathbf{S}(\boldsymbol{\xi}) := \frac{\varphi'(|\boldsymbol{\xi}|)}{|\boldsymbol{\xi}|}\boldsymbol{\xi}, \quad \boldsymbol{\xi} \in \mathbb{R}^{N \times n}$$

$$\mathbf{V}(\boldsymbol{\xi}) := \sqrt{\frac{\varphi'(|\boldsymbol{\xi}|)}{|\boldsymbol{\xi}|}}\,\boldsymbol{\xi}, \quad \boldsymbol{\xi} \in \mathbb{R}^{N \times n}.$$

We define the *shifted N-function* φ_a for $a \geq 0$ by

$$(A.2) \qquad\qquad \varphi_a(t) := \int_0^t \frac{\varphi'(a+s)}{a+s}s\,\mathrm{d}s.$$

In our application (cf. (2.2), (2.3) and (4.14)) φ is given by

$$\varphi(t) = \int\limits_0^t (\kappa + s)^{p-2} s \, \mathrm{d}s$$

and the tensors are

$$\mathbf{S}(\boldsymbol{\xi}) = (\kappa + |\boldsymbol{\xi}|)^{p-2} \boldsymbol{\xi} \quad \text{and} \quad \mathbf{V}(\boldsymbol{\xi}) = (\kappa + |\boldsymbol{\xi}|)^{\frac{p-2}{2}} \boldsymbol{\xi}.$$

Lemma 13 (Equivalence lemma)**.** *We have*

$$\big(\mathbf{S}(\mathbf{P}) - \mathbf{S}(\mathbf{Q})\big) \cdot \big(\mathbf{P} - \mathbf{Q}\big) \approx \big|\mathbf{V}(\mathbf{P}) - \mathbf{V}(\mathbf{Q})\big|^2$$
$$\approx \varphi_{|\mathbf{P}|}(|\mathbf{P} - \mathbf{Q}|)$$
$$\approx \varphi''(|\mathbf{P}| + |\mathbf{Q}|)|\mathbf{P} - \mathbf{Q}|^2$$

uniformly in $\mathbf{P}, \mathbf{Q} \in \mathbb{R}^{N \times n}$. *Moreover, uniformly in* $\mathbf{Q} \in \mathbb{R}^{N \times n}$,

$$\mathbf{S}(\mathbf{Q}) \cdot \mathbf{Q} \approx |\mathbf{V}(\mathbf{Q})|^2 \approx \varphi(|\mathbf{Q}|)$$
$$|\mathbf{S}(\mathbf{P}) - \mathbf{S}(\mathbf{Q})| \approx \big(\varphi_{|\mathbf{P}|}\big)'(|\mathbf{P} - \mathbf{Q}|).$$

The constants depend only on the characteristics of φ.

Lemma 14. *Let* φ *be an uniformly convex N-function. Then for each* $\delta > 0$ *there exists* $C_\delta \geq 1$ *(only depending on* δ *and the characteristics of* φ*) such that*

$$\big(\mathbf{S}(\mathbf{P}) - \mathbf{S}(\mathbf{Q})\big) \cdot \big(\mathbf{R} - \mathbf{Q}\big) \leq \delta\big|\mathbf{V}(\mathbf{P}) - \mathbf{V}(\mathbf{Q})\big|^2 + C_\delta\big|\mathbf{V}(\mathbf{R}) - \mathbf{V}(\mathbf{Q})\big|^2$$

for all $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in \mathbb{R}^{N \times n}$.

Lemma 15 (Change of Shift)**.** *Let* φ *be an uniformly convex N-function. Then for each* $\delta > 0$ *there exists* $C_\delta \geq 1$ *(only depending on* δ *and the characteristics of* φ*) such that*

$$\varphi_{|\mathbf{a}|}(t) \leq C_\delta \, \varphi_{|\mathbf{b}|}(t) + \delta \, |\mathbf{V}(\mathbf{a}) - \mathbf{V}(\mathbf{b})|^2,$$

for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^{N \times n}$ *and* $t \geq 0$.

REFERENCES

[AKM18]   Benny Avelin, Tuomo Kuusi, and Giuseppe Mingione, *Nonlinear Calderón-Zygmund theory in the limiting case*, Arch. Ration. Mech. Anal. **227** (2018), no. 2, 663–714. MR 3740385

[BDK12]   Liudmila Belenki, Lars Diening, and Christian Kreuzer, *Optimality of an adaptive finite element method for the p-Laplacian equation*, IMA J. Numer. Anal. **32** (2012), no. 2, 484–510. MR 2911397

[BDN18]   Sören Bartels, Lars Diening, and Ricardo H. Nochetto, *Unconditional stability of semi-implicit discretizations of singular flows*, SIAM J. Numer. Anal. **56** (2018), no. 3, 1896–1914.

[BH19]   Dominic Breit and Martina Hofmanova, *Space-time approximation of stochastic p-Laplace systems*, arXiv e-prints (2019), arXiv:1904.03134.

[BL93]   John W. Barrett and W. B. Liu, *Finite element approximation of the p-Laplacian*, Math. Comp. **61** (1993), no. 204, 523–537.

