Mimetic finite difference approximation of quasilinear elliptic problems

Paola F. Antonietti · Nadia Bigoni · Marco Verani

1 Introduction

In this paper we address the discretization of the following quasilinear elliptic problem:

$$-\mathrm{div}(\kappa(|\nabla u|^2)\nabla u) = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega, \tag{1}$$

P. F. Antonietti · N. Bigoni (✉) · M. Verani
MOX, Dipartimento di Matematica "F. Brioschi", Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy
e-mail: nadia.bigoni@mail.polimi.it

P. F. Antonietti
e-mail: paola.antonietti@polimi.it

M. Verani
e-mail: marco.verani@polimi.it

where f is a given function, and Ω is an open, bounded set of $\mathbb{R}^2$. The approximation of problem (1) is a challenging standalone problem, see, e.g., [27,30,32,45] and [36,37] in the context of conforming and discontinuous finite element methods, respectively. For the finite volume approximation of nonlinear elliptic problems we refer for example to [1,19].

At the the same time, efficient discretizations of problem (1) represent a necessary intermediate step towards the solution of non-newtonian flow problems, see, e.g., [6–10,18,39,44] and [28] again in the conforming and discontinuous finite element context, respectively. It is important to remark that non-newtonian fluids are ubiquitous in industrial applications. For example, for polymeric extrusion, robust and reliable numerical algorithms are mandatory in order to predict and, possibly, to optimize the production process, c.f. [5,43], for example.

The aim of this paper is to show that the flexibility of the Mimetic Finite Difference (MFD) method (see, e.g., [21–24] for a detailed introduction) can be successfully exploited to discretize quasilinear elliptic problems, and that the Kačanov algorithm can be successfully employed to solve the discrete nonlinear system resulting from MFD discretizations; we refer to [47] for more details on the Kačanov method and to [34] for an application in the context of adaptive finite element methods. The MFD method can be naturally employed on very general decompositions made of (possibly non convex) polyhedrals which do not have to fulfill matching conditions. Thanks to such a flexibility, the MFD method has been rapidly applied to a wide range of problems [2,11,15–17,21,41,42], and to Maxwell's equations and magnetostatic fields problems [38,40], Stokes equations [13,14], eigenvalue problems [26], problems governed by variational inequalities [3] and optimal control problems [4]. Very recently, a very promising evolution of the MFD scheme, namely the Virtual Element Method, has been introduced in [12,25,29]. However, the application to nonlinear elliptic problems has not been taken into account yet. The aim of this work is to approximate the solution of a quasilinear elliptic problem of monotone type by using the MFD method. We show that the MFD solution converges to the exact analytical solution in a suitable discrete energy norm. The convergence holds under a suitable approximation assumption which is verified through numerical experiments. The resulting nonlinear problem is then solved iteratively via linearization by applying the Kačanov method, and the convergence of the Kačanov algorithm is proved in the discrete mimetic framework.

The rest of the paper is organized as follows. In Sect. 2 we introduce the weak form of (1) and state its well–posedness. The MFD discrete formulation and its well-posedness is discussed in Sect. 3, and the a priori error estimates are derived in Sect. 4. In Sect. 5, we introduce the Kačanov method in the mimetic context, and prove that the algorithm converges to the mimetic discrete solution. Finally, in Sect. 6 we present some numerical results to confirm our theoretical analysis.

2 A quasilinear elliptic problem

Throughout the paper we will use standard notation for Sobolev spaces, norms and seminorms (see, e.g., [33]). Moreover, the symbol $\lesssim$ denotes an upper bound that holds up to a positive constant independent of the discretization parameters.

We first introduce the following semilinear form

$$b(u; v, w) := \int_\Omega \kappa(|\nabla u|^2)\, \nabla v \cdot \nabla w \, dx \quad \forall u, v, w \in H_0^1(\Omega), \tag{2}$$

where $\kappa : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is a nonlinear function whose properties will be stated below (cf. Assumption 21), and Ω is a sufficiently smooth two-dimensional domain. We are interested in solving the following quasilinear elliptic problem: find $u \in H_0^1(\Omega)$ such that

$$b(u; u, v) = F(v) \quad \forall v \in H_0^1(\Omega), \tag{3}$$

where

$$F(v) := \int_\Omega f v \, dx,$$

for a given $f \in L^2(\Omega)$. We remark that relevant physical applications, as the modeling of non-Newtonian fluids governed by the Carreau law:

$$\kappa(t) := \eta_\infty + (\eta_0 - \eta_\infty)(1 + \lambda t)^{\frac{p-2}{2}}, \quad t \geq 0, \tag{4}$$

with $\eta_0 \geq \eta_\infty \geq 0$, $\lambda > 0$ and $p \in (1, \infty)$, fit in this framework.

Throughout the paper, the nonlinear function $\kappa : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is assumed to satisfy the following.

Assumption 21 *(Assumptions on the nonlinearity)* The nonlinear function κ : $\mathbb{R}^+ \longrightarrow \mathbb{R}^+$ satisfies the following

 i) $\kappa(\cdot)$ is continuous over $[0, +\infty)$;
 ii) there exist two positive constants k_*, k^* such that:

$$k_*(t - s) \leq \kappa(t^2)t - \kappa(s^2)s \leq k^*(t - s) \quad \forall t > s \geq 0. \tag{5}$$

We observe that, by choosing $s = 0$ into (5) we get

$$k_* \leq \kappa(t^2) \leq k^* \quad \forall t > 0. \tag{6}$$

Remark 1 The Carreau law (4) satisfies Assumption (21) provided that $p \in (1, 2)$ and $\eta_\infty > 0$. To prove the upper bound in (5), it is sufficient to observe that the function

$$\varphi(t) := \frac{t}{(1 + \lambda t^2)^\nu} \quad \nu \in \left(0, \frac{1}{2}\right)$$

is Lipschitz continuous with constant equal to one. Indeed,

$$\varphi'(t) = \frac{1}{(1 + \lambda t^2)^\nu} \left[1 - \frac{2\nu\lambda t^2}{1 + \lambda t^2}\right]$$

is positive and bounded from above (the maximum is equal to one) and it tends to zero as t goes to infinity. Therefore, setting $\nu := \frac{2-p}{2}$, we have

$$\eta(t^2)t - \eta(s^2)s = \eta_\infty(t-s) + (\eta_0 - \eta_\infty)\left[\frac{t}{(1+\lambda t^2)^\nu} - \frac{s}{(1+\lambda s^2)^\nu}\right] \le \eta_0(t-s).$$

The lower bound in (5) follows straightforwardly

$$\eta(t^2)t - \eta(s^2)s = \eta_\infty(t-s) + (\eta_0 - \eta_\infty)\left[\frac{t}{(1+\lambda t^2)^\nu} - \frac{s}{(1+\lambda s^2)^\nu}\right] \ge \eta_\infty(t-s),$$

since the term in the square brackets is positive.

