CONVERGENCE OF A FULL DISCRETIZATION OF QUASILINEAR PARABOLIC EQUATIONS IN ISOTROPIC AND ANISOTROPIC ORLICZ SPACES*

ETIENNE EMMRICH[†] AND ANETA WRÓBLEWSKA[‡]

November 10, 2011

Abstract. Convergence of a full discretization is shown for a general class of nonlinear parabolic problems. The numerical method combines the backward Euler method for the time discretization with a generalized internal approximation scheme for the spatial discretization. The governing monotone elliptic differential operator is described by a nonlinearity that may have anisotropic and non-polynomial growth but fulfills a coercivity condition in terms of a generalized $\mathscr{N}$-function.

Key words. Nonlinear parabolic equation, monotone operator, non-standard growth condition, Orlicz space, time discretization, internal approximation, finite element method, convergence.

AMS subject classifications. 65M12, 35K55, 47H05, 47J35, 65M06, 65M60

1. Introduction.

We are concerned with the approximation of the initial-boundary value problem for a quasilinear parabolic equation that reads

$$\partial_t u - \nabla \cdot a(\nabla u) = f \text{ in } Q := \Omega \times (0,T)\,, \tag{1.1a}$$

$$u = 0 \text{ on } \partial\Omega \times (0,T)\,, \quad u(\cdot, 0) = u_0 \text{ in } \Omega\,. \tag{1.1b}$$

Here, $\Omega \subset \mathbb{R}^d$ is a bounded domain with Lipschitz boundary $\partial\Omega$ and $[0,T]$ is the time interval under consideration. For given functions $f : \overline{Q} \to \mathbb{R}$ and $u_0 : \overline{\Omega} \to \mathbb{R}$, we look for a solution $u : \overline{Q} \to \mathbb{R}$. Throughout this paper, we assume that the nonlinearity $a : \mathbb{R}^d \to \mathbb{R}^d$ is continuous as well as monotone such that

$$(a(\xi) - a(\eta)) \cdot (\xi - \eta) \geq 0 \qquad \text{for all } \xi, \eta \in \mathbb{R}^d\,.$$

Moreover, we assume that there exists an $\mathscr{N}$-function $M : \mathbb{R}^d \to \mathbb{R}_0^+$ (see Definition 2.1 below) and a constant $\mu > 0$ such that

$$a(\xi) \cdot \xi \geq \mu \left(M(\xi) + M^*(a(\xi)) \right) \qquad \text{for all } \xi \in \mathbb{R}^d\,, \tag{1.2}$$

where M^* denotes the conjugate function to M, the dot means the Euclidean inner product and $|\xi|^2 = \xi \cdot \xi$. Typical examples are (see also [6, 9, 17, 20, 31])

1) $a(\xi) = |\xi|^{p-2}\xi$ $(p > 1)$ with $M(\xi) = \frac{1}{p}|\xi|^p$, $M^*(\eta) = \frac{1}{q}|\eta|^q$ $(\frac{1}{p} + \frac{1}{q} = 1)$, which leads to the p-Laplacian (including the Laplacian for $p = 2$);
2) $a(\xi) = \xi e^{|\xi|}$ with $M(\xi) = (|\xi| - 1)\, e^{|\xi|} + 1$, $M^*(\eta) = \left(|\xi(\eta)|^2 - |\xi(\eta)| + 1\right) e^{|\xi(\eta)|} - 1$;
3) $a(\xi) = \xi \log(|\xi| + 1)$ with $M(\xi) = \frac{1}{2}(|\xi|^2 - 1)\log(|\xi| + 1) + \frac{1}{4}|\xi|(2 - |\xi|)$, $M^*(\eta) = \frac{1}{2}(|\xi(\eta)|^2 + 1)\log(|\xi(\eta)| + 1) - \frac{1}{4}|\xi(\eta)|(2 - |\xi(\eta)|)$;
4) $a(\xi) = \frac{\xi}{|\xi|}\log(|\xi| + 1) + \frac{\xi}{|\xi|+1}$ with $M(\xi) = |\xi|\log(|\xi| + 1)$, $M^*(\eta) = \frac{|\xi(\eta)|^2}{|\xi(\eta)|+1}$;

*This work was supported by the Collaborative Research Center 701, which is funded by DFG (German Research Council).

[†]Universität Bielefeld, Fakultät für Mathematik, Universitätsstraße 25, 33615 Bielefeld, Germany. (emmrich@math.uni-bielefeld.de)

[‡]University of Warsaw, Institute of Applied Mathematics and Mechanics, ul. Banacha 2, Warszawa, Poland. A.W. is a Ph.D. student in the International Ph.D. Projects Programme of Foundation for Polish Science within the Innovative Economy Operational Programme 2007-2013 (Ph.D. Programme: Mathematical Methods in Natural Sciences). (a.wroblewska@mimuw.edu.pl)

5) $a(\xi) = [|\xi_1|^{p_1-2}\xi_1, |\xi_2|^{p_2-2}\xi_2]$ $(1 < p_1,\, p_2 < \infty)$ for $\xi = [\xi_1, \xi_2] \in \mathbb{R}^2$ with $M(\xi) = \frac{1}{p_1}|\xi_1|^{p_1} + \frac{1}{p_2}|\xi_2|^{p_2}$, $M^*(\eta) = \frac{1}{q_1}|\eta_1|^{q_1} + \frac{1}{q_2}|\eta_2|^{q_2}$;

6) $a(\xi) = 2[2\xi_1 - \xi_2, 2\xi_2 - \xi_1]$ for $\xi = [\xi_1, \xi_2] \in \mathbb{R}^2$ with $M(\xi) = \xi_1^2 + \xi_2^2 + (\xi_1 - \xi_2)^2$, $M^*(\eta) = \frac{1}{6}\left(\eta_1^2 + \eta_1\eta_2 + \eta_2^2\right)$;

7) $a(\xi) = [\xi_1 e^{\xi_1^2}, \xi_2 e^{\xi_2^2}]$ for $\xi = [\xi_1, \xi_2] \in \mathbb{R}^2$ with $M(\xi) = \frac{1}{2}e^{\xi_1^2} + \frac{1}{2}e^{\xi_2^2} - 1$, $M^*(\eta) = (\xi_1(\eta)^2 - \frac{1}{2})e^{\xi_1(\eta)^2} + (\xi_2(\eta)^2 - \frac{1}{2})e^{\xi_2(\eta)^2} + 1$;

where $\xi(\eta)$ always solves the equation $\eta = a(\xi(\eta))$. Whereas the first example is well understood and can be studied by employing the standard theory of monotone operators relying on the Sobolev space $W_0^{1,p}(\Omega)$, the other examples lead to operators satisfying non-polynomial or anisotropic growth conditions instead of having the usual p-structure. These latter examples require another functional setting, namely to consider monotone operators in (isotropic or anisotropic) Orlicz spaces. Since, in general, Orlicz spaces are neither reflexive nor separable, additional difficulties arise. Applications can be found, e.g., in fluid dynamics and rheology (see [16, 18]) as well as in electrodynamics (see [4]), of course for u being $\mathbb{R}^d$-valued then, which still requires further research. Indeed, example 5) appears in the description of Prandtl–Eyring fluids. Note that the same type of equations also arises in image processing with, e.g., $a(\xi) = \xi/(1 + |\xi|^2)$ or $a(\xi) = \xi\exp(-|\xi|^2)$ (Perona–Malik equation, see [24]). Unfortunately, the underlying potential then is not convex and thus does not fit into our framework. Also the minimal surface or prescribed mean curvature equation with $a(\xi) = \xi/\sqrt{1 + |\xi|^2}$ does not fit into our framework since the corresponding potential does not grow superlinearly at infinity.

Existence of global weak solutions (solutions in the sense of distributions) has been shown in [9] if the conjugate of the underlying $\mathscr{N}$-function satisfies a Δ_2-condition (see (2.4) below). Similar results under somehow restrictive assumptions have also been obtained in [11, 25]. More recently, existence has been proved in [17] for the general case avoiding any restrictive growth or Δ_2-condition and allowing anisotropy but for problems with homogeneous right-hand side only. The method of proof relies upon a Galerkin approximation with eigenfunctions of the Laplacian. See also [12] for a similar result, including a uniqueness result, but in the isotropic case, and [13] for a generalization of [12] allowing lower order terms but requiring a more restrictive growth condition. For the case that the nonlinearity a is the gradient of a continuously differentiable potential, existence and uniqueness of the homogeneous problem is also shown in [6]. The method of proof there relies upon a time discretization and considering each time step as the Euler–Lagrange equation for a corresponding variational problem. (The limit of the sequence of approximate solutions is then identified to be an exact solution by testing with the solution itself and employing the classical Minty trick. This is said to be allowed because of an approximation argument, which is, unfortunately, not carried out.)

A main problem, which also arises in our studies, is the lack of a tensor structure of Orlicz spaces over the time-space cylinder. In the standard treatment, it is this tensor structure which allows to reduce the parabolic partial differential equation to an operator differential equation for functions in time taking values in an appropriate Banach space of functions in space.

In this paper, we study the convergence of a fully discrete approximation. Apart from the work in [4, 10, 26] on the Galerkin finite element approximation of elliptic problems described by monotone operators in Orlicz spaces, there is, to the best knowledge of the authors, no other study of numerical approximations available, especially not for problems of parabolic type.

In this first attempt to analyze time-dependent problems, we restrict our considerations to the scalar case without non-monotone perturbations (such as lower order terms). Moreover, to keep the presentation readable, we do not consider the case where a is a Carathéodory function that explicitly depends on (x, t) although we believe that this case can be treated similarly.

Our main result is the convergence of a sequence of approximate solutions towards an exact solution. The numerical approximation here comes from combining the backward Euler (or Rothe) method with a generalized internal approximation scheme. This approximation scheme covers the abstract Galerkin method but also standard conforming finite element methods as we show. The assumptions on the underlying differential operator are as general as in [17] avoiding any restrictive growth or Δ_2-condition and allowing anisotropy. In contrast to [17], we also allow non-homogeneous right-hand sides by employing estimates relying on the Bogovskii operator.

It should be noted that the convergence result provided here also implies existence of a weak solution. We also provide a uniqueness result. Moreover, we should emphasize that the method of proof here differs somehow from that in [12, 17] not only because of the full discretization. In particular, we use a certain characterization of Orlicz spaces as a weak closure (together with results on mollification and the continuity of the translation) and we omit employing knowledge about the sequence of time derivatives of the approximate solutions. The latter would require to have the boundedness of the sequence of L^2-orthogonal projections onto the finite dimensional subspaces with respect to the operator norm induced by the norm $\|\cdot\|_{2,\Omega} + \|\nabla\cdot\|_{M,\Omega}$, where $\|\cdot\|_{2,\Omega}$ denotes the L^2-norm and $\|\cdot\|_{M,\Omega}$ the Luxemburg norm. This, however, is by no way obvious for an arbitrary internal approximation scheme or a particular finite element method (but was implicitly used in [17] for the special Galerkin approximation employed there). Instead, we employ the centered Steklov average for a regularization in time and avoid compactness arguments of the Lions–Aubin type (that, e.g., have been used in [13]).

We are aware of the fact that it would be desirable to have an analysis at hand for a semi-implicit variant of the time discretization. So far, we are not able to prove convergence for such a method. We also emphasize that error estimates were not in the scope of this paper since such would require to assume higher regularity of the exact solution, which is, in general, not known even for smooth data. Moreover, available estimates of the interpolation error in [10] require the restrictive Δ_2-condition.

The outline of the paper is as follows: In Section 2, we introduce the necessary notation, recall basic facts about Orlicz spaces, and prove some auxiliary results. It follows Section 3 with the description of the numerical method, the proof of existence of a numerical solution, and the derivation of a priori estimates for the fully discrete solution. Convergence towards an exact solution is then shown in Section 4. In Section 5, we finally illustrate the numerical method for a simple example.