[BL94]   ――――, *Finite element approximation of the parabolic p-Laplacian*, SIAM J. Numer. Anal. **31** (1994), no. 2, 413–428.

[BM19]   Dominic Breit and Prince Romeo Mensah, *Space-time approximation of parabolic systems with variable growth*, IMA Journal of Numerical Analysis (2019), drz039.

[Bom06]   Mats Boman, *Estimates for the L₂-projection onto continuous finite element spaces in a weighted Lₚ-norm*, BIT **46** (2006), no. 2, 249–260.

[BR20]   Luigi C. Berselli and Michael Růžička, *Space-time discretization for nonlinear parabolic systems with p-structure*, arXiv e-prints (2020), arXiv:2001.09888.

[CFPP14]   C. Carstensen, M. Feischl, M. Page, and D. Praetorius, *Axioms of adaptivity*, Comput. Math. Appl. **67** (2014), no. 6, 1195–1253. MR 3170325

[CM20]   Andrea Cianchi and Vladimir G. Maz'ya, *Second-order regularity for parabolic p-laplace problems*, The Journal of Geometric Analysis **30** (2020).

[DDH⁺16] Stephan Dahlke, Lars Diening, Christoph Hartmann, Benjamin Scharf, and Markus Weimar, *Besov regularity of solutions to the p-Poisson equation*, Nonlinear Anal. **130** (2016), 298–329. MR 3424623

[DER07]   Lars Diening, Carsten Ebmeyer, and Michael Růžička, *Optimal convergence for the implicit space-time discretization of parabolic systems with p-structure*, SIAM J. Numer. Anal. **45** (2007), no. 2, 457–472. MR 2300281

[DFTW20] Lars Diening, M. Fornasier, R. Tomasi, and M. Wank, *A relaxed kaanov iteration for the p-poisson problem*, Numerische Mathematik (2020).

[DK08]   Lars Diening and Christian Kreuzer, *Linear convergence of an adaptive finite element method for the p-Laplacian equation*, SIAM J. Numer. Anal. **46** (2008), no. 2, 614–638. MR 2383205

[DM05]   Frank Duzaar and Giuseppe Mingione, *Second order parabolic systems, optimal regularity, and singular sets of solutions*, Ann. Inst. H. Poincaré Anal. Non Linéaire **22** (2005), no. 6, 705–751. MR 2172857

[DMS11]   Frank Duzaar, Giuseppe Mingione, and Klaus Steffen, *Parabolic systems with polynomial growth and regularity*, Mem. Amer. Math. Soc. **214** (2011), no. 1005, x+118. MR 2866816

[DR07]   L. Diening and M. Růžička, *Interpolation operators in Orlicz–Sobolev spaces*, Numer. Math. **107** (2007), no. 1, 107–129. MR 2317830

[DSS19]   Lars Diening, Toni Scharle, and Sebastian Schwarzacher, *Regularity for parabolic systems of Uhlenbeck type with Orlicz growth*, J. Math. Anal. Appl. **472** (2019), no. 1, 46–60. MR 3906361

[EF01]   Carsten Ebmeyer and Jens Frehse, *Mixed boundary value problems for nonlinear elliptic equations with p-structure in nonsmooth domains*, Differential Integral Equations **14** (2001), no. 7, 801–820. MR 1828325

[EJ95]   Kenneth Eriksson and Claes Johnson, *Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal. **32** (1995), no. 3, 706–740.

[EL05]   Carsten Ebmeyer and WB. Liu, *Quasi-norm interpolation error estimates for the piecewise linear finite element approximation of p-Laplacian problems*, Numer. Math. **100** (2005), no. 2, 233–258.

[LMW12]   A. Logg, K.-A. Mardal, and G. N. Wells (eds.), *Automated solution of differential equations by the finite element method*, Lecture Notes in Computational Science and Engineering, vol. 84, Springer, Heidelberg, 2012, The FEniCS book. MR 3075806

[Sav98]   Giuseppe Savaré, *Regularity results for elliptic equations in Lipschitz domains*, J. Funct. Anal. **152** (1998), no. 1, 176–201. MR 1600081

[SZ90]   L. Ridgway Scott and Shangyou Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp. **54** (1990), no. 190, 483–493.

[Wei92]   Dongming Wei, *Existence, uniqueness, and numerical analysis of solutions of a quasilinear parabolic problem*, SIAM J. Numer. Anal. **29** (1992), no. 2, 484–497. MR 1154277

(D. Breit) DEPARTMENT OF MATHEMATICS, HERIOT-WATT UNIVERSITY, EDINBURGH EH14 4AS, UK.
E-mail address: `d.breit@hw.ac.uk`

(L. Diening, J. Storn, J. Wichmann) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BIELEFELD, POSTFACH 10 01 31, 33501 BIELEFELD, GERMANY
E-mail address: `lars.diening@uni-bielefeld.de`
E-mail address: `jstorn@math.uni-bielefeld.de`
E-mail address: `jwichmann@math.uni-bielefeld.de`