Remark 2 The nonlinear function $\kappa(\cdot)$ could also be a function of the space position, i.e., $\kappa : \Omega \times \mathbb{R}^+ \longrightarrow \mathbb{R}^+$. In such cases, Assumption 21 has to be modified as follows:

i) $\kappa(\cdot, \cdot)$ is continuous over $\overline{\Omega} \times [0, +\infty)$;
ii) there exist two positive constants k_*, k^* such that:

$$k_*(t - s) \le \kappa(x, t^2)t - \kappa(x, s^2)s \le k^*(t - s) \quad \forall t > s \ge 0 \quad \forall x \in \Omega.$$

Since we are interested in non-Newtonian fluids governed by the Carreau law (4) we will not address in this paper this case specifically.

Before going after let us recall here some useful results. Let us introduce the function $\beta : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ given by

$$\beta(s) := \frac{1}{2}\int_0^{s^2} \kappa(t)\,dt.$$

By observing that $\beta'(s) = \kappa(s^2)s$, for all $s \in \mathbb{R}^+$, and using Assumption 21, we get

$$k_*(t - s) \le \beta'(t) - \beta'(s) \le k^*(t - s) \quad \forall t > s \ge 0.$$

Next, let us state the following Lemma (we refer to [47, Lemma 25.26] for the proof).

Lemma 1 *Let $\gamma(x) := \beta(|x|)$ for $x \in \mathbb{R}^d$. Under Assumption 21, for all $x, y \in \mathbb{R}^d$ it holds*

$$|\gamma'(x) - \gamma'(y)| \le 3k^*|x - y| \tag{7}$$
$$\langle \gamma'(x) - \gamma'(y), x - y \rangle_{\mathbb{R}^d} \ge k_*|x - y|^2, \tag{8}$$

where $\langle \cdot, \cdot \rangle_{\mathbb{R}^d}$ denotes the Euclidean scalar product of $\mathbb{R}^d$.

Next, we address the well-posedness of problem (3). To this aim, we introduce the nonlinear operator

$$\mathcal{B} : H_0^1(\Omega) \longrightarrow H^{-1}(\Omega) \quad \langle \mathcal{B}u, v \rangle := b(u; u, v) \ \forall v \in H_0^1(\Omega). \tag{9}$$

The existence and uniqueness of such an operator follows from the Riesz's Theorem. Hence, problem (3) can be equivalently stated as follows: find $u \in H_0^1(\Omega)$ such that

$$\mathcal{B}u = F, \tag{10}$$

where $F \in H^{-1}(\Omega)$ is given. The next result is a key point towards the proof of the well-posedness of problem (10).

Lemma 2 *The operator $\mathcal{B}$ defined as in (9) is Lipschitz continuous and strongly monotone with respect to the $H^1(\Omega)$-norm, i.e.,*

$$\|\mathcal{B}u - \mathcal{B}v\|_{H^{-1}(\Omega)} \leq 3k^* \|u - v\|_{H^1(\Omega)} \qquad \forall u, v \in H_0^1(\Omega), \tag{11}$$
$$\langle \mathcal{B}u - \mathcal{B}v, u - v \rangle \geq k_* \|u - v\|_{H^1(\Omega)}^2 \quad \forall u, v \in H_0^1(\Omega).$$

Proof From (7), we obtain that, for all $u, v, z \in H_0^1(\Omega)$, it holds

$$|\langle \mathcal{B}u - \mathcal{B}v, z \rangle| \leq \int_\Omega \left| \kappa(|\nabla u|^2)\nabla u - \kappa(|\nabla v|^2)\nabla v \right| |\nabla z| \, \mathrm{d}x$$
$$\leq 3k^* \int_\Omega |\nabla u - \nabla v||\nabla z| \, \mathrm{d}x \leq 3k^* \|u - v\|_{H^1(\Omega)} \|z\|_{H^1(\Omega)},$$

and

$$\langle \mathcal{B}u - \mathcal{B}v, u - v \rangle = \int_\Omega \left(\kappa(|\nabla u|^2)\nabla u - \kappa(|\nabla v|^2)\nabla v \right) \nabla(u - v) \, \mathrm{d}x$$
$$\geq k_* \int_\Omega |\nabla u - \nabla v|^2 \, \mathrm{d}x = k_* \|u - v\|_{H^1(\Omega)}^2.$$

This easily implies that

$$\|\mathcal{B}u - \mathcal{B}v\|_{H^{-1}(\Omega)} \leq 3k^* \|u - v\|_{H^1(\Omega)} \qquad \forall u, v \in H_0^1(\Omega),$$
$$\langle \mathcal{B}u - \mathcal{B}v, u - v \rangle \geq k_* \|u - v\|_{H^1(\Omega)}^2 \qquad \forall u, v \in H_0^1(\Omega).$$

Finally, we can state the well-posedness of problem (10).

Proposition 1 *Let $\mathcal{B}$ be the operator defined as in (9) and let $F \in H^{-1}(\Omega)$ be given. Under Assumption 21, Eq. (10) admits a unique solution $u \in H_0^1(\Omega)$.*

Proof By using the Zarantonello's Theorem (see, e.g., [47, Theorem 25.B]) and Lemma 2 we can conclude that Eq. (10) admits a unique solution $u \in H_0^1(\Omega)$.

3 The MFD discretization

In this section we derive a mimetic discretization of problem (3) by introducing suitable mesh assumptions and the discrete space endowed with a proper scalar product. We refer [3,21,24] for more details on the MFD method for elliptic problems.

3.1 Mesh assumptions, discrete space and norms

Let Ω_h be a non-overlapping partition of Ω into, possibly non-convex polygonal elements E with granularity $h := \sup_{E \in \Omega_h} h_E$, being h_E the diameter of $E \in \Omega_h$. We denote by $\mathcal{N}_h^\circ$ and $\mathcal{N}_h^\partial$ the sets of interior and boundary mesh vertexes, respectively, and set $\mathcal{N}_h = \mathcal{N}_h^\circ \cup \mathcal{N}_h^\partial$. Proceeding as in [24] we also assume the following.

Assumption 31 (*Mesh regularity assumptions*) There exist an integer number N, independent of h, and a compatible sub-decomposition $\mathcal{T}_h$ of every Ω_h into *shape-regular* triangles in such a way that every element E can be decomposed in at most N triangles.