2. Notation and preliminaries.

2.1. General notation. By $L^p(\Omega)$ ($p \in [1, \infty]$), we denote the usual Lebesgue space, for $\mathbb{R}^d$-valued functions, we write $L^p(\Omega; \mathbb{R}^d)$, both equipped with the standard norm $\|\cdot\|_{p,\Omega}$. Moreover, we rely upon the usual notation for Sobolev spaces. In particular, we have $W^{1,p}(\Omega) = \{v \in L^p(\Omega) : \nabla v \in L^p(\Omega; \mathbb{R}^d)\}$, and $W_0^{1,p}(\Omega)$ ($p \in [1, \infty)$) denotes the closure of $\mathscr{C}_c^\infty(\Omega)$ with respect to the $W^{1,p}$-norm. Here, $\mathscr{C}_c^\infty(\Omega)$ denotes the space of infinitely differentiable functions with compact support in Ω. By

$\gamma_0 v$, we denote the trace of $v : \overline{\Omega} \to \mathbb{R}$ such that $\gamma_0 v = v$ on $\partial\Omega$ for smooth v.

For a Banach space X, we denote by $L^p(0,T;X)$ ($p \in [1,\infty]$) the usual Bochner–Lebesgue space equipped with the standard norm. We recall that $L^p(0,T;L^p(\Omega)) = L^p(Q)$ if $p < \infty$. Here, we identify the abstract function $u : [0,T] \to L^p(\Omega)$ with the function $u : \overline{Q} \to \mathbb{R}$ via $[u(t)](x) = u(x,t)$. The standard norm is then denoted by $\|\cdot\|_{p,Q}$. By $\mathscr{C}([0,T];X)$, we denote the usual space of continuous functions $u : [0,T] \to X$, whereas $\mathscr{C}_w([0,T];X)$ denotes the space of demicontinuous functions (i.e., continuous with respect to the weak topology in X). See also [14] for more details. By $\langle\cdot,\cdot\rangle$, we denote the duality pairing. Finally, c denotes a generic positive constant.

2.2. Orlicz spaces. In this section, we recall the definition of Orlicz spaces and some of their properties (see, especially, [20] for a very readable introduction as well as [1, 15, 19, 27, 28, 31]). Let us emphasize that our considerations include noninearities with anisotropic growth. We, therefore, rely upon anisotropic Orlicz classes and spaces defined by $\mathscr{N}$-functions with vector-valued arguments (see, in particular, [9, 27, 28]).

DEFINITION 2.1 ($\mathscr{N}$-function). *A function $M : \mathbb{R}^d \to \mathbb{R}$ is said to be an $\mathscr{N}$-function if it satisfies the following conditions:*

(i) M is continuous, $M(\xi) = 0$ if and only if $\xi = 0$, $M(\xi) = M(-\xi)$ for all $\xi \in \mathbb{R}^d$;

(ii) M is convex;

(iii) M has superlinear growth such that $\lim_{|\xi|\to 0} \frac{M(\xi)}{|\xi|} = 0$, $\lim_{|\xi|\to\infty} \frac{M(\xi)}{|\xi|} = \infty$.

Some authors prefer the term *generalized $\mathscr{N}$-function* in order to emphasize the dependence on ξ and not only on $|\xi|$. Note that (i) and (ii) imply $M(\xi) \geq 0$ for all $\xi \in \mathbb{R}^d$. Because of the anisotropic character, the function M need not be a function that is increasing with respect to the components of its vector-valued argument (see, e.g., example 6) in the introduction).

For an $\mathscr{N}$-function M, we denote by M^* the conjugate function given by the Legendre–Fenchel transform $M^*(\eta) = \sup_{\xi\in\mathbb{R}^d} (\xi\cdot\eta - M(\xi))$ ($\eta \in \mathbb{R}^d$). The conjugate function M^* is also an $\mathscr{N}$-function (see [27]). An important tool in deriving a priori estimates will be the Fenchel–Young inequality

$$|\xi\cdot\eta| \leq M(\xi) + M^*(\eta) \qquad \text{for all } \xi,\eta \in \mathbb{R}^d\,. \tag{2.1}$$

The *anisotropic Orlicz class* $\mathscr{L}_M(\Omega;\mathbb{R}^d)$ is the set of all (equivalence classes of almost everywhere equal) measurable functions $\xi : \Omega \to \mathbb{R}^d$ such that

$$\rho_{M,\Omega}(\xi) := \int_\Omega M(\xi(x))\,dx < \infty\,.$$

Although $\mathscr{L}_M(\Omega;\mathbb{R}^d)$ is a convex set it may not be a linear space. The mapping $\rho_{M,\Omega}$ is a modular in the sense of [20, p. 208].

Since the function $M : \mathbb{R}^d \to \mathbb{R}$ is continuous, $\xi = \xi(x) \in L^\infty(\Omega;\mathbb{R}^d)$ implies $x \mapsto M(\xi(x)) \in L^\infty(\Omega)$, which shows that $L^\infty(\Omega;\mathbb{R}^d) \subseteq \mathscr{L}_M(\Omega;\mathbb{R}^d)$.

The *anisotropic Orlicz space* $L_M(\Omega;\mathbb{R}^d)$ is defined as the linear hull of $\mathscr{L}_M(\Omega;\mathbb{R}^d)$. It is a Banach space with respect to the Luxemburg norm

$$\|\xi\|_{M,\Omega} := \inf\left\{\lambda > 0 : \int_\Omega M\left(\frac{\xi(x)}{\lambda}\right)dx \leq 1\right\};$$

the infimum is attained if $\xi \neq 0$. In general, $L_M(\Omega;\mathbb{R}^d)$ is neither separable nor reflexive. Note that $\rho_{M,\Omega}(\xi) \leq \|\xi\|_{M,\Omega}$ if $\|\xi\|_{M,\Omega} \leq 1$, $\rho_{M,\Omega}(\xi) \geq \|\xi\|_{M,\Omega}$ if $\|\xi\|_{M,\Omega} > 1$ for all $\xi \in L_M(\Omega;\mathbb{R}^d)$, and thus $\|\xi\|_{M,\Omega} \leq \rho_{M,\Omega}(\xi) + 1$. Moreover, if $\xi \in L_M(\Omega;\mathbb{R}^d)$

then there exists $\lambda > 0$ such that $\rho_{M,\Omega}(\xi/\lambda) < \infty$. Finally, because of the superlinear growth of M, there holds

$$L_M(\Omega;\mathbb{R}^d) \subseteq L^1(\Omega;\mathbb{R}^d)\,. \tag{2.2}$$

This can be seen from the following observations: Let $\xi \in L_M(\Omega;\mathbb{R}^d)$, $\xi \neq 0$, and set $\lambda = \|\xi\|_{M,\Omega} > 0$ such that $\rho_{M,\Omega}(\xi/\lambda) \leq 1$. We set

$$\Omega_1 := \left\{ x \in \Omega : M\left(\frac{\xi(x)}{\lambda}\right) \geq \frac{|\xi(x)|}{\lambda} \right\}, \quad \Omega_2 := \Omega \setminus \Omega_1\,.$$

Since $M(\eta)/|\eta| \to \infty$ as $|\eta| \to \infty$, there exists $C > 0$ such that $|\xi(x)| \leq C$ for all $x \in \Omega_2$. We, therefore, find

$$\begin{aligned}\int_\Omega |\xi(x)|\,dx &= \lambda \int_{\Omega_1} \frac{|\xi(x)|}{\lambda}\,dx + \int_{\Omega_2} |\xi(x)|\,dx \leq \lambda \int_{\Omega_1} M\left(\frac{\xi(x)}{\lambda}\right)\,dx + C|\Omega_2| \\ &\leq \lambda\,\rho_{M,\Omega}\left(\frac{\xi}{\lambda}\right) + C|\Omega_2| \leq \|\xi\|_{M,\Omega} + C|\Omega_2| < \infty\,.\end{aligned}$$

Clearly, the anisotropic Orlicz class and space coincide with the *isotropic Orlicz class and space*, respectively, if the $\mathscr{N}$-function $M = M(\xi)$ only depends on $|\xi|$ rather than on ξ.

Let us denote by $E_M(\Omega;\mathbb{R}^d)$ the closure with respect to the Luxemburg norm of the set of bounded measurable functions defined on Ω. It turns out that $E_M(\Omega;\mathbb{R}^d)$ is the largest linear space contained in the Orlicz class $\mathscr{L}_M(\Omega;\mathbb{R}^d)$ such that

$$E_M(\Omega;\mathbb{R}^d) \subseteq \mathscr{L}_M(\Omega;\mathbb{R}^d) \subseteq L_M(\Omega;\mathbb{R}^d)\,,$$

with, in general, strict inclusion. From the equivalence of the Luxemburg and the Orlicz norm

$$|||\xi|||_{M,\Omega} := \sup\left\{ \int_\Omega \xi \cdot \eta\,dx : \eta \in \mathscr{L}_{M^*}(\Omega;\mathbb{R}^d) \text{ with } \rho_{M^*,\Omega}(\eta) \leq 1 \right\},$$

one immediately finds that $L^\infty(\Omega;\mathbb{R}^d)$ is continuously embedded in $E_M(\Omega;\mathbb{R}^d)$.

The space $E_M(\Omega;\mathbb{R}^d)$ is separable and $\mathscr{C}_c^\infty(\Omega;\mathbb{R}^d)$ is dense in $E_M(\Omega;\mathbb{R}^d)$. The space $L_M(\Omega;\mathbb{R}^d)$ is the dual of $E_{M^*}(\Omega;\mathbb{R}^d)$, the duality pairing is given by

$$\langle \xi, \eta \rangle = \int_\Omega \xi \cdot \eta\,dx\,, \quad \xi \in L_M(\Omega;\mathbb{R}^d)\,,\ \eta \in E_{M^*}(\Omega;\mathbb{R}^d)\,.$$

At this point, we may recall the *generalized Hölder inequality*

$$\int_\Omega \xi \cdot \eta\,dx \leq 2\,\|\xi\|_{M,\Omega}\,\|\eta\|_{M^*,\Omega} \qquad \text{for all } \xi \in L_M(\Omega;\mathbb{R}^d)\,,\ \eta \in L_{M^*}(\Omega;\mathbb{R}^d)\,,$$

which shows that $\xi \cdot \eta \in L^1(\Omega)$ if $\xi \in L_M(\Omega;\mathbb{R}^d)$ and $\eta \in L_{M^*}(\Omega;\mathbb{R}^d)$. (The factor 2 is due to the use of the Luxemburg norm instead of the Orlicz norm.)

It is worth to mention that for any $\xi \in L_M(\Omega;\mathbb{R}^d)$

$$\lim_{k\to\infty} \int_\Omega (\xi - \xi_k) \cdot \eta\,dx = 0 \qquad \text{for all } \eta \in L_{M^*}(\Omega;\mathbb{R}^d)\,, \tag{2.3}$$

where $\xi_k(x) = \xi(x)$ if $|\xi(x)| \le k$ and $\xi_k(x) = 0$ otherwise. This shows that $L_M(\Omega;\mathbb{R}^d)$ is the closure of $E_M(\Omega;\mathbb{R}^d)$ with respect to the weak convergence in $E_M(\Omega;\mathbb{R}^d)$ (see, e.g., [19, p. 131]). It will later be important to see that (2.3) not only holds for all $\eta \in E_{M^*}(\Omega;\mathbb{R}^d)$ but for all $\eta \in L_{M^*}(\Omega;\mathbb{R}^d)$. This is seen as follows: Because of the generalized Hölder inequality, we already know that $\xi \cdot \eta \in L^1(\Omega)$. Therefore,

$$\int_\Omega (\xi - \xi_k)\cdot \eta \, dx = \int_{\Omega_k} \xi \cdot \eta \, dx \quad \text{with } \Omega_k := \{x \in \Omega : |\xi(x)| > k\}$$

is well-defined. In view of Chebyshev's inequality, we have that $|\Omega_k| \le \frac{1}{k}\,\|\xi\|_{1,\Omega}$. The absolute continuity of the integral over the integrable function $\xi\cdot\eta$ finally proves

$$\int_{\Omega_k} \xi \cdot \eta \, dx \to 0 \quad \text{as } k \to \infty\,.$$

If the $\mathscr{N}$-function M satisfies the so-called Δ_2-*condition*, i.e., if there exists $c > 0$ such that

$$M(2\xi) \le cM(\xi) \quad \text{for all } \xi \in \mathbb{R}^d, \tag{2.4}$$

then $\mathscr{L}_M(\Omega;\mathbb{R}^d) = L_M(\Omega;\mathbb{R}^d) = E_M(\Omega;\mathbb{R}^d)$ (see [1, 20, 28]). The Δ_2-condition is, however, rather restrictive.