We point out that Assumption 31 only requires the existence of a *compatible* sub-mesh that does not have to be constructed in practice. Moreover, Assumption 31 guarantees that the following mesh regularity properties are satisfied (cf. [21]).

i) There exists $N_{\mathsf{e}} > 0$ such that every element E has at most N_{e} edges;
ii) There exists $\gamma > 0$ which only depends on N_{e} such that, for every element E and for every edge e of E, it holds

$$|\mathsf{e}| \geq \gamma h_E \quad |E| \geq \gamma h_E^2,$$

where $|\mathsf{e}|$ is the length of e and $|E|$ denotes the area of the element E;
iii) For every $E \in \Omega_h$ and for every edge e of E, the following *trace inequality* holds

$$\|\psi\|_{L^2(\mathsf{e})}^2 \lesssim h_E^{-1} \|\psi\|_{L^2(E)}^2 + h_E |\psi|_{H^1(E)}^2 \quad \forall \psi \in H^1(E). \tag{12}$$

Next, we introduce the discrete space by choosing as degrees-of-freedom the nodal values in each vertex $\mathsf{v} \in \mathcal{N}_h$. More precisely, every discrete function v_h is a vector of real components $v_h := \{v^{\mathsf{v}}\}_{\mathsf{v} \in \mathcal{N}_h}$, one per mesh vertex. The space of unknowns V_h is then defined taking into account the homogeneous Dirichlet boundary conditions, i.e.,

$$V_h := \left\{ v_h := \{v^{\mathsf{v}}\}_{\mathsf{v} \in \mathcal{N}_h} : v_h^{\mathsf{v}} = 0 \; \forall \mathsf{v} \in \mathcal{N}_h^\partial \right\}.$$

Clearly, the number of unknowns is equal to the number of interior mesh vertexes. The space V_h is endowed with the following discrete norm that mimics the usual $H_0^1(\Omega)$-norm:

$$\|v_h\|_{1,h}^2 := \sum_{E \in \Omega_h} \|v_h\|_{1,h,E}^2 = \sum_{E \in \Omega_h} |E| \sum_{\substack{e \in \mathcal{E}_h \\ e \subset \partial E}} \left[\frac{1}{|e|} (v^{v_2} - v^{v_1}) \right]^2, \qquad (13)$$

where v_1 and v_2 are the two vertexes of $e \in \mathcal{E}_h$, and $|E|$ is the area of the element $E \in \Omega_h$.

Finally, we introduce the standard nodal interpolation operator satisfying classical approximation estimates [20]. For every function $v \in \mathcal{C}^0(\overline{\Omega}) \cap H_0^1(\Omega)$, the interpolation $v_I \in V_h$ is given by

$$v_I^v := v(v) \quad \forall v \in \mathcal{N}_h. \qquad (14)$$

A local interpolation operator from $\mathcal{C}^0(\overline{E}) \cap H^1(E)$ into V_h^E can be defined analogously, where we denote by V_h^E the restriction of V_h to the $E \in \Omega_h$, i.e., $V_h^E = V_h|_E$.

3.2 The discrete formulation of the problem

The mimetic approximation of problem (3) is essentially based on the following observation. Let us consider the local semilinear form given by the restriction of the form (2) on each element E of the partition, i.e.,

$$b^E(u; v, w) := \int_E \kappa(|\nabla u|^2) \nabla v \cdot \nabla w \, dx \quad \forall u, v, w \in H^1(E). \qquad (15)$$

Next, denoting by $\mathbb{P}^1(E)$ the space of linear polynomials on $E \in \Omega_h$, we can observe that it holds

$$b^E(\varphi; v, w) = \kappa(|\nabla \varphi|^2) \int_E \nabla v \cdot \nabla w \, dx \quad \forall \varphi \in \mathbb{P}^1(E) \ \forall v, w \in H^1(E).$$

In view of the above relation, the idea is to construct a suitable discrete approximation of the nonlinear term $\kappa(|\nabla \varphi|^2)$ and of the integral one.

To this aim, we introduce a symmetric bilinear form $a_h^E(\cdot, \cdot) : V_h^E \times V_h^E \longrightarrow \mathbb{R}$ that locally mimics

$$a_h^E(v_h, w_h) \sim \int_E \nabla \widetilde{v}_h \cdot \nabla \widetilde{w}_h \, dx. \qquad (16)$$

In this context, $\widetilde{v}_h, \widetilde{w}_h$ can be interpreted as regular functions living on E which "extend the data" v_h, w_h inside the element and the symbol $\sim$ stands for approximation. As shown in [21], the local forms $a_h^E(\cdot, \cdot)$ can be constructed in such a way that the following properties are satisfied.

Proposition 2 *For every $E \in \Omega_h$, it holds*

$$\|v_h\|_{1,h,E}^2 \lesssim a_h^E(v_h, v_h), \qquad a_h^E(u_h, v_h) \lesssim \|u_h\|_{1,h,E} \|v_h\|_{1,h,E} \quad \forall u_h, v_h \in V_h^E.$$

Moreover, for every $E \in \Omega_h$ and every linear function $\varphi \in \mathbb{P}^1(E)$

$$a_h^E(v_h, \varphi_I) = \sum_{\substack{e \in \mathcal{E}_h \\ e \subset \partial E}} (\nabla \varphi \cdot \mathbf{n}_E^e) \frac{|e|}{2} \left(v_h^{v_1} + v_h^{v_2} \right) \quad \forall v_h \in V_h^E, \tag{17}$$

where v_1 and v_2 are the two vertexes of $e \in \mathcal{E}_h$, and $\mathbf{n}_E^e$ is the unit normal vector to $e \in \mathcal{E}_h$ that points outside E.

Identity (17) is usually refereed to as *local consistency* property and is related to the following observation. Let E any element in Ω_h. For all $v \in H^1(E)$ and for all linear functions φ on E, using an integration by parts formula, we obtain

$$\int_E \nabla v \cdot \nabla \varphi \, dx = -\int_E \Delta \varphi \, v \, dx + \sum_{e \in \mathcal{E}_h^E} \int_e (\nabla \varphi \cdot \mathbf{n}_E^e) v \, ds$$

$$= \sum_{e \in \mathcal{E}_h^E} \nabla \varphi \cdot \mathbf{n}_E^e \int_e v \, ds. \tag{18}$$

If we substitute the integral in the last term of (18) with a trapezium integration rule, we get condition (17) which means that that the discrete bilinear form $a_h^E(\cdot, \cdot)$ respects integration by parts formula whenever tested with linear functions. As pointed out in [3], the local form $a_h^E(\cdot, \cdot)$ can be written in an algebraic manner by introducing a suitable symmetric and positive definite local matrix $\mathbb{A}_h^E \in \mathbb{R}^{k_E \times k_E}$ such that

$$a_h^E(u_h, v_h) = u_h^T \mathbb{A}_h^E v_h \quad \forall u_h, v_h \in V_h^E,$$

where k_E denotes the number of vertexes of $E \in \Omega_h$.

Next, we discuss the mimetic discretization of the nonlinear term within each element. To this aim, for every $E \in \Omega_h$, we introduce the operator

$$\mathcal{G}_h^E : V_h^E \longrightarrow \mathbb{R}^+ \ \mathcal{G}_h^E(u_h) := \frac{a_h^E(u_h, u_h)}{|E|} = \frac{u_h^T \mathbb{A}_h^E u_h}{|E|}. \tag{19}$$

Bearing in mind (16), the above operator turns out to be a good candidate to approximate $|\nabla u_h|^2$ within each element, as the following heuristics shows:

$$\frac{\int_E |\nabla u|^2}{|E|} \sim \mathcal{G}_h^E(u_I) \forall u \in \mathcal{C}^0(\overline{E}) \cap H^1(E),$$

where $u_I \in V_h^E$ is defined according to (14).