In the sequel, we also consider Orlicz classes and spaces over the time-space cylinder Q (the definitions and results from above are the same, just replace Ω by Q). We emphasize that, without strong assumption on M and M^*, there is no tensor structure in L_M. Thus, in general, $L_M(Q) \ne L_M(0,T;L_M(\Omega))$ (see [9, Proposition 1.3]).

2.3. Preliminary results. In this section, we summarize a few preliminary results such as the weak sequential lower semicontinuity of the modular with respect to the weak convergence in $L^1(\Omega;\mathbb{R}^d)$, an approximation result and a useful estimate relying on the Bogovskii operator.

LEMMA 2.2. *Let $\{\xi_\ell\} \subset \mathscr{L}_M(Q;\mathbb{R}^d)$ be a bounded sequence, i.e., there exists $C > 0$ such that $\rho_{M,Q}(\xi_\ell) \le C$ for all $\ell \in \mathbb{N}$. Then there exists $\xi \in \mathscr{L}_M(Q;\mathbb{R}^d)$ and a subsequence, denoted by ℓ', such that $\xi_{\ell'} \rightharpoonup \xi$ in $L^1(Q;\mathbb{R}^d)$ and $\rho_{M,Q}(\xi) \le \liminf_{\ell'\to\infty} \rho_{M,Q}(\xi_{\ell'})$.*

Proof. In a first step, we prove that the sequence $\{\xi_\ell\}$ is weakly relatively compact in $L^1(Q;\mathbb{R}^d)$. Since $\mathscr{L}_M(Q;\mathbb{R}^d) \subseteq L^1(Q;\mathbb{R}^d)$ (see (2.2)) and in view of the Dunford–Pettis theorem (see, e.g., [2, Thm. 2.4.5]), it remains to prove equi-integrability of the sequence. This, however, follows from a result analogous to the de la Vallée–Poussin theorem, and we closely follow [2, Thm. 2.4.4 on p. 58]. Since M has superlinear growth, there exists for every $K > 0$ a constant $C_K > 0$ such that

$$0 \le |\xi| \le \frac{M(\xi)}{K} + C_K \quad \text{for all } \xi \in \mathbb{R}^d\,.$$

Let $A \subseteq Q$ be a measurable subset with measure $|A|$. Then for all $\ell \in \mathbb{N}$

$$\int_A |\xi_\ell(x)|\,dx \le \frac{1}{K}\int_A M(\xi_\ell(x))\,dx + C_K\,|A| \le \frac{1}{K}\,\rho_{M,Q}(\xi_\ell) + C_K\,|A| \le \frac{C}{K} + C_K\,|A|\,.$$

For $A = \Omega$, this shows the boundedness of the sequence $\{\xi_\ell\}$ in $L^1(Q;\mathbb{R}^d)$. Let $\varepsilon > 0$ be arbitrary and set $K = 2C/\varepsilon$, $\delta = \varepsilon/(2C_K)$. We then obtain for any A with $|A| < \delta$

$$\int_A |\xi_\ell(x)|\,dx \le \frac{\varepsilon}{2}\left(1 + \frac{|A|}{\delta}\right) < \varepsilon\,,$$

which finally proves equi-integrablity. Hence, a subsequence of $\{\xi_\ell\}$ converges weakly in $L^1(Q;\mathbb{R}^d)$ towards an element $\xi \in L^1(Q;\mathbb{R}^d)$.

In a second step, we show the weak sequential lower semicontinuity of the modular in $L^1(Q;\mathbb{R}^d)$. This, however, is an immediate consequence of the convexity and continuity of $M = M(\xi)$ together with [2, Thm. 13.1.1 on p. 498] (see also [8, Thm. 3.20 on p. 94] upon noting that $M(\xi) \geq \eta \cdot \xi - M^*(\eta)$ for any $\eta \in \mathbb{R}^d$ and all $\xi \in \mathbb{R}^d$). It also proves $\xi \in \mathscr{L}_M(Q;\mathbb{R}^d)$. ◻

Unfortunately, the method of truncation as employed in (2.3) is not always appropriate when working with gradients. We, therefore, provide the following result. Let

$$J_0(x,t) := \begin{cases} c_0 \exp\left(-\dfrac{1}{1-|x|^2-t^2}\right) & \text{if } |x|^2+t^2<1\,,\\ 0 & \text{otherwise,}\end{cases}$$

where $c_0 > 0$ is such that $\int_{\mathbb{R}^d\times\mathbb{R}} J_0(x,t)\,dxdt = 1$, and set for sufficiently small $\delta > 0$

$$J_\delta(x,t) = \delta^{-(d+1)} J_0(\delta^{-1}x, \delta^{-1}t)\,, \quad (x,t) \in \mathbb{R}^d \times \mathbb{R}\,.$$

For any locally integrable function $u = u(x,t)$, the mollification $J_\delta * u$ is then a smooth function with compact support on the ball with $|x|^2 + t^2 \leq \delta^2$.

Lemma 2.3. *Let $w \in \mathscr{W} := \{w \in W^{1,1}(0,T;L^2(\Omega)) : \nabla w \in \mathscr{L}_M(Q;\mathbb{R}^d)\,,\ \gamma_0 w(\cdot,t) = 0$ a.e. in $(0,T) \ni t\}$. For any $\varepsilon > 0$ there is then a smooth function w_ε, which vanishes at $\partial\Omega \times [0,T]$ such that*

$$\|w_\varepsilon - w\|_{W^{1,1}(0,T;L^2(\Omega))} < \frac{\varepsilon}{2}$$

and such that for all $\eta \in L_{M^*}(Q;\mathbb{R}^d)$

$$\Big|\int_Q \nabla w_\varepsilon \cdot \eta\,dxdt - \int_Q \nabla w \cdot \eta\,dxdt\Big| < \frac{\varepsilon}{2}\,.$$

Proof. Note that $W^{1,1}(0,T;L^2(\Omega))$ is continuously embedded in $\mathscr{C}([0,T];L^2(\Omega))$. Moreover, since $\nabla w \in \mathscr{L}_M(Q;\mathbb{R}^d) \subset L^1(0,T;L^1(\Omega))$, the trace $\gamma_0 w(\cdot,t)$ is well-defined for almost all $t \in (0,T)$. The proof follows, in particular, from the continuity of mollification and translation of a function in $L_M(Q;\mathbb{R}^d)$ with respect to the weak convergence in $E_M(Q;\mathbb{R}^d)$ (see [15, Lemma 1.5, 1.6] and [9, Prop. 1.2]) together with standard arguments.

Let $\varepsilon > 0$. Then there is $n \in \mathbb{N}$ such that

$$\|T_n(w) - w\|_{W^{1,1}(0,T;L^2(\Omega))} < \frac{\varepsilon}{4}\,, \tag{2.5}$$

where for $(x,t) \in \overline{Q}$

$$[T_n(w)](x,t) := \begin{cases} w(x,t) & \text{if } |w(x,t)| \leq n\,,\\ n & \text{if } w(x,t) > n\,,\\ -n & \text{if } w(x,t) < -n\,.\end{cases}$$

In order to prove (2.5), we recall that $w \in \mathscr{C}([0,T];L^2(\Omega))$ and that, for each $t \in [0,T]$, the set $\Omega_n(t) := \{x \in \Omega : |w(x,t)| > n\}$ is measurable with measure

$|\Omega_n(t)| \leq \frac{1}{n^2}\|w\|^2_{\mathscr{C}([0,T];L^2(\Omega))}$. An application of Lebesgue's theorem on dominated convergence then shows, in particular, that

$$\int_0^T \Big(\int_{\Omega_n(t)} |\partial_t w(x,t)|^2 \, dx \Big)^{1/2} dt \to 0 \text{ as } n \to \infty\,.$$

Note that the truncation above is in $L^\infty(Q) \subset L_M(Q)$. Obviously, $\gamma_0[T_n(w)](\cdot,t) = 0$ for almost all $t \in (0,T)$. Furthermore, we have $[\nabla T_n(w)](x,t) = \nabla w(x,t)$ if $|w(x,t)| \leq n$ and $[\nabla T_n(w)](x,t) = 0$ otherwise.

Since $\nabla w \cdot \eta \in L^1(Q)$ for any $\eta \in L_{M^*}(Q;\mathbb{R}^d)$, the absolute continuity of the integral also shows (using Chebyshev's inequality and the same argumentation as on page 6) for sufficiently large n that

$$\left| \int_Q \Big(\nabla T_n(w) - \nabla w \Big) \cdot \eta \, dxdt \right| < \frac{\varepsilon}{4}\,.$$

Since Ω is a Lipschitz domain and $\partial\Omega$ is compact, there is a finite number of points $x^j \in \partial\Omega$, radii $r^j > 0$ and Lipschitz continuous functions $\lambda^j : \mathbb{R}^{d-1} \to \mathbb{R}$ $(j = 1,2,\ldots,J)$ such that –up to a rigid motion if necessary–

$$\Omega \cap \Omega_j = \{x = [x_1,\ldots,x_{d-1},x_d] \in \Omega_j : x_d < \lambda^j(x_1,\ldots,x_{d-1})\}\,,$$

where $\Omega_j \subset \mathbb{R}^d$ denotes the open ball of radius r^j with origin x^j. For sufficiently small $\delta_0 > 0$, we may also assume that $\{\Omega_j\}_{j=0}^J$ with $\Omega_0 := \{x \in \Omega : \operatorname{dist}(x,\partial\Omega) > \delta_0\}$ is an open cover of $\overline{\Omega}$. Let $\{\chi_j\}_{j=0}^J$ be a smooth partition of unity for $\overline{\Omega}$ subordinate to this open cover.

For sufficiently small $r > 0$, the intervals $I_0 = (r/2, T - r/2)$, $I_1 = (-r,r)$ and $I_2 = (T-r, T+r)$ build an open cover of $[0,T]$. Let $\{\zeta_k\}_{k=0}^2$ be a smooth partition of unity for $[0,T]$ subordinate to this open cover.

It is clear that $\{\Omega_j \times I_k\}_{j=0,k=0}^{J,2}$ is an open cover of $\overline{Q}$ and that $\{\chi_j\zeta_k\}_{j=0,k=0}^{J,2}$ is a smooth partition of unity subordinate to this open cover. In particular, we have $T_n(w) = \sum_{j,k} w_{jk}$, where $w_{jk} := \chi_j\zeta_k T_n(w)$, and $\operatorname{supp} w_{jk} \subset \Omega_j \times I_k$.

We observe that $w_{00} \in W^{1,1}(0,T;L^2(\Omega))$ with $\nabla w_{00} = \nabla\chi_0\zeta_0 T_n(w) + \chi_0\zeta_0\nabla T_n(w) \in L_M(Q;\mathbb{R}^d)$ and $\operatorname{supp} w_{00} \subset Q$. The mollification is continuous in $W^{1,1}(0,T;L^2(\Omega))$ with respect to the strong convergence, which can be shown by standard arguments (employing, in particular, the continuity of the translation in $L^1(0,T;L^2(\Omega))$, which follows from Lusin's theorem). Moreover, the mollification of a function in the Orlicz space $L_M(Q;\mathbb{R}^d)$ is continuous with respect to the weak convergence in $E_M(Q;\mathbb{R}^d)$. There exists, therefore, a sufficiently small number $\delta_{00} > 0$ such that

$$\|J_{\delta_{00}} * w_{00} - w_{00}\|_{W^{1,1}(0,T;L^2(\Omega))} < \frac{\varepsilon}{12(J+1)}$$

and such that for all $\eta \in L_{M^*}(Q;\mathbb{R}^d)$.

$$\left| \int_Q \Big(\nabla(J_{\delta_{00}} * w_{00}) - \nabla w_{00} \Big) \cdot \eta \, dxdt \right| < \frac{\varepsilon}{12(J+1)}\,.$$

Here we have also used that $\nabla(J_{\delta_{00}} * T_n(w)) = \nabla J_{\delta_{00}} * T_n(w) \in E_M(Q;\mathbb{R}^d)$.

For $(j,k) \neq (0,0)$, we observe the following. Since the translation is continuous in $W^{1,1}(0,T;L^2(\Omega))$ with respect to the strong convergence and continuous in

$L_M(Q;\mathbb{R}^d)$ with respect to the weak convergence in $E_M(Q;\mathbb{R}^d)$ and since translation and derivative commute, there exist sufficiently small numbers $\delta_j > 0$ and $\tau_k > 0$ such that

$$\|\widetilde{w}_{jk} - w_{jk}\|_{W^{1,1}(0,T;L^2(\Omega))} < \frac{\varepsilon}{24(J+1)}$$

and such that for all $\eta \in L_{M^*}(Q;\mathbb{R}^d)$.

$$\left|\int_Q \Big(\nabla \widetilde{w}_{jk} - \nabla w_{jk}\Big)\cdot \eta \, dxdt\right| < \frac{\varepsilon}{24(J+1)}\,,$$

where $\widetilde{w}_{jk}(x_1,\dots,x_{d-1},x_d,t) := \overline{w}_{jk}(x_1,\dots,x_{d-1},x_d+\delta_j, t-(-1)^k\tau_k)$ for $(x,t) \in \mathbb{R}^d \times \mathbb{R}$ and where $\overline{w}_{jk}$ is the extension of w_{jk} by zero outside $\overline{Q}$. Note that the translation with respect to space is inwards whereas the translation with respect to time is outwards. This takes into account that $T_n(w)$ and thus w_{jk} has vanishing trace at $\partial\Omega$. By construction, the restriction of $\widetilde{w}_{jk}$ to $K \times [-\tau_1, T+\tau_2]$ for any compact subset $K \subset \mathbb{R}^d$ has the same regularity as w on $\overline{Q}$, and $\operatorname{supp} \widetilde{w}_{jk} \subset \Omega \times I'_k$, where $I'_1 = [-\tau_1, r-\tau_1]$ and $I'_2 = [T-r+\tau_2, T+\tau_2]$.

There also exist sufficiently small numbers $\delta_{jk} > 0$ such that

$$\|J_{\delta_{jk}} * \widetilde{w}_{jk} - \widetilde{w}_{jk}\|_{W^{1,1}(0,T;L^2(\Omega))} < \frac{\varepsilon}{24(J+1)}$$

and such that for all $\eta \in L_{M^*}(Q;\mathbb{R}^d)$.

$$\left|\int_Q \Big(\nabla (J_{\delta_{jk}} * \widetilde{w}_{jk}) - \nabla \widetilde{w}_{jk}\Big)\cdot \eta \, dxdt\right| < \frac{\varepsilon}{24(J+1)}\,.$$

Putting altogether shows (with the convention $\widetilde{w}_{00} = w_{00}$) that $w_\varepsilon := \sum_{j,k} J_{\delta_{jk}} * \widetilde{w}_{jk}$ satisfies the asserted estimates. Moreover, the restriction of w_ε to $\overline{Q}$ vanishes at $\partial\Omega \times [0,T]$ because of $\operatorname{supp} \widetilde{w}_{jk} \subset \Omega \times I'_k$. □

LEMMA 2.4. *Let $\varepsilon \in (0,1)$ and $g \in L^q(\Omega)$ with $q > d$ be given. Then there exists $C > 0$ such that*

$$\int_\Omega gv \, dx \le C|\Omega| + \frac{1}{2|\Omega|}\|g\|_{1,\Omega}^2 + \frac{1}{2}\|v\|_{2,\Omega}^2 + \varepsilon \int_\Omega M(\nabla v)\, dx \tag{2.6}$$

for all $v \in \mathscr{V} := \{v \in L^2(\Omega) : \nabla v \in \mathscr{L}_M(\Omega;\mathbb{R}^d),\ \gamma_0 v = 0\}$. Here, C is given by $C = \sup_{\eta\in\mathbb{R}^d,\, |\eta| \le c\varepsilon^{-1}\|g\|_{q,\Omega}} M^*(\eta)$, where $c > 0$ only depends on Ω, d, and q.

Proof. First, we recall that $\{v \in L^2(\Omega) : \nabla v \in \mathscr{L}_M(\Omega;\mathbb{R}^d)\} \subset W^{1,1}(\Omega) \cap L^2(\Omega)$. Therefore, also the trace $\gamma_0 v$ is well-defined (see, e.g., [22, Thm. 4.2 on p. 84]). The proof uses properties of the Bogovskii operator (see [23, Lemma 3.17], [29, Lemma 2.1.1 on p. 68]). In fact, if $g \in L^q(\Omega)$ then there exists $\tilde{g} \in W_0^{1,q}(\Omega;\mathbb{R}^d)$ such that

$$g = \frac{1}{|\Omega|}\int_\Omega g\, dx + \nabla \cdot \tilde{g}$$

and

$$\|\tilde{g}\|_{W_0^{1,q}(\Omega;\mathbb{R}^d)} = \|\nabla \tilde{g}\|_{q,\Omega} \le c\left\| g - \frac{1}{|\Omega|}\int_\Omega g\, dx\right\|_{q,\Omega} \le c\,\|g\|_{q,\Omega}\,,$$

where $c > 0$ depends on q and Ω. It then follows from integration by parts that

$$\int_\Omega gv\,dx = \frac{1}{|\Omega|}\int_\Omega g\,dx \int_\Omega v\,dx - \int_\Omega \tilde{g}\cdot\nabla v\,dx\,. \tag{2.7}$$

With the Fenchel–Young inequality (2.1) and the properties of M (convexity and $M(0)=0$), we find

$$\begin{aligned}\left|\int_\Omega \tilde{g}\cdot\nabla v\,dx\right| &\leq \int_\Omega M^*(\varepsilon^{-1}\,\tilde{g})\,dx + \int_\Omega M(\varepsilon\,\nabla v)\,dx\\ &\leq \int_\Omega M^*(\varepsilon^{-1}\,\tilde{g})\,dx + \varepsilon\int_\Omega M(\nabla v)\,dx\,.\end{aligned}$$

Since $q > d$, we have the continuous embedding $W_0^{1,q}(\Omega;\mathbb{R}^d) \hookrightarrow \mathscr{C}(\overline{\Omega};\mathbb{R}^d)$ such that for all $x \in \Omega$

$$|\tilde{g}(x)| \leq \|\tilde{g}\|_{\infty,\Omega} \leq c\,\|\tilde{g}\|_{W_0^{1,q}(\Omega;\mathbb{R}^d)} \leq c\,\|g\|_{q,\Omega}\,,$$

where $c > 0$ depends on q, Ω, d. Since $M^* : \mathbb{R}^d \to \mathbb{R}_0^+$ is continuous, we find that

$$\sup_{x\in\Omega} M^*(\varepsilon^{-1}\,\tilde{g}(x)) \leq \sup_{\eta\in\mathbb{R}^d,\,|\eta|\leq c\varepsilon^{-1}\,\|g\|_{q,\Omega}} M^*(\eta) =: C < \infty\,.$$

This, together with

$$\int_\Omega g\,dx\int_\Omega v\,dx \leq \|g\|_{1,\Omega}\|v\|_{1,\Omega} \leq |\Omega|^{1/2}\|g\|_{1,\Omega}\|v\|_{2,\Omega} \leq \frac{1}{2}\,\|g\|_{1,\Omega}^2 + \frac{|\Omega|}{2}\,\|v\|_{2,\Omega}^2\,,$$

proves the assertion. □

We should remark that we do not claim that the assumptions in the previous lemma are optimal in order to get the estimate (2.6).

3. A full discretization. In this section, we describe the numerical method that combines a generalized internal approximation scheme (such as a Galerkin scheme or a conforming finite element method, see [30]) for the spatial discretization with the backward Euler scheme for the temporal discretization.

3.1. Discretization. We consider an equidistant time grid: For $N \in \mathbb{N}$ ($N \geq 1$), let $\tau = T/N$ and $t_n = n\tau$ ($n = 0, 1, \ldots, N$). Moreover, we consider a generalized internal approximation of the space

$$V := \{v \in L^2(\Omega) : \nabla v \in E_M(\Omega;\mathbb{R}^d)\,,\ \gamma_0 v = 0\}\,,\quad \|v\|_V := \|v\|_{2,\Omega} + \|\nabla v\|_{M,\Omega}\,,$$

which is given by a sequence of (not necessarily nested) finite dimensional subspaces $V_m \subset V$ ($m \in \mathbb{N}$) and restriction operators $R_m : V \to V_m$ such that for any sequence $\{m_\ell\}_{\ell\in\mathbb{N}}$ with $m_\ell \to \infty$ as $\ell \to \infty$ there holds

$$R_{m_\ell} v \to v \quad \text{in } V \text{ as } \ell \to \infty \text{ for all } v \in V\,. \tag{3.1}$$

Since V is a separable Banach space there always exists a Galerkin basis and thus an internal approximation scheme for V. Note that it suffices if the restriction operators are defined on (and the strong convergence takes place for) a dense subset of V (see, e.g., [30, pp. 25ff.]).

EXAMPLE 3.1 (Finite element approximation). We shortly describe how to construct a generalized internal approximation scheme that satisfies (3.1) from finite elements. Let $\{V_m\}_{m\in\mathbb{N}}$ be a sequence of finite element spaces such that $V_m \subset W^{1,\infty}(\Omega)$ and let I_m denote the corresponding global interpolation operator (see, e.g., [7, Sect. 12]). We assume that I_m can be defined at least on $\mathscr{C}^2(\overline{\Omega})$ and that

$$\|I_{m_\ell} v - v\|_{W^{1,\infty}(\Omega)} \to 0 \text{ as } \ell \to \infty \text{ for all } v \in \mathscr{C}^2(\overline{\Omega})$$

for any sequence $\{m_\ell\}_{\ell\in\mathbb{N}}$ with $m_\ell \to \infty$ as $\ell \to \infty$, which implies

$$\|I_{m_\ell} v - v\|_V \to 0 \text{ as } \ell \to \infty \text{ for all } v \in \mathscr{C}^2(\overline{\Omega})\,.$$

This is, e.g., fulfilled for conforming $\mathscr{P}^1$ (or rectangular $\mathscr{Q}^1$) elements corresponding to a regular affine family of triangulations of a polyhedral domain Ω (or a domain Ω that is the union of d-dimensional rectangles), see [5, Thm. 4.4.20, 4.6.14], [7, Thm. 16.2].

For the construction of the restriction operators $R_m : V \to V_m$, we follow [30, p. 28]). If $v \in \mathscr{C}^2(\overline{\Omega})$ then $R_m v := I_m v$ for all $m \in \mathbb{N}$. Otherwise, there is a sequence $\{v_n\}_{n\in\mathbb{N}} \subset \mathscr{C}^2(\overline{\Omega})$ such that $\|v - v_n\|_V < 1/n$ (note that $\mathscr{C}^2(\overline{\Omega})$ is dense in V) and a sequence $\{m_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$ such that $\|I_m v_n - v\|_V < 1/n$ for all $m \geq m_n$. We may suppose that $\{m_n\}$ is increasing. We now set $R_m v := I_m v_n$ if $m_n \leq m < m_{n+1}$. It then follows $\|R_m v - v\|_V < 2/n$, which shows (3.1).