In view of the above discussion, we define the mimetic discretization of the local semi-linear form (15) as follows:

$$b_h^E(u_h; v_h, w_h) := \kappa(\mathcal{G}_h^E(u_h)) \, a_h^E(v_h, w_h) \quad \forall u_h, v_h, w_h \in V_h^E,$$

and set

$$b_h(u_h; v_h, w_h) := \sum_{E \in \Omega_h} b_h^E(u_h; v_h, w_h) \quad \forall u_h, v_h, w_h \in V_h. \tag{20}$$

Then, the mimetic discretization of problem (3) reads: find $u_h \in V_h$, such that

$$b_h(u_h; u_h, v_h) = F_h(v_h) \quad \forall v_h \in V_h. \tag{21}$$

The right-hand-side of problem (21) is defined as

$$F_h(v_h) := \sum_{E \in \Omega_h} \overline{f}_E \sum_{i=1}^{k_E} v^{\mathsf{v}_i} \omega_E^i,$$

where $\mathsf{v}_1, \ldots, \mathsf{v}_{k_E}$ are the vertexes of E, $\omega_E^1, \ldots, \omega_E^{k_E}$ are positive weights such that $\sum_{i=1}^{k_E} \omega_E^i = |E|$, and

$$\overline{f}_E := \frac{1}{|E|} \int_E f \, \mathrm{d}x.$$

3.2.1 Well-posedness of the discrete problem

This section is devoted to prove the well-posedness of the discrete problem (21).

The following straightforward generalization of Lemma 1 will be useful in the forthcoming analysis.

Corollary 1 *Let* $\mathbb{M} \in \mathbb{R}^{d \times d}$ *be a symmetric and positive definite matrix, and let* $\langle \cdot, \cdot \rangle_{\mathbb{M}}$ *and* $\| \cdot \|_{\mathbb{M}}$ *the induced scalar product and norm, respectively, i.e.,*

$$\langle x, y \rangle_{\mathbb{M}} := x^T \mathbb{M} y, \quad \|x\|_{\mathbb{M}}^2 := x^T \mathbb{M} x \quad \forall x, y \in \mathbb{R}^d.$$

Under Assumption 21, the function $\gamma_{\mathbb{M}}(x) := \beta(\|x\|_{\mathbb{M}})$ *satisfies*

$$|\langle \gamma_{\mathbb{M}}'(x) - \gamma_{\mathbb{M}}'(y), x - y \rangle_{\mathbb{M}}| \leq 3k_* \|x - y\|_{\mathbb{M}}^2$$
$$\langle \gamma_{\mathbb{M}}'(x) - \gamma_{\mathbb{M}}'(y), x - y \rangle_{\mathbb{M}} \geq k_* \|x - y\|_{\mathbb{M}}^2,$$

for all $x, y \in \mathbb{R}^d$.

Proceeding as before, we introduce the discrete operator

$$\mathcal{B}_h : V_h \longrightarrow \mathbb{R}, \langle \mathcal{B}_h u_h, v_h \rangle := b_h(u_h; u_h, v_h) \forall v_h \in V_h, \tag{22}$$

and rewrite problem (21) as: find $u_h \in V_h$ such that

$$\mathcal{B}_h u_h = F_h, \tag{23}$$

where F_h is given.

Lemma 3 *The operator $\mathcal{B}_h$ defined as in* (22) *is Lipschitz continuous and strongly monotone with respect to the discrete norm* (13)*, i.e.,*

$$|\langle \mathcal{B}_h u_h - \mathcal{B}_h v_h, u_h - v_h \rangle| \lesssim \|u_h - v_h\|_{1,h}^2 \qquad \forall\, u_h, v_h \in V_h \tag{24}$$

$$\langle \mathcal{B}_h u_h - \mathcal{B}_h v_h, u_h - v_h \rangle \gtrsim \|u_h - v_h\|_{1,h}^2 \qquad \forall\, u_h, v_h \in V_h. \tag{25}$$

Proof The proof follows by applying Corollary 1 to the following function

$$\widetilde{\gamma}(u_h) := |E|\beta \left(\frac{\|u_h\|_{\mathbb{A}_h^E}}{|E|^{1/2}} \right) = |E|\beta \left(\sqrt{\mathcal{G}_h^E u_h} \right),$$

after having observed that it holds

$$\langle \widetilde{\gamma}'(u_h), v_h \rangle = \kappa(\mathcal{G}_h^E u_h) a_h^E(u_h, v_h) = b_h^E(u_h; u_h, v_h) \quad \forall u_h, v_h \in V_h.$$

Finally, the following well-posedness result holds.

Proposition 3 *Let $\mathcal{B}_h$ be the operator defined as in* (22)*. Under Assumption 21, Eq.* (23) *admits a unique solution $u_h \in V_h$.*

Proof The thesis follows by the Zarantonello's Theorem and the Lipschitz continuity and strongly monotonicity of the operator $\mathcal{B}_h$ proved in Lemma 3. □

4 Error analysis

In this section we derive the a-priori error estimates in the mesh-dependent norm (13) to prove the convergence of the MFD approximate solution to the exact solution of problem (3).

Let us preliminary prove the following result that will be useful in the subsequent analysis.

Lemma 4 *For every $u_h \in V_h$ it holds*

$$|b_h(u_h; v_h, w_h)| \lesssim \|v_h\|_{1,h} \|w_h\|_{1,h} \quad \forall w_h, v_h \in V_h. \tag{26}$$

and

$$b_h(u_h; w_h - v_h, w_h - v_h) \gtrsim \|w_h - v_h\|_{1,h}^2 \quad \forall w_h, v_h \in V_h. \tag{27}$$

Proof We first show (26). By observing that $b_h(\cdot; \cdot, \cdot)$ is symmetric and linear in its second and third arguments and by using (6) and Proposition 2, we get

$$
\begin{aligned}
|b_h(u_h; v_h, w_h)| &\leq \sum_{E \in \Omega_h} |\kappa(\mathcal{G}_h^E(u_h)) a_h^E(v_h, w_h)| \\
&\lesssim \sum_{E \in \Omega_h} \|v_h\|_{1,h,E} \|w_h\|_{1,h,E} \lesssim \|v_h\|_{1,h} \|w_h\|_{1,h}.
\end{aligned}
$$

Similarly, we have

$$b_h(u_h; w_h - v_h, w_h - v_h) \gtrsim \sum_{E \in \Omega_h} \|w_h - v_h\|_{1,h,E}^2 = \|w_h - v_h\|_{1,h}^2,$$

which is (27), and the proof is complete.

Next, we introduce the following lifting operator.