For another construction of restriction operators and for estimates of the interpolation error assuming the Δ_2-condition, we refer to [10]. □

The numerical method under consideration now reads as follows: *Find* $\{u^n\}_{n=1}^N \subset V_m$ *such that for* $n = 1, 2, \ldots, N$

$$\int_\Omega \left(\frac{u^n - u^{n-1}}{\tau} v + a(\nabla u^n) \cdot \nabla v \right) dx = \int_\Omega f(\cdot, t_n) v \, dx \quad \text{for all } v \in V_m\,. \tag{3.2}$$

Here, $u^0 \in V_m$ denotes a suitable approximation of the initial datum $u_0 \in L^2(\Omega)$. Moreover, we have assumed that f is continuous with respect to time. If this is not the case, one may work with $f^n = \frac{1}{\tau}\int_{t_{n-1}}^{t_n} f(\cdot, t)dt$ $(n = 1, 2, \ldots, N)$ instead of $f(\cdot, t_n)$.

3.2. Solvability. We are now going to show that there exists a solution to the numerical scheme (3.2).

THEOREM 3.2 (Existence of discrete solution). *Let* $f \in \mathscr{C}([0,T]; L^q(\Omega))$ *with* $q > d$ *and* $u^0 \in V_m$. *Let* $\tau < 1$. *Then there exists a solution* $\{u^n\}_{n=1}^N \subset V_m$ *to (3.2).*

The proof relies upon the following auxiliary result, which is a direct consequence of Brouwer's fixed point theorem (see, e.g., [14, p. 74]).

LEMMA 3.3. *For some* $R > 0$, *let* $\boldsymbol{h} : \overline{B}(0,R) \to \mathbb{R}^m$ *be continuous, where* $\overline{B}(0,R) \subset \mathbb{R}^m$ *denotes the closed ball with respect to some norm* $\|\cdot\|_{\mathbb{R}^m}$ *on* $\mathbb{R}^m$. *If*

$$\boldsymbol{h}(\boldsymbol{v}) \cdot \boldsymbol{v} \geq 0 \quad \text{for all } \boldsymbol{v} \in \mathbb{R}^m \text{ with } \|\boldsymbol{v}\|_{\mathbb{R}^m} = R$$

then there exists $\tilde{\boldsymbol{v}} \in \overline{B}(0,R)$ *such that* $\boldsymbol{h}(\tilde{\boldsymbol{v}}) = 0$.

Proof. [of Theorem 3.2] We prove the existence step-by-step. So let us assume we are given $u^{n-1} \in L^2(\Omega)$. We then show existence of $u^n \in V_m$ satisfying (3.2).

Since V_m is finite dimensional, we have $V_m = \text{span}\{\varphi_1, \varphi_2, \ldots, \varphi_m\}$ for a suitable set of basis functions (without loss of generality, we may assume that the index m in the notation of V_m equals the dimension of V_m). We then have a one-to-one mapping between V_m and $\mathbb{R}^m$ given by

$$\boldsymbol{v} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_m] \in \mathbb{R}^m \quad \leftrightsquigarrow \quad V_m \ni v = \sum_{j=1}^{m} \boldsymbol{v}_j \varphi_j \, ,$$

and $\|\boldsymbol{v}\|_{\mathbb{R}^m} := \|v\|_{2,\Omega}$ defines a norm on $\mathbb{R}^m$. We now define the mapping $\boldsymbol{h}$ via

$$\boldsymbol{h}_i(\boldsymbol{v}) := \int_\Omega \left(\frac{v - u^{n-1}}{\tau} \varphi_i + a(\nabla v) \cdot \nabla \varphi_i - f(\cdot, t_n)\varphi_i \right) dx \, .$$

Obviously, any solution $u^n \in V_m$ corresponds to a zero $\boldsymbol{u}^n$ of $\boldsymbol{h}$ and vice versa.

Due to the continuity of the nonlinearity a, the function $\boldsymbol{h} : \mathbb{R}^m \to \mathbb{R}^m$ is continuous. Moreover, we have with the simple but crucial relation (which reflects the stability of the backward Euler method)

$$(a - b) \cdot a = \frac{1}{2} \left(a^2 - b^2 + (a - b)^2 \right) , \qquad a, b \in \mathbb{R} \, , \tag{3.3}$$

the coercivity assumption (1.2), and (2.6) (taking $\varepsilon = \mu/2$)

$$\begin{aligned} \boldsymbol{h}(\boldsymbol{v}) \cdot \boldsymbol{v} &= \int_\Omega \left(\frac{v - u^{n-1}}{\tau} v + a(\nabla v) \cdot \nabla v - f(\cdot, t_n) v \right) dx \\ &\geq \frac{1}{2\tau} \left(\|v\|_{2,\Omega}^2 - \|u^{n-1}\|_{2,\Omega}^2 \right) + \frac{\mu}{2} \int_\Omega M(\nabla v) dx + \mu \int_\Omega M^*(a(\nabla v)) dx - \frac{1}{2} \|v\|_{2,\Omega}^2 - C(f(\cdot, t_n)) \, , \end{aligned}$$

where

$$C(f(\cdot, t_n)) := C\, |\Omega| + \frac{1}{2|\Omega|} \|f(\cdot, t_n)\|_{1,\Omega}^2 \tag{3.4a}$$

with $C > 0$ given by Lemma 2.4. It is clear that

$$\begin{aligned} &\max_{n=1,2,\dots,N} C(f(\cdot, t_n)) \\ &\leq |\Omega| \sup_{\eta \in \mathbb{R}^d, |\eta| \leq 2c\mu^{-1} \|f\|_{L^\infty(0,T;L^q(\Omega))}} M^*(\eta) + \frac{1}{2|\Omega|} \|f\|_{L^\infty(0,T;L^1(\Omega))}^2 =: C(f) \, , \end{aligned} \tag{3.4b}$$

where $c > 0$ only depends on Ω, d, and q.

Taking now R such that $R^2 > \left(\|u^{n-1}\|_{2,\Omega}^2 + 2\tau\, C(f(\cdot, t_n)) \right) / (1 - \tau)$, the assumptions of Lemma 3.3 are fulfilled, and there exists a zero of $\boldsymbol{h}$. This zero, however, solves (3.2) at level n. ◻

Remark 3.4. *It is straightforward to show uniqueness of the discrete solution if the nonlinearity is strictly monotone.*

3.3. A priori estimates. The following a priori estimates are the essential prerequisite for the proof of convergence.

Theorem 3.5 (Uniform boundedness of discrete solution). *Let $u^0 \in V_m$ and $f \in \mathscr{C}([0, T]; L^q(\Omega))$ for some $q > d$. Let $\{u^n\} \subset V_m$ be any solution to (3.2). Let $\tau \leq \tau_0 < 1$. Then there holds for all $n = 1, 2, \dots, N$*

$$\begin{aligned} \|u^n\|_{2,\Omega}^2 + \sum_{j=1}^{n} \|u^j - u^{j-1}\|_{2,\Omega}^2 + \tau \sum_{j=1}^{n} \int_\Omega M(\nabla u^j)\, dx + \tau \sum_{j=1}^{n} \int_\Omega M^*(a(\nabla u^j))\, dx \\ \leq c \left(\|u^0\|_{2,\Omega}^2 + C(f) \right) , \end{aligned} \tag{3.5}$$

where $c > 0$ depends on μ, T and τ_0, and $C(f)$ is given by (3.4).

Proof. We take $v = u^n$ in (3.2), employ the relation (3.3) for the discrete time derivative, invoke the coercivity assumption (1.2), and use (2.6) with $\varepsilon = \mu/2$. This leads to

$$\frac{1}{2\tau}\left(\|u^n\|_{2,\Omega}^2 - \|u^{n-1}\|_{2,\Omega}^2 + \|u^n - u^{n-1}\|_{2,\Omega}^2\right) \\ + \frac{\mu}{2}\int_\Omega M(\nabla u^n)\,dx + \mu\int_\Omega M^*(a(\nabla u^n))\,dx \le \frac{1}{2}\|u^n\|_{2,\Omega}^2 + C(f(\cdot,t_n))\,,$$

where $C(f(\cdot,t_n))$ is given by (3.4). Summation then implies for all $n = 1,2,\ldots,N$

$$\|u^n\|_{2,\Omega}^2 + \tau\sum_{j=1}^n \|u^j - u^{j-1}\|_{2,\Omega}^2 + \mu\tau\sum_{j=1}^n\int_\Omega M(\nabla u^j)\,dx + 2\mu\tau\sum_{j=1}^n\int_\Omega M^*(a(\nabla u^j))\,dx \\ \le \|u^0\|_{2,\Omega}^2 + \tau\sum_{j=1}^n\|u^j\|_{2,\Omega}^2 + 2\tau\sum_{j=1}^n C(f(\cdot,t_j))\,.$$

With (3.4), we have $2\tau\sum_{j=1}^n C(f(\cdot,t_j)) \le 2TC(f)$. Applying a discrete Gronwall lemma now proves the assertion. ∎

If the approximation of the initial datum is taken from a bounded set, the theorem above shows indeed uniform boundedness of the discrete solution. The application of a discrete Gronwall lemma cannot be avoided but is not too problematic here from the numerical point of view since it results in a constant that behaves like $\exp(T/(1-\tau_0))$.

4. Convergence of the numerical solution. In what follows, we consider a sequence $\{(m_\ell, N_\ell)\}_{\ell\in\mathbb{N}}$ such that $m_\ell \to \infty$ as well as $N_\ell \to \infty$ as $\ell \to \infty$. Moreover, we suppose that $\tau_\ell \le \tau_0 < 1$ for all $\ell \in \mathbb{N}$. (When writing t_n or u^n, we omit calling the dependence on ℓ if no confusion is likely to arise.)

Furthermore, we consider a sequence $\{u_\ell^0\}_{\ell\in\mathbb{N}}$ of approximations of the initial datum $u_0 \in L^2(\Omega)$ such that $u_\ell^0 \in V_{m_\ell}$ and

$$u_\ell^0 \to u_0 \quad \text{in } L^2(\Omega) \text{ as } \ell\to\infty\,. \tag{4.1}$$

From a fully discrete solution $\{u^n\}$ corresponding to the space V_{m_ℓ} and the time grid with step size $\tau_\ell = T/N_\ell$, we now construct numerical approximations that are defined on the whole time interval: Let u_ℓ be the piecewise constant function with

$$u_\ell(\cdot,t) = u^n \quad \text{if } t\in(t_{n-1},t_n]\ (n = 1,2,\ldots,N_\ell)\,,\ u_\ell(\cdot,0) = u^1\,.$$

Moreover, let $\hat{u}_\ell$ denote the linear spline interpolating $(t_0,u^0),(t_1,u^1),\ldots,(t_{N_\ell},u^{N_\ell})$.

We also use the piecewise constant in time approximation f_ℓ defined by

$$f_\ell(\cdot,t) = f(\cdot,t_n) \quad \text{if } t\in(t_{n-1},t_n]\ (n = 1,2,\ldots,N_\ell)\,,\ f_\ell(\cdot,0) = f(\cdot,t_1)\,.$$

It is clear that if $f\in\mathscr{C}([0,T];L^q(\Omega))$ then

$$\|f - f_\ell\|_{L^\infty(0,T;L^q(\Omega))} \to 0 \quad \text{as } \ell\to\infty\,. \tag{4.2}$$

The main result of the paper now reads as follows:

THEOREM 4.1 (Convergence of approximate solution). *Let $u_0 \in L^2(\Omega)$ and $f \in \mathscr{C}([0,T];L^q(\Omega))$ with $q > d$ be given. Consider the numerical solution of (1.1)* *by the scheme (3.2) on a sequence of finite dimensional subspaces, such that (3.1) is satisfied, and time step sizes, which tend to zero and are bounded away from 1. For the approximation of the initial datum, assume (4.1).*

Then there is a subsequence, denoted by ℓ', such that the sequences $\{u_{\ell'}\}$ and $\{\hat{u}_{\ell'}\}$ of piecewise constant in time and piecewise linear in time prolongations, respectively, of the numerical solutions converge weakly in $L^\infty(0,T;L^2(\Omega))$ towards an exact solution $u \in \mathscr{C}_w([0,T];L^2(\Omega))$ to (1.1). Moreover, $u_{\ell'}(\cdot,T) = \hat{u}_{\ell'}(\cdot,T)$ converges weakly in $L^2(\Omega)$ towards $u(\cdot,T)$, $\nabla u_{\ell'}$ converges weakly* in $L_M(Q;\mathbb{R}^d)$ towards $\nabla u \in \mathscr{L}_M(Q;\mathbb{R}^d)$ and $a(\nabla u_{\ell'})$ converges weakly* in $L_{M^*}(Q;\mathbb{R}^d)$ towards $a(\nabla u) \in \mathscr{L}_{M^*}(Q;\mathbb{R}^d)$.*

We remark that, without assuming higher regularity of the exact solution (which is, in general, not known) no better convergence can be expected.

The proof will be prepared by the following lemma.

Lemma 4.2. *Under the assumptions of Theorem 4.1, there is a subsequence, denoted by ℓ', and elements $u \in L^\infty(0,T;L^2(\Omega))$ with $\nabla u \in \mathscr{L}_M(Q;\mathbb{R}^d)$ and $\gamma_0 u(\cdot,t) = 0$ for almost all $t \in (0,T)$, $z \in L^2(\Omega)$, $\alpha \in \mathscr{L}_{M^*}(Q;\mathbb{R}^d)$ such that, as $\ell \to \infty$,*

$$\begin{aligned}
&u_\ell - \hat{u}_\ell \to 0 \text{ in } L^2(Q)\,, \quad u_{\ell'},\, \hat{u}_{\ell'} \overset{*}{\rightharpoonup} u \text{ in } L^\infty(0,T;L^2(\Omega))\,,\\
&\hat{u}_{\ell'}(\cdot,T) = u_{\ell'}(\cdot,T) \rightharpoonup z \text{ in } L^2(\Omega)\,,\\
&\nabla u_{\ell'} \overset{*}{\rightharpoonup} \nabla u \text{ in } L_M(Q;\mathbb{R}^d)\,, \quad a(\nabla u_{\ell'}) \overset{*}{\rightharpoonup} \alpha \text{ in } L_{M^*}(Q;\mathbb{R}^d)\,.
\end{aligned}$$

Proof. Because of (4.1), the sequence $\{u_\ell^0\}$ is bounded in $L^2(\Omega)$. Therefore, the right-hand side of the a priori estimate in Theorem 3.5 is also bounded.

A simple calculation (employing the definition of u_ℓ and $\hat{u}_\ell$) shows that

$$\|u_\ell - \hat{u}_\ell\|_{2,Q}^2 = \frac{\tau}{3} \sum_{n=1}^{N_\ell} \|u^n - u^{n-1}\|_{2,\Omega}^2\,,$$

and, in view of Theorem 3.5, the right-hand side tends to zero as $\ell \to \infty$.

An immediate consequence of the definition of the approximate solutions is

$$\|u_\ell\|_{L^\infty(0,T;L^2(\Omega))} = \max_{n=1,2,\ldots,N_\ell} \|u^n\|_{2,\Omega}\,, \quad \|\hat{u}_\ell\|_{L^\infty(0,T;L^2(\Omega))} = \max_{n=0,1,\ldots,N_\ell} \|u^n\|_{2,\Omega}\,,$$

and Theorem 3.5 shows the boundedness of $\{u_\ell\}$ and $\{\hat{u}_\ell\}$ in $L^\infty(0,T;L^2(\Omega))$, which is the dual of the separable Banach space $L^1(0,T;L^2(\Omega))$. We thus have weak* convergence of a subsequence in $L^\infty(0,T;L^2(\Omega))$. The limits of both the sequences must coincide since their difference tends to zero in $L^2(Q)$.

Since $\|\hat{u}_\ell(\cdot,T)\|_{2,\Omega} = \|u_\ell(\cdot,T)\|_{2,\Omega} = \|u^{N_\ell}\|_{2,\Omega}$, the a priori estimate in Theorem 3.5 also proves the asserted weak convergence of a subsequence of $\{\hat{u}_\ell(\cdot,T)\} = \{u_\ell(\cdot,T)\}$ in $L^2(\Omega)$.

With respect to the sequence of gradients of u_ℓ, we observe that

$$\int_Q M(\nabla u_\ell)dxdt = \tau \sum_{n=1}^{N_\ell} \int_\Omega M(\nabla u^n)dx$$

is uniformly bounded, see again Theorem 3.5. From the boundedness of the modular, however, boundedness of the Luxemburg norm follows. Therefore, $\{\nabla u_\ell\} \subset$

$\mathscr{L}_M(Q;\mathbb{R}^d) \subseteq L_M(Q;\mathbb{R}^d)$ is bounded with respect to $\|\cdot\|_{M,Q}$. Since $L_M(Q;\mathbb{R}^d)$ is the dual of the separable Banach space $E_{M^*}(Q;\mathbb{R}^d)$, we obtain weak* convergence of a subsequence in $L_M(Q;\mathbb{R}^d)$ towards an element $\xi \in L_M(Q;\mathbb{R}^d)$ such that $\nabla u_{\ell'} \overset{*}{\rightharpoonup} \xi$.

It remains to show $\xi = \nabla u$. However, since $\mathscr{C}_c^\infty(\Omega;\mathbb{R}^d)\otimes\mathscr{C}_c^\infty(0,T) \subset E_{M^*}(Q;\mathbb{R}^d)$, we find for all $\Phi \in \mathscr{C}_c^\infty(\Omega;\mathbb{R}^d)$ and all $\psi \in \mathscr{C}_c^\infty(0,T)$ with integration by parts

$$\begin{aligned}\int_Q \xi\cdot\Phi\,\psi dxdt &= \lim_{\ell'\to\infty}\int_Q \nabla u_{\ell'}\cdot\Phi\,\psi dxdt\\ &= -\lim_{\ell'\to\infty}\int_Q u_{\ell'}\,\nabla\cdot\Phi\,\psi dxdt = -\int_Q u\,\nabla\cdot\Phi\,\psi dxdt\,.\end{aligned}$$

In the last step, we have used that $u_{\ell'}$ converges weakly* in $L^\infty(0,T;L^2(\Omega))$ towards u.

In view of Lemma 2.2, we finally get $\nabla u \in \mathscr{L}_M(Q;\mathbb{R}^d)$.

Since the trace operator $\gamma_0 : W^{1,1}(\Omega) \to L^1(\partial\Omega)$ is linear and bounded, it can be extended to a linear bounded (and thus weakly-weakly continuous) operator mapping $L^1(0,T;W^{1,1}(\Omega))$ into $L^1(0,T;L^1(\partial\Omega))$. By employing the Fenchel–Young inequality, it is easy to show that the uniform boundedness of the modular of ∇u_ℓ implies the uniform boundedness of the $L^1(0,T;L^1(\Omega;\mathbb{R}^d))$-norm of ∇u_ℓ. Therefore, the subsequence can be chosen such that $u_{\ell'}$ converges also weakly in $L^1(0,T;W^{1,1}(\Omega))$ towards u. Since $u_{\ell'}$ has vanishing trace for almost all $t \in (0,T)$, also the weak limit u must have vanishing trace for almost all $t \in (0,T)$.

A similar argumentation as for $\{\nabla u_\ell\}$ proves the remaining assertion for $\{a(\nabla u_\ell)\}$ since

$$\int_Q M^*(a(\nabla u_\ell))dxdt = \tau\sum_{n=1}^{N_\ell}\int_\Omega M^*(a(\nabla u^n))dx$$

is uniformly bounded in view of Theorem 3.5. We infer that there exists $\alpha \in L_{M^*}(Q;\mathbb{R}^d)$ such that $a(\nabla u_{\ell'}) \overset{*}{\rightharpoonup} \alpha$ in $L_{M^*}(Q;\mathbb{R}^d)$ for a subsequence. Because of Lemma 2.2, α belongs to the Orlicz class $\mathscr{L}_{M^*}(Q;\mathbb{R}^d)$. ☐

We are now ready to prove the main result.

Proof. [of Theorem 4.1] We omit writing ℓ' for the subsequence from Lemma 4.2. Using the approximations $\hat{u}_\ell$ and u_ℓ, the numerical scheme (3.2) can be written as

$$\int_\Omega (\partial_t\hat{u}_\ell v + a(\nabla u_\ell)\cdot\nabla v)\,dx = \int_\Omega f_\ell v dx \qquad \text{for all } v \in V_{m_\ell}\,. \tag{4.3}$$

With respect to time, this equation holds almost everywhere in $(0,T)$ as well as in the weak sense. This immediately implies

$$\begin{aligned}&-\int_Q \hat{u}_\ell R_{m_\ell} v\psi' dxdt + \int_\Omega \hat{u}_\ell(\cdot,T)R_{m_\ell} v dx\,\psi(T) - \int_\Omega \hat{u}_\ell(\cdot,0)R_{m_\ell} v dx\,\psi(0)\\ &+\int_Q a(\nabla u_\ell)\cdot\nabla R_{m_\ell} v\psi dxdt = \int_Q f_\ell R_{m_\ell} v\psi dxdt \quad \text{for all } v\in V\,,\ \psi\in\mathscr{C}^1([0,T])\,.\end{aligned}$$

Note that $\hat{u}_\ell(\cdot,T) = u^{N_\ell}$ and $\hat{u}_\ell(\cdot,0) = u_\ell^0$.

With Lemma 4.2, relation (3.1), (4.1) and (4.2), we obtain in the limit

$$\begin{aligned}-\int_Q uv\psi' dxdt + \int_\Omega zv dx\,\psi(T) - \int_\Omega u_0 v dx\,\psi(0) + \int_Q \alpha\cdot\nabla v\psi dxdt = \int_Q fv\psi dxdt\\ \text{for all } v\in V\,,\ \psi\in\mathscr{C}^1([0,T])\,.\end{aligned} \tag{4.4}$$

In particular, we have employed that (with $1/q + 1/q' = 1$), as $\ell \to \infty$,

$$R_{m_\ell} v\psi' \to v\psi' \quad \text{in } L^1(0,T;L^2(\Omega))\,, \quad R_{m_\ell} v \to v \quad \text{in } L^2(\Omega)\,,$$
$$\nabla R_{m_\ell} v\psi \to \nabla v\psi \quad \text{in } E_M(Q;\mathbb{R}^d)\,, \quad R_{m_\ell} v\psi \to v\psi \quad \text{in } L^1(0,T;L^{q'}(\Omega))\,.$$

This follows from (3.1) and the definition of the norm in V. Moreover, we observe that $V \hookrightarrow W^{1,1}(\Omega) \cap L^2(\Omega)$, and for $d = 1$, we obtain $V \hookrightarrow L^{q'}(\Omega)$ for any $q > d = 1$.

Relation (4.4) implies, by density arguments,

$$-\int_Q u\partial_t w dxdt + \int_\Omega zw(\cdot,T)dx - \int_\Omega u_0 w(\cdot,0)dx + \int_Q \alpha\cdot\nabla w dxdt = \int_Q fwdxdt$$
$$\text{for all } w \in \mathscr{W}\,, \tag{4.5}$$

where $\mathscr{W}$ was defined in Lemma 2.3. This is a crucial step. We first observe that the tensor product $V \otimes \mathscr{C}^1([0,T])$ is included in $\mathscr{W}$, which shows that (4.4) is a particular case of (4.5). The function w_ε that exists in view of Lemma 2.3 can be approximated, with respect to the strong convergence in $\mathscr{C}^1(\overline{Q})$, by a polynomial vanishing at $\partial\Omega \times [0,T]$, which possesses a tensor structure and thus belongs to $V \otimes \mathscr{C}^1([0,T])$. For any $u \in L^\infty(0,T;L^2(\Omega))$, $z, u_0 \in L^2(\Omega)$, $\alpha \in \mathscr{L}_{M^*}(Q;\mathbb{R}^d)$, $f \in \mathscr{C}([0,T];L^q(\Omega))$, any $\varepsilon > 0$ and any $w \in \mathscr{W}$, there is hence (recalling also the continuous embedding of $W^{1,1}(0,T;L^2(\Omega))$ into $\mathscr{C}([0,T];L^2(\Omega))$) an element $w_\varepsilon \in V \otimes \mathscr{C}^1([0,T])$ such that

$$\left|\int_Q u\partial_t(w_\varepsilon - w)dxdt\right| + \left|\int_\Omega z(w_\varepsilon(\cdot,T) - w(\cdot,T))dx\right|$$
$$+\left|\int_\Omega u_0(w_\varepsilon(\cdot,0) - w(\cdot,0))dx\right| + \left|\int_Q \alpha\cdot\nabla(w_\varepsilon - w)dxdt\right| + \left|\int_Q f(w_\varepsilon - w)dxdt\right| < \varepsilon\,.$$

For the last term, we have to apply (2.7) upon noting that for $f \in \mathscr{C}([0,T];L^q(\Omega))$ there is a function $\tilde{f} \in \mathscr{C}([0,T];W_0^{1,q}(\Omega;\mathbb{R}^d)) \hookrightarrow \mathscr{C}([0,T];L^\infty(\Omega;\mathbb{R}^d)) \subset L_{M^*}(Q;\mathbb{R}^d)$.