Lemma 5 *For every $E \in \Omega_h$, there exists a local lifting operator*

$$\mathcal{R}_h^E \; : \; V_h^E \longrightarrow H^1(E) \cap \mathcal{C}^0(\overline{E}),$$

that satisfies the following properties:

(L1) it vanishes on the boundary of Ω_h;
(L2) $\mathcal{R}_h^E v_h|_{\mathsf{e}}$ is a linear function $\forall \mathsf{e} \in \mathcal{E}_h$, $\mathsf{e} \subseteq \partial E$ and $\forall v_h \in V_h^E$;
(L3) $|\mathcal{R}_h^E v_h|_{H^1(E)}^2 \lesssim \|v_h\|_{1,h,E}^2 \quad \forall v_h \in V_h^E$;
(L4) $\|R_h^E v_h - v_h^{\vee}\|_{L^2(E)} \lesssim h_E \|v_h\|_{1,h,E} \; \forall v_h \in V_h^E$.

We observe that it is not necessary to build explicitly the lifting operator, but it is sufficient to show that it exists. A detailed way to construct it and more details about its properties con be found in [24]. We can then extend this definition, by introducing the *global lifting operator* given by

$$
\begin{aligned}
\mathcal{R}_h \; &: \; V_h \longrightarrow H^1(\Omega_h) \cap \mathcal{C}^0(\overline{\Omega}_h), \\
\mathcal{R}_h(v_h)|_E &:= \mathcal{R}_h^E(v_h|_E) \quad \forall v_h \in V_h, \; E \in \Omega_h.
\end{aligned}
$$

Finally, the following result states the convergence of the MFD approximation towards the exact solution of problem (3).

Theorem 1 *Let $u \in H^2(\Omega) \cap H_0^1(\Omega)$ be the solution of the continuous problem (3) and let u_h be the solution of the discrete problem (21). Moreover, let u_I be the*

interpolation of the exact solution u defined according to (14). *Finally, assume that there exists $\alpha > 0$ so that*

$$\|\kappa(|\nabla v|^2) - \kappa(\mathcal{G}_h^E(v_I))\|_{L^\infty(\Omega)} \lesssim h^\alpha \quad \forall v \in \mathcal{C}^0(\overline{\Omega}) \cap H_0^1(\Omega). \tag{28}$$

Then, it holds

$$\|u_I - u_h\|_{1,h} \lesssim h^{\min(1,\alpha)}. \tag{29}$$

.

Proof Throughout the proof we set $e_h := u_h - u_I$. For every $E \in \Omega_h$, let u_E^1 be the $L^2(E)$-projection of u onto the space of linear polynomials on E. We then define the piecewise discontinuous linear function u^1 as $u^1|_E := u_E^1$ for all $E \in \Omega_h$. With a little abuse of notation, in the following we indicate with $(u^1)_I$ a collection of nodal values such that for all elements E the restriction of $(u^1)_I$ to $E \in \Omega_h$ is given by the local interpolation $(u_E^1)_I$. Note that both the $\|\cdot\|_{1,h}$ norm and the bilinear form $a_h(\cdot, \cdot)$ can be immediately extended to $(u^1)_I$, since both operators are a sum of local terms. By using (25) and the discrete problem (21) we get

$$
\begin{aligned}
\|e_h\|_{1,h}^2 &\lesssim b_h(u_h; u_h, e_h) - b_h(u_I; u_I, e_h) = F_h(e_h) - b_h(u_I; u_I, e_h) \\
&= F_h(e_h) - b_h(u_I; u_I - (u^1)_I, e_h) - b_h(u_I; (u^1)_I, e_h) \\
&= F_h(e_h) - b_h(u_I; u_I - (u^1)_I, e_h) + \sum_{E \in \Omega_h} \delta_E,
\end{aligned} \tag{30}
$$

where $\delta_E := -\kappa(\mathcal{G}_h^E(u_I)) \, a_h^E((u^1)_I, e_h)$. We next estimate the local term δ_E. We first observe that, thanks to (L2)

$$\frac{|\mathsf{e}|}{2}\left(e_h^{v_1} + e_h^{v_2}\right) = \int_{\mathsf{e}} \mathcal{R}_h^E e_h \, dx \quad \forall E \in \Omega_h, \mathsf{e} \in \mathcal{E}_h^E.$$

Then, by integrating twice by parts and using the fact that $\mathcal{R}_h e_h$ vanishes on the boundary of Ω_h, it follows

$$
\begin{aligned}
\delta_E &= -\kappa(\mathcal{G}_h^E(u_I)) \sum_{\substack{\mathsf{e} \in \mathcal{E}_h \\ \mathsf{e} \subseteq \partial E}} \frac{\partial u^1}{\partial \mathbf{n}_E^\mathsf{e}} \int_{\mathsf{e}} \mathcal{R}_h^E e_h \, dx \\
&= -\kappa(\mathcal{G}_h^E(u_I)) \int_E \nabla \mathcal{R}_h^E e_h \cdot \nabla u^1 \, dx \\
&= \kappa(\mathcal{G}_h^E(u_I)) \int_E \nabla \mathcal{R}_h^E e_h \cdot \nabla (u - u^1) \, dx - \kappa(\mathcal{G}_h^E(u_I)) \int_E \nabla \mathcal{R}_h^E e_h \cdot \nabla u \, dx \\
&= \kappa(\mathcal{G}_h^E(u_I)) \int_E \nabla \mathcal{R}_h^E e_h \cdot \nabla (u - u^1) \, dx
\end{aligned}
$$

$$+ \int_E \left\{ \kappa(|\nabla u|^2) - \kappa(\mathcal{G}_h^E(u_\mathtt{I})) \right\} \nabla \mathcal{R}_h^E e_h \cdot \nabla u \, dx$$

$$+ \int_E \mathrm{div}\left(\kappa(|\nabla u|^2)\nabla u \right) \mathcal{R}_h^E e_h \, dx. \tag{31}$$

Finally, by using problem (3) and employing (31) into (30) together with Proposition 2 we have

$$\|e_h\|_{1,h}^2 \lesssim \left| F_h(e_h) - (f, \mathcal{R}_h e_h) \right| + \|u_\mathtt{I} - (u^1)_\mathtt{I}\|_{1,h} \|e_h\|_{1,h}$$

$$+ \left| \sum_{E \in \Omega_h} \kappa(\mathcal{G}_h^E(u_\mathtt{I})) \int_E \nabla \mathcal{R}_h^E e_h \cdot \nabla(u - u^1) \, dx \right|$$

$$+ \left| \sum_{E \in \Omega_h} \int_E \left\{ \kappa(|\nabla u|^2) - \kappa(\mathcal{G}_h^E(u_\mathtt{I})) \right\} \nabla \mathcal{R}_h^E e_h \cdot \nabla u \, dx \right|$$

$$:= \mathtt{I} + \mathtt{II} + \mathtt{III} + \mathtt{IV}.$$

Let us estimate separately the terms $\mathtt{I}$–$\mathtt{IV}$. Proceeding as in the estimate of the *First Piece* in [21], there holds

$$\mathtt{I} = \left| F_h(e_h) - (f, \mathcal{R}_h e_h) \right| \lesssim h \|f\|_{L^2(\Omega)} \|e_h\|_{1,h} \lesssim h \|e_h\|_{1,h}.$$

Applying the trace inequality (12) and employing a standard interpolation error estimate yield to

$$\mathtt{II} \lesssim \left(\sum_{E \in \Omega_h} \|\nabla(u - u^1)\|_{L^2(E)}^2 + h_E^2 \, |u|_{H^2(E)}^2 \right)^{1/2} \|e_h\|_{1,h} \lesssim h \, |u|_{H^2(\Omega)} \|e_h\|_{1,h}.$$

Term $\mathtt{III}$ can be estimated using the Cauchy–Schwarz inequality, assumption (L3) and again standard interpolation error estimates

$$\mathtt{III} \lesssim \|\nabla \mathcal{R}_h^E e_h\|_{L^2(\Omega_h)} \|\nabla(u - u^1)\|_{L^2(\Omega_h)} \lesssim h \|e_h\|_{1,h} \, |u|_{H^2(\Omega)} \lesssim h \|e_h\|_{1,h} \,.$$

Finally, by using Assumption (28) and (L3), we get

$$\mathtt{IV} \leq \sum_{E \in \Omega_h} \|\kappa(|\nabla u|^2) - \kappa(\mathcal{G}_h^E(u_\mathtt{I}))\|_{L^\infty(\Omega)} \int_E \left| \nabla \mathcal{R}_h^E e_h \cdot \nabla u \right| dx$$

$$\lesssim h^\alpha \|\nabla \mathcal{R}_h^E e_h\|_{L^2(\Omega_h)} \|\nabla u\|_{L^2(\Omega_h)} \lesssim h^\alpha \|e_h\|_{1,h} \, |u|_{H^1(\Omega)} \,.$$

Combining the estimates of $\mathtt{I}$–$\mathtt{IV}$ with (4) yields the result.

Remark 3 (*On Assumption* (28)) Theorem 1 relies on assumption (28), whose theoretical justification is still object of an on-going research. However, the numerical investigation performed in Sect. 6 suggests the validity of (28) with $\alpha = 1$, so that the error estimate (29) reduces to $\|u_{\mathrm{I}} - u_h\|_{1,h} \lesssim h$.

5 The Kačanov method

In this section we present the Kačanov method (see, e.g., [47]) to solve the nonlinear problem (21). For $k \geq 0$, the idealized algorithm without any stopping criterion (that will discussed later) reads as follows: given $u_h^{(k)} \in V_h$, find $u_h^{(k+1)} \in V_h$ such that

$$b_h(u_h^{(k)}; u_h^{(k+1)}, v_h) = F_h(v_h) \quad \forall v_h \in V_h . \tag{32}$$

The convergence of the sequence $\{u_h^{(k)}\}_{k \geq 0}$ to the "exact" discrete solution u_h of problem (21) is stated in Theorem 2, below. The proof takes the steps from the general ideas of [47, Theorem 25.B] and requires some preliminary results that will be collected in the following.

We introduce the functional $E_h : V_h \longrightarrow \mathbb{R}$ defined as

$$E_h(v_h) := \frac{1}{2} \sum_{E \in \Omega_h} |E| \beta \left(\sqrt{\mathcal{G}_h^E(v_h)} \right) \tag{33}$$

that represents the energy associated to system (21). It is easy to see that

$$\langle E_h'(u_h), v_h \rangle = b_h(u_h; u_h, v_h) \quad \forall u_h, v_h \in V_h. \tag{34}$$

Finally, we state two auxiliary lemmas that will be crucial to prove the convergence of the Kačanov algorithm.

Lemma 6 *Let $E_h(\cdot)$ be defined according to* (33) *and let assume that $\kappa(\cdot)$ is monotone decreasing, i.e., $\kappa(t_1) \geq \kappa(t_2)$ for all $0 \leq t_1 \leq t_2$. Then,*

$$E_h(u_h) - E_h(v_h) \leq \frac{1}{2} [b_h(v_h; u_h, u_h) - b_h(v_h; v_h, v_h)], \tag{35}$$

$$\langle E'(u_h) - E'(v_h), u_h - v_h \rangle \gtrsim \|u_h - v_h\|_{1,h}^2 \tag{36}$$

for all $u_h, v_h \in V_h$.

Proof Inequality (35) can be proved by using that $k(\cdot)$ is monotone decreasing and (19), (20) and (33)

$$E_h(u_h) - E_h(v_h) = \frac{1}{2} \sum_{E \in \Omega_h} |E| \left[\int_0^{\mathcal{G}_h^E(u_h)} \kappa(s) \, ds - \int_0^{\mathcal{G}_h^E(v_h)} \kappa(s) \, ds \right]$$

$$\leq \frac{1}{2} \sum_{E \in \Omega_h} |E| \kappa \left(\mathcal{G}_h^E(v_h) \right) \left[\frac{a_h^E(u_h, u_h)}{|E|} - \frac{a_h^E(v_h, v_h)}{|E|} \right]$$

$$\leq \frac{1}{2} \left[b_h(v_h; u_h, u_h) - b_h(v_h; v_h, v_h) \right].$$

Inequality (36) easily follows from (34) and the strongly monotonicity of the operator $\mathcal{B}_h$ shown in Lemma 3. □

We also have the following result that is a consequence of Lemma 4.

Lemma 7 *For every $k \geq 0$, given $u_h^{(k)} \in V_h$, the variational problem (32) admits a unique solution $u_h^{(k+1)} \in V_h$ that is also the minimum of the quadratic variational problem*

$$\min_{v_h \in V_h} \mathcal{H}(v_h) = \min_{v_h \in V_h} \frac{1}{2} b_h \left(u_h^{(k)}; v_h, v_h \right) - F_h(v_h),$$

where $\mathcal{H} : V_h \longrightarrow \mathbb{R}$. On the other hand, the functional $\mathcal{H}(\cdot)$ admits a unique minimum that is also solution to (32).

Proof The first part of the proof follows immediately from Lemma 4. On the other hand, for $u_h^{(k)} \in V_h$ fixed, we observe that

$$\langle \mathcal{H}'(v_h), w_h \rangle = b_h(u_h^{(k)}; v_h, w_h) - F_h(w_h) \quad \forall v_h, w_h \in V_h ,$$

and from Lemma 4 we deduce

$$\langle \mathcal{H}'(v_h) - \mathcal{H}'(w_h), v_h - w_h \rangle \gtrsim \|v_h - w_h\|_{1,h}^2 \quad \forall v_h, w_h \in V_h .$$

By following [47, Problem 25.3] we get that the functional $\mathcal{H}$ is weakly coercive, strictly convex and continuous and therefore it admits a unique minimum u_h^*. Moreover $\langle \mathcal{H}'(u_h^*), w_h - u_h^* \rangle \geq 0$, for all $w_h \in V_h$ and this implies that $u_h^* = u_h^{(k+1)}$ is the unique solution to (32). □

We are now ready to prove the convergence of the Kačanov method (32).