We are now going to derive further properties of the limit u.

Recalling that $u \in L^\infty(0,T;L^2(\Omega))$ with $\nabla u \in \mathscr{L}_M(Q;\mathbb{R}^d) \subset L^1(0,T;L_M(\Omega;\mathbb{R}^d))$, $\alpha \in \mathscr{L}_{M^*}(Q;\mathbb{R}^d) \subset L^1(0,T;L_{M^*}(\Omega;\mathbb{R}^d))$, and $f \in \mathscr{C}([0,T];L^q(\Omega))$ (with $q > d$), we see that for any $v \in V$ the functions

$$t \mapsto \int_\Omega u(\cdot,t)v\,dx\,,\ t \mapsto \int_\Omega \alpha(\cdot,t)\cdot\nabla v\,dx\,,\ t \mapsto \int_\Omega f(\cdot,t)v\,dx$$

are at least in $L^1(0,T)$. This observation, together with (4.4), shows that

$$\frac{d}{dt}\int_\Omega u(\cdot,t)v\,dx = \int_\Omega (f(\cdot,t)v - \alpha(\cdot,t)\cdot\nabla v)\,dx \tag{4.6}$$

holds true in the weak sense. Moreover, the function $t \mapsto \int_\Omega u(\cdot,t)v\,dx$ then is absolutely continuous. Hence, since V is dense in $L^2(\Omega)$ with respect to the strong convergence in $L^2(\Omega)$ and since $u \in L^\infty(0,T;L^2(\Omega))$, there holds $u \in \mathscr{C}_w([0,T];L^2(\Omega))$.

We can now prove $u(\cdot,0) = u_0 \in L^2(\Omega)$. For arbitrary $v \in V$, we have with (4.3)

$$\int_\Omega u_\ell^0 R_{m_\ell} v\,dx = \left[\int_\Omega \hat{u}_\ell(\cdot,t)R_{m_\ell} v\,dx\,\frac{t-T}{T}\right]_{t=0}^T$$
$$= \int_0^T \left(\int_\Omega \partial_t\hat{u}_\ell R_{m_\ell} v\,dx\,\frac{t-T}{T} + \int_\Omega \hat{u}_\ell R_{m_\ell} v\,dx\,\frac{1}{T}\right)dt$$
$$= \int_0^T \left(\int_\Omega (f_\ell R_{m_\ell} v - a(\nabla u_\ell)\cdot\nabla R_{m_\ell} v)\,dx\,\frac{t-T}{T} + \int_\Omega \hat{u}_\ell R_{m_\ell} v\,dx\,\frac{1}{T}\right)dt\,.$$

In the limit (see Lemma 4.2), we thus obtain with integration by parts (using (4.6))

$$\begin{aligned}\int_\Omega u_0 v\,dx &= \int_0^T \left(\int_\Omega (fv - \alpha\cdot\nabla v)\,dx\,\frac{t-T}{T} + \int_\Omega uv\,dx\,\frac{1}{T}\right) dt \\ &= \left[\int_\Omega uv\,dx \frac{t-T}{T}\right]_{t=0}^T = \int_\Omega u(\cdot,0)v\,dx\,.\end{aligned}$$

Using the function $t \mapsto t/T$ instead of $t \mapsto (t-T)/T$, the same argumentation as above provides that the weak in $L^2(\Omega)$ limit z of $\hat{u}_\ell(\cdot,T) = u_\ell(\cdot,T)$ is indeed $u(\cdot,T)$,

$$\hat{u}_\ell(\cdot,T) \rightharpoonup z = u(\cdot,T) \in L^2(\Omega) \text{ as } \ell\to\infty\,.$$

It remains to identify α, i.e., to show that $\alpha = a(\nabla u)$. For proving this, we employ a variant of Minty's monotonicity trick. Unfortunately, a direct application of Minty's trick is not possible since we are working in spaces which are not reflexive and so we cannot just take the limit u as a test function in the limit equation (4.4).

Using (3.3), we find

$$\begin{aligned}\int_Q \partial_t \hat{u}_\ell\, u_\ell dxdt &= \sum_{n=1}^{N_\ell} \int_\Omega (u^n - u^{n-1})\,u^n dx \ge \frac{1}{2}\left(\|u^{N_\ell}\|_{2,\Omega}^2 - \|u_\ell^0\|_{2,\Omega}^2\right) \\ &= \frac{1}{2}\left(\|u_\ell(\cdot,T)\|_{2,\Omega}^2 - \|u_\ell^0\|_{2,\Omega}^2\right),\end{aligned}$$

which implies, because of the weak lower semicontinuity of the norm, the weak convergence of $u_\ell(\cdot,T)$ towards $z = u(T)$ in $L^2(\Omega)$, and the strong convergence (4.1),

$$\frac{1}{2}\left(\|u(\cdot,T)\|_{2,\Omega}^2 - \|u_0\|_{2,\Omega}^2\right) \le \liminf_{\ell\to\infty} \int_Q \partial_t \hat{u}_\ell\, u_\ell dxdt\,. \tag{4.7}$$

On the other hand, since a is monotone, we know that for all $\eta \in L^\infty(Q;\mathbb{R}^d)$

$$\begin{aligned}\int_Q a(\nabla u_\ell)\cdot\nabla u_\ell\,dxdt &\ge \int_Q a(\nabla u_\ell)\cdot\nabla u_\ell\,dxdt - \int_Q (a(\nabla u_\ell) - a(\eta))\cdot(\nabla u_\ell - \eta)\,dxdt \\ &= \int_Q a(\nabla u_\ell)\cdot\eta\,dxdt + \int_Q a(\eta)\cdot(\nabla u_\ell - \eta)\,dxdt\,.\end{aligned}$$

Note that $a(\eta) \in E_{M^*}(Q;\mathbb{R}^d)$ since $\eta \in L^\infty(Q;\mathbb{R}^d)$ and a is continuous. In the limit, we thus obtain (see again Lemma 4.2)

$$\int_Q \alpha\cdot\eta\,dxdt + \int_Q a(\eta)\cdot(\nabla u - \eta)\,dxdt \le \liminf_{\ell\to\infty}\int_Q a(\nabla u_\ell)\cdot\nabla u_\ell\,dxdt\,. \tag{4.8}$$

Finally, we know that

$$\int_Q f_\ell u_\ell\,dxdt \to \int_Q fu\,dxdt \text{ as } \ell\to\infty\,.$$

Taking $v = u_\ell(\cdot,t) \in V_{m_\ell}$ in (4.3), using (4.7) and (4.8), we thus come up with

$$\begin{aligned}&\frac{1}{2}\left(\|u(\cdot,T)\|_{2,\Omega}^2 - \|u_0\|_{2,\Omega}^2\right) + \int_Q \alpha\cdot\eta\,dxdt \\ &+ \int_Q a(\eta)\cdot(\nabla u - \eta)\,dxdt \le \int_Q fu\,dxdt\,.\end{aligned} \tag{4.9}$$

Unfortunately, we cannot take $w = u$ in (4.5) due to the lack of regularity in time. We, therefore, consider the centered Steklov average of u, given by

$$(S_h u)(\cdot, t) = \frac{1}{2h} \int_{t-h}^{t+h} u(\cdot, s)\, ds\,, \quad t \in [0, T]\,,$$

where $h > 0$ and where u is extended by zero outside $[0, T]$. The properties of u imply that $S_h u \in \mathscr{W}$. It is known that

$$\lim_{h \to 0} \int_Q f S_h u \, dxdt = \int_Q f u \, dxdt\,.$$

On the other hand, we find with (4.5)

$$\begin{aligned} \int_Q f S_h u \, dxdt = &- \int_Q u \partial_t S_h u \, dxdt + \int_\Omega u(\cdot, T) S_h u(\cdot, T)\, dx \\ &- \int_\Omega u(\cdot, 0) S_h u(\cdot, 0)\, dx + \int_Q \alpha \cdot \nabla S_h u \, dxdt\,, \end{aligned}$$

where $(\partial_t S_h u)(\cdot, t) = (u(\cdot, t+h) - u(\cdot, t-h))\,/(2h)$ and thus

$$\begin{aligned} &\int_Q u \partial_t S_h u \, dxdt = \frac{1}{2h} \int_0^T \int_\Omega u(\cdot, t)\,(u(\cdot, t+h) - u(\cdot, t-h))\, dxdt \\ &= \frac{1}{2h} \int_0^{T-h} \int_\Omega u(\cdot, t) u(\cdot, t+h)\, dxdt - \frac{1}{2h} \int_h^T \int_\Omega u(\cdot, t) u(\cdot, t-h)\, dxdt = 0\,. \end{aligned} \tag{4.10}$$

Moreover, we have

$$\begin{aligned} \int_\Omega u(\cdot, T) S_h u(\cdot, T) dx &= \frac{1}{2h} \int_{T-h}^T \int_\Omega u(\cdot, T) u(\cdot, s)\, dxds \\ &\to \frac{1}{2} \int_\Omega u(\cdot, T)^2\, dx = \frac{1}{2} \left\| u(\cdot, T) \right\|_{2,\Omega}^2 \text{ as } h \to 0\,. \end{aligned} \tag{4.11}$$

Recall here that $u \in \mathscr{C}_w([0, T]; L^2(\Omega))$ and thus $s = T$ is a Lebesgue point of the mapping $s \mapsto \int_\Omega u(\cdot, T) u(\cdot, s)\, dx$. Analogously, we have

$$\int_\Omega u(\cdot, 0) S_h u(\cdot, 0)\, dx \to \frac{1}{2} \left\| u_0 \right\|_{2,\Omega}^2 \text{ as } h \to 0\,.$$

Finally, we observe that

$$\begin{aligned} &\int_Q \alpha \cdot \nabla S_h u \, dxdt - \int_Q \alpha \cdot \nabla u \, dxdt \\ &= \frac{1}{2h} \int_0^T \int_{t-h}^{t+h} \int_\Omega \alpha(\cdot, t) \cdot \nabla\, (u(\cdot, s) - u(\cdot, t))\, dxdsdt \\ &= \frac{1}{2} \int_{-1}^1 \int_0^T \int_\Omega \alpha(\cdot, t) \cdot (\nabla u(\cdot, t + rh) - \nabla u(\cdot, t))\, dxdtdr \to 0 \text{ as } h \to 0 \end{aligned} \tag{4.12}$$

since the translation of a function in the Orlicz space $L_M(Q; \mathbb{R}^d)$ is continuous with respect to the weak convergence in $E_M(Q; \mathbb{R}^d)$ (see [15, Lemma 1.5] and [9, Prop. 1.2]).