Theorem 2 *Let $\{u_h^{(k)}\}_{k \geq 0}$ be the sequence built by the Kačanov method, then $u_h^{(k)} \longrightarrow u_h$ in V_h, as $k \to +\infty$.*

Proof We preliminary observe that, by defining $\mathcal{L}_h(u_h) := E_h(u_h) - F_h(u_h)$, the following minimization problem

$$\min_{v_h \in V_h} \mathcal{L}(v_h)$$

admits a unique solution thanks to the well-posedness of problem (21) together with (34). Next, let us define

$$g(u_h) := \frac{1}{2} \left(b_h(u_h^{(k)}; u_h, u_h) - b_h(u_h^{(k)}; u_h^{(k)}, u_h^{(k)}) \right) + E_h(u_h^{(k)}) - F_h(u_h).$$

Thanks to (35), we get

$$E_h(u_h^{(k+1)}) - F_h(u_h^{(k+1)}) \le \frac{1}{2} \left(b_h(u_h^{(k)}; u_h^{(k+1)}, u_h^{(k+1)}) - b_h(u_h^{(k)}; u_h^{(k)}, u_h^{(k)}) \right)$$
$$+ E_h(u_h^{(k)}) - F_h(u_h^{(k+1)}),$$

that is

$$\mathcal{L}_h(u_h^{(k+1)}) \le g(u_h^{(k+1)}).$$

From Lemma 7 we deduce that $u_h^{(k+1)} \in V_h$ is the minimum of the functional $g(\cdot)$. Then, we get

$$\mathcal{L}_h(u_h^{(k+1)}) \le g(u_h^{(k+1)}) \le g(u_h^{(k)}) = \mathcal{L}_h(u_h^{(k)}). \tag{37}$$

So, we can conclude that the sequence $\{\mathcal{L}_h(u_h^{(k)})\}_{k \ge 0}$ is monotone decreasing and is limited from below, since the operator $\mathcal{L}_h(\cdot)$ admits a minimum. Consequently,

$$\lim_{k \to \infty} \left\{ \mathcal{L}_h(u_h^{(k+1)}) - \mathcal{L}_h(u_h^{(k)}) \right\} = 0.$$

Finally, by using (37), the linearity of the form $b_h(\cdot; \cdot, \cdot)$ in its second and third arguments and property (27) we get

$$\mathcal{L}(u_h^{(k)}) - \mathcal{L}_h(u_h^{(k+1)}) \ge \mathcal{L}(u_h^{(k)}) - g(u_h^{(k+1)})$$
$$= F_h(u_h^{(k+1)} - u_h^{(k)}) - \frac{1}{2}(b_h(u_h^{(k)}; u_h^{(k+1)}, u_h^{(k)}) - b_h(u_h^{(k)}; u_h^{(k)}, u_h^{(k)}))$$
$$= \frac{1}{2} b_h(u_h^{(k)}; u_h^{(k+1)} - u_h^{(k)}, u_h^{(k+1)} - u_h^{(k)}) \gtrsim \|u_h^{(k+1)} - u_h^{(k)}\|_{1,h}^2.$$

So we can conclude that

$$\|u_h^{(k+1)} - u_h^{(k)}\|_{1,h} \to 0, \tag{38}$$

as k tends to $+\infty$. Next, we prove that the sequence $\{u_h^{(k)}\}_{k \ge 0}$ converges to the solution u_h, as k tends to $+\infty$. By using (36), (34) and that u_h is the solution of problem (21) we have

$$\|u_h^{(k)} - u_h\|_{1,h}^2 \lesssim \langle E_h'(u_h^{(k)}) - E_h'(u_h), u_h^{(k)} - u_h \rangle$$
$$\leq \langle E_h'(u_h^{(k)}), u_h^{(k)} - u_h \rangle + F_h(u_h - u_h^{(k)})$$
$$= b_h(u_h^{(k)}; u_h^{(k)}, u_h^{(k)} - u_h) + F_h(u_h - u_h^{(k)}). \qquad (39)$$

Then, by using twice (32) we get

$$b_h(u_h^{(k)}; u_h^{(k)}, u_h^{(k)} - u_h) + F_h(u_h - u_h^{(k)})$$
$$= b_h(u_h^{(k)}; u_h^{(k)} - u_h^{(k+1)}, u_h^{(k)} - u_h) \pm b_h(u_h^{(k)}; u_h^{(k+1)}, u_h^{(k+1)} - u_h^{(k)})$$
$$= b_h(u_h^{(k)}; u_h^{(k)} - u_h^{(k+1)}, u_h^{(k)} - u_h + u_h^{(k+1)}) + F_h(u_h^{(k+1)} - u_h^{(k)}).$$

Finally, by inserting the above estimate into (39), using (26) and that $|F_h(v_h)| \lesssim \|f\|_{L^2(\Omega)} \|v_h\|_{1,h}$ we get

$$\|u_h^{(k)} - u_h\|_{1,h}^2 \lesssim \|u_h^{(k+1)} - u_h^{(k)}\|_{1,h} \left(\|u_h^{(k)} - u_h + u_h^{(k+1)}\|_{1,h} + \|f\|_{L^2(\Omega)} \right).$$
$$(40)$$

The proof is complete by observing that the quantity $\|u_h^{(k)} - u_h + u_h^{(k+1)}\|_{1,h}$ is bounded from above since each term is bounded thanks to the well-posedness of problem (21) and of the corresponding linearized ones (see Proposition 7). Therefore, thanks to (38) we get $\|u_h^{(k)} - u_h\| \to 0$ as k tends to ∞.

The following result, stemming from the approach described in [35], contains a computable estimate for the error $\|u_h - u_h^{(k)}\|_{1,h}$, which can be employed as a reliable stopping criterion for the Kačanov method (see also Sect. 6 for further details).