Altogether, we infer from (4.9) that for all $\eta \in L^\infty(Q;\mathbb{R}^d)$

$$0 \le \int_Q (a(\eta) - \alpha) \cdot (\eta - \nabla u)\, dxdt\,.$$

Following the modification of Minty's trick in [18] (see also [21]), we set $Q_k = \{(x,t) \in Q : |\nabla u(x,t)| > k\}$ for any $k \in \mathbb{N}$. For arbitrary $i, j \in \mathbb{N}$ with $j < i$, arbitrary $\lambda > 0$, and arbitrary $\zeta \in L^\infty(Q;\mathbb{R}^d)$, we take

$$\eta = (\nabla u)\mathbb{1}_{Q\setminus Q_i} + \lambda\zeta\mathbb{1}_{Q\setminus Q_j} = \begin{cases} 0 & \text{in } Q_i\,, \\ \nabla u & \text{in } Q_j \setminus Q_i\,, \\ \nabla u + \lambda\zeta & \text{in } Q \setminus Q_j\,. \end{cases}$$

This shows that

$$0 \le -\int_{Q_i} (a(0) - \alpha) \cdot \nabla u\, dxdt + \lambda \int_{Q\setminus Q_j} (a(\nabla u + \lambda\zeta) - \alpha) \cdot \zeta\, dxdt\,.$$

As in (2.3), we see that the first term on the right-hand side tends to zero as $i \to \infty$ since $(a(0) - \alpha) \cdot \nabla u \in L^1(Q)$. We, therefore, come up with

$$0 \le \int_{Q\setminus Q_j} (a(\nabla u + \lambda\zeta) - \alpha) \cdot \zeta\, dxdt\,.$$

Recalling that a is continuous, we immediately find that

$$\int_{Q\setminus Q_j} (a(\nabla u + \lambda\zeta) - \alpha) \cdot \zeta\, dxdt \to \int_{Q\setminus Q_j} (a(\nabla u) - \alpha) \cdot \zeta\, dxdt \text{ as } \lambda \to 0$$

and thus

$$0 \le \int_{Q\setminus Q_j} (a(\nabla u) - \alpha) \cdot \zeta\, dxdt$$

for any $j \in \mathbb{N}$ and any $\zeta \in L^\infty(Q;\mathbb{R}^d)$. The choice $\zeta = -\frac{a(\nabla u) - \alpha}{|a(\nabla u) - \alpha|}$ if $a(\nabla u) \neq \alpha$ and $\zeta = 0$ otherwise provides

$$\int_{Q\setminus Q_j} |a(\nabla u) - \alpha|\, dxdt \le 0$$

and thus $\alpha = a(\nabla u)$ almost everywhere in $Q \setminus Q_j$. Since j was arbitrary, this proves $\alpha = a(\nabla u)$ almost everywhere in Q, which finishes the proof. ☐

REMARK 4.3. *If the exact solution is unique, which is the case if the nonlinearity a is strictly monotone, then the whole sequences of approximate solutions converge.*

Uniqueness in case of a strictly monotone nonlinearity is seen as follows: Let u and v be two different solutions to the problem with the same data (u_0, f). From the proof above, we already know that then for all $w \in \mathscr{W}$

$$-\int_Q (u - v)\partial_t w dxdt + \int_\Omega (u(\cdot,T) - v(\cdot,T))w(\cdot,T)dx + \int_Q (a(\nabla u) - a(\nabla v)) \cdot \nabla w dxdt = 0\,.$$

With $w = S_h(u - v)$ and observations analogous to (4.10), (4.11), and (4.12), we find

$$\frac{1}{2}\, \|u(\cdot,T) - v(\cdot,T)\|_{2,\Omega}^2 + \int_Q (a(\nabla u) - a(\nabla v)) \cdot (\nabla u - \nabla v) dxdt = 0$$

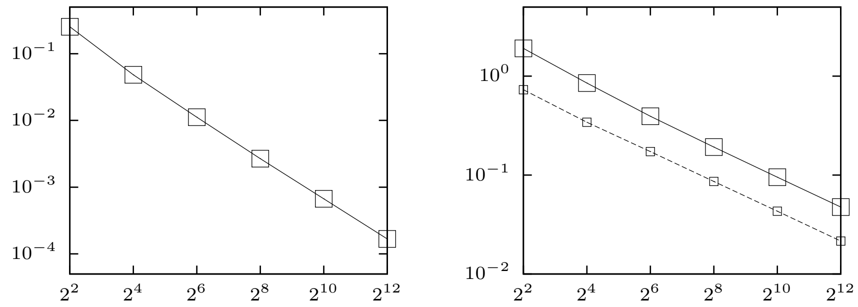

FIG. 5.1. *Error between exact and numerical solution*

TABLE 5.1
Error between exact and numerical solution

mesh size h	time step τ	$\|u-u_\ell\|_{L^\infty(L^2)}$	$\|\nabla(u-u_\ell)\|_{L^1(Q)}$	$\sqrt{\rho(\nabla(u-u_\ell))}$
1.00E+00	1.00E-01	2.53E-01	1.91E+00	7.30E-01
5.00E-01	2.50E-02	4.83E-02	8.53E-01	3.42E-01
2.50E-01	6.25E-03	1.12E-02	3.93E-01	1.73E-01
1.25E-01	1.56E-03	2.67E-03	1.92E-01	8.62E-02
6.25E-02	3.91E-04	6.70E-04	9.52E-02	4.31E-02
3.13E-02	9.77E-05	1.68E-04	4.75E-02	2.15E-02

as $h \to 0$. On the other hand, the strict monotonicity of a shows that

$$\int_Q (a(\nabla u) - a(\nabla v)) \cdot (\nabla u - \nabla v) dxdt = 0$$

if and only if $\nabla u = \nabla v$ almost everywhere. Recalling here that $\gamma_0 u = \gamma_0 v = 0$, this is in contradiction to $u \neq v$.

5. Numerical illustration. We consider example 7) from page 2 on $Q = (-1,1)^2 \times (0,1)$ with $u_0(x,y) = \mathrm{e}^{-1}\sin(\pi x)\sin(\pi y)$ and f such that the exact solution is given by $u(x,y,t) = \mathrm{e}^{-t-1}u_0(x,y)$.

For the spatial discretization, we employ standard $\mathscr{Q}^1$ finite elements (here indeed uniform squares), which fit into our framework because of Example 3.1. The arising nonlinear system of equations in each time step is solved by a Newton iteration where we use the exact Jacobian. The time step size is taken proportional to the square of the spatial mesh size. The computations have been carried out using `dealII` (see [3]).

In Figure 5.1 (left), the error between the exact solution u and the numerical solution u_ℓ is shown in the norm of $L^\infty(0,T;L^2(\Omega))$. Moreover in Figure 5.1 (right), the difference in the gradient of the exact and the numerical solution is shown in the norm of $L^1(Q)$ (big boxes) and in $\sqrt{\rho_{M,Q}}$ (small boxes). (We consider the square root of the modular since $M(\xi) \sim \frac{1}{2}|\xi|^2$ for $|\xi| \to 0$.) On the x-axis, we have the number of finite elements. Table 5.1 shows the corresponding numbers.

It turns out that, for the smooth solution we consider here, the convergence is even better than provided by the theoretical result.

Acknowledgements. The authors gratefully acknowledge fruitful discussions with Filip Rindler and help with the numerical computations by David Šiška.

REFERENCES

[1] R. A. Adams, *Sobolev spaces*, Academic Press, New York, 1975.

[2] H. Attouch, G. Buttazzo and G. Michaille, *Variational analysis in Sobolev and BV spaces: applications to PDEs and optimization*, SIAM, Philadelphia, 2006.

[3] W. Bangerth, R. Hartmann, and G. Kanschat, *deal.II – a general purpose object oriented finite element library*, ACM Trans. Math. Softw., 33 (2007) 4, pp. 24/1–24/27.

[4] K. Barg, *Asymptotische Fehlerabschätzungen für das Galerkin-Verfahren zur numerischen Behandlung von nichtlinearen Randwertaufgaben in Sobolev- Orlicz-Räumen*, Hannover Univ., Diss., 1989.

[5] S.C. Brenner and L.R. Scott, *The mathematical theory of finite element methods*, Springer, Berlin, 2002.

[6] Y. Cai and S. Zhou *Existence and uniqueness of weak solutions for a non-uniformly parabolic equation*, J. Funct. Anal. 257 (2009), no. 10, pp. 3021–3042.

[7] P.G. Ciarlet, *Finite element methods (Part 1)*, In: P.G. Ciarlet and J.L. Lions (eds.), Handbook of numerical analysis, volume II, North-Holland, Amsterdam, 1991.

[8] B. Dacorogna, *Direct methods in the calculus of variations*, Springer, New York, 2008.

[9] T. Donaldson, *Inhomogeneous Orlicz-Sobolev spaces and nonlinear parabolic initial value problems*, J. Diff. Eqs. 16 (1974), pp. 201–256.

[10] L. Diening and M. Růžička, *Interpolation operators in Orlicz–Sobolev spaces*, Numer. Math. 107 (2007) 1, pp. 107–129.

[11] A. Elmahi, *Strongly nonlinear parabolic initial-boundary value problems in Orlicz spaces*, Electron. J. Differ. Eqs. 9 (2002), pp. 203-220.

[12] A. Elmahi and D. Meskine, *Parabolic equations in Orlicz spaces*, J. London Math. Soc. (2), 72 (2005), pp. 410–428.

[13] A. Elmahi and D. Meskine, *Strongly nonlinear parabolic equations with natural growth term in Orlicz spaces*, Nonlinear Analysis 60(2005), pp. 1–35.

[14] H. Gajewski, K. Gröger, and K. Zacharias, *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, Akademie-Verlag, Berlin, 1974.

[15] J.-P. Gossez, *Nonlinear elliptic boundary value problems for equations with rapidly (or slowly) increasing coefficients*, Transactions Amer. Mathem. Soc. 190 (1974), pp. 163–205.

[16] P. Gwiazda and A. Świerczewska-Gwiazda, *On non-Newtonian fluids with a property of rapid thickening under different stimulus*, Math. Mod. Meth. Appl. Sci. 18 (2008) 7, pp. 1073–1092.

[17] P. Gwiazda and A. Świerczewska-Gwiazda, *Parabolic equations in anisotropic Orlicz spaces with general $\mathscr{N}$-function*, Nonlinear Parabolic Problems: H. Amann Festschrift, to appear.

[18] P. Gwiazda, A. Świerczewska-Gwiazda, and A. Wróblewska, *Monotonicity methods in generalized Orlicz spaces for a class of non-Newtonian fluids*, Math. Meth. Appl. Sci. 33 (2010) 2, pp. 125–135.

[19] M.A. Krasnosel'skij and Ya.B. Rutitskij, *Convex functions and Orlicz spaces*, Noordhoff, Groningen, 1961.

[20] A. Kufner, O. John, and S. Fučik, *Function spaces*, Noordhoff, Publishing House of the Czechoslovak Academy of Sciences, Prag, 1977.

[21] V. Mustonen and M. Tienari, *On monotone-like mappings in Orlicz-Sobolev spaces*, Math. Bohem. 124 (1999) 2-3, pp. 255–271.

[22] J. Nečas, *Les méthodes directes en théorie des équations elliptiques*, Masson, Paris, 1967.

[23] A. Novotný and I. Straškraba, *Introduction to the mathematical theory of compressible flow*, Oxford Univ. Press, Oxford, 2004.

[24] P. Perona and J. Malik, *Scale-space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Machine Intell., 12 (1990), pp. 629–639.

[25] J. Robert, *Inéquations variationnelles paraboliques fortement non linéaires*, J. Math. Pures Appl. 53 (1974), pp. 299–321.

[26] R. Schumann, *The Galerkin approximation for quasilinear elliptic equations with rapidly (or slowly) increasing coefficients*, Zeitschr. für Anal. und ihre Anwend. 1 (1982) 4, pp. 73–85.

[27] M. S. Skaff, *Vector valued Orlicz spaces. Generalized N-Functions, I*, Pacific J. Math. 28 (1969) 1,pp. 193–206.

[28] M. S. Skaff, *Vector valued Orlicz spaces, II*, Pacific J. Math. 28 (1969) 2, pp. 413–430.

[29] H. Sohr, *Navier–Stokes equations: an elementaty functional analytic approach*, Birkhauser Verlag, Basel, 2001.

[30] R. Temam, *Numerical analysis*, D. Reidel Publ. Company, Dordrecht, 1973.

[31] E. Zeidler, *Nonlinear functional analysis and its applications III: variational methods and optimization*, Springer, New York, 1985.