Proposition 4 *Assume the function $\kappa(\cdot)$ satisfies Assumption 21 and also $\kappa'(t) \leq 0$ and $2\kappa'(t)t + \kappa(t) \gtrsim 1$. Then*

$$\|u_h - u_h^{(k)}\|_{1,h}^2 \lesssim E_h(u_h^{(k)}) - F_h(u_h^{(k)}) + E_h^*(p_h^*)$$

where $p_{h|E}^* := \kappa(\mathcal{G}_h^E(u_h^{(k-1)}))(\mathbb{A}_h^E)^{1/2} u_h^{(k)}{}_{|E}$, and the conjugate functional $E_h^*(\cdot)$ is defined as

$$E_h^*(p_h^*) := \sum_{E \in \Omega_h} |E| \beta^* \left(\sqrt{\mathcal{G}_h^E(p_h^*)} \right).$$

Here, $\beta^*(s^*) := \kappa(\bar{s}^2)\bar{s}^2 - \beta(\bar{s})$ for every $s^* \in \mathbb{R}$ being $\bar{s} = \bar{s}(s^*)$ the unique solution to the algebraic equation

$$\beta'(\bar{s}) = s^*.$$

Proof Setting, as before, $\mathcal{L}_h(\cdot) := E_h(\cdot) - F_h(\cdot)$, we first observe that it holds

$$\mathcal{D}_h(u_h^{(k)}) := \mathcal{L}_h(u_h^{(k)}) - \mathcal{L}_h(u_h) - \langle \mathcal{L}_h'(u_h), u_h^{(k)} - u_h \rangle$$

$$= \int_0^1 (1-t) \langle \mathcal{L}_h''(u_h + t(u_h^{(k)} - u_h))(u_h^{(k)} - u_h), u_h^{(k)} - u_h \rangle \, \mathrm{d}t, \quad (41)$$

where the second order Fréchet derivative satisfies

$$\langle \mathcal{L}_h''(u_h)v_h, v_h \rangle = \sum_E 2\kappa' \left(\frac{\|u_h\|_{\mathbb{A}_h^E}^2}{|E|} \right) \frac{(a_h^E(u_h, v_h))^2}{|E|} + \kappa \left(\frac{\|u_h\|_{\mathbb{A}_h^E}^2}{|E|} \right) a_h^E(v_h, v_h)$$

$$\gtrsim \sum_E \left(2\kappa' \left(\frac{\|u_h\|_{\mathbb{A}_h^E}^2}{|E|} \right) \frac{\|u_h\|_{1,h,E}^2}{|E|} + \kappa \left(\frac{\|u_h\|_{\mathbb{A}_h^E}^2}{|E|} \right) \right) \|v_h\|_{1,h,E}^2$$

$$\gtrsim \|v_h\|_{1,h}^2, \quad (42)$$

where we used Proposition 2 together with the fact that $\kappa'(z) \leq 0$ and $2\kappa'(t)t + \kappa(t) \gtrsim 1$. Observing that u_h is the unique solution to (21) yields $\langle \mathcal{L}_h'(u_h), v_h \rangle = 0$ for every $v_h \in V_h$. Hence, combining (41) and (42) we get

$$\mathcal{L}_h(u_h^{(k)}) - \mathcal{L}_h(u_h) = \mathcal{D}_h(u_h^{(k)}) \gtrsim \|u_h^{(k)} - u_h\|_{1,h}^2. \quad (43)$$

Using the standard theory of duality in optimization (see, e.g., [31, Ch. 3],[46, Sec. 51.5] or [35, Sec. 2.7] for a general introduction) and following the approach outlined in [35, Sec. 5.3] (see also [46, Theorem 51.B] and [47, Theorem 25.K] for similar results), we obtain

$$\mathcal{L}_h(u_h^{(k)}) - \mathcal{L}_h(u_h) \leq \mathcal{L}_h(u_h^{(k)}) + E_h^*(p_h^*), \quad (44)$$

where it is crucial to observe that p_h^* has been chosen to solve the linear elliptic problem $p_h^T \mathbb{A}_h^{1/2} v_h = F(v_h)$ for every $v_h \in V_h$. A closer inspection reveals that p_h^* is indeed the solution of the k-th linearized problem (32). Finally, combining (43) and (44) yields the thesis.

Remark 4 It is easy to see that the Carreau law defined in (4) satisfies the assumptions on the function $\kappa(\cdot)$ of Proposition 4.

We conclude this section by observing that Theorems 1 and 2 can be combined to obtain an estimate of the error $\|u_{\mathrm{I}} - u_h^{(k)}\|_{1,h}$.

Corollary 2 *Let $u_{\mathrm{I}} \in V_h$ be the interpolation of the exact solution of problem (3) as defined in (14). Then, for every $h > 0$ there exists $k \in \mathbb{N}$, $k = k(h)$, such that*

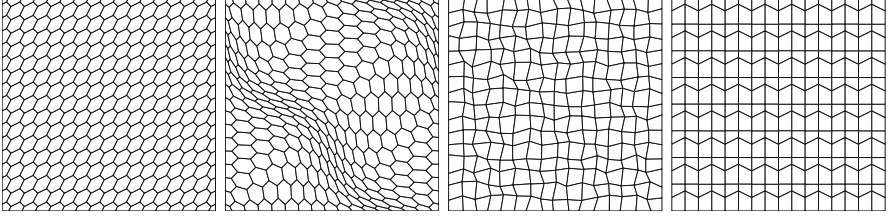$$\|u_{\mathrm{I}} - u_h^{(k)}\|_{1,h} \lesssim O(h) + \|u_h - u_h^{(k)}\|_{1,h},$$

Fig. 1 Examples of the considered decompositions of $\Omega = (0, 1)^2$. From left to right: *hexagons type1*, *hexagons type2*, *quadrilaterals* and *trapezes*

where $u_h^{(k)} \in V_h$ is the k-th iterate obtained with the Kačanov method (32).

6 Numerical experiments

This section is devoted to investigate the performance of the MFD method for the approximate solution of problem (3) and to validate the convergence analysis stated in Theorem 2.

Throughout the section, the function $\kappa(\cdot)$ is chosen equal to the Carreau law defined in (4), with $\eta_0 = 3$, $\eta_\infty = 1$ and $p = 1.7$. We consider four different sequences of decompositions of the domain $\Omega = (0, 1)^2$. Figure 1 shows an example of all the considered computational grids, that we denote by *hexagons type1*, *hexagons type2*, *quadrilaterals* and *trapezes*, respectively.

To solve problem (21) we have employed the *feasible* Kačanov method supplemented with a suitable stopping criterion, as shown in Algorithm 6.1. The tolerance in the stopping criterion has been set equal to 10^{-8}. We remark that, the reliability of the stopping criterion is justified in view of inequality (40). Alternatively, one can resort to the computable estimate presented in Proposition 4.

Algorithm 6.1: *Feasible* Kačanov method

---

1 Given the initial guess $u_h^{(0)}$, set `toll`, $k = -1$, $u_h^{(-1)} = u_h^{(0)}$;

2 **while** $\|u_h^{(k+1)} - u_h^{(k)}\|_{1,h} \geq$ `toll` **do**

3      $k + 1 \leftarrow k$;

4   SOLVE $b_h(u_h^{(k)}; u_h^{(k+1)}, v_h) = F_h(v_h) \quad \forall v_h \in V_h$;

5 **end**

6 SET $u_h := u_h^{(k+1)}$;

6.1 Example 1

In the first example we choose $u(x, y) = \log(1+x+y)+x+y$ as the exact solution and set the data and the non-homogeneous boundary conditions accordingly. The relative errors in the discrete norm (13) computed on all the sequences of the considered grids are reported in Fig. 2a (loglog scale). We clearly observe a linear convergence as predicted by our a-priori error estimates stated in Theorem 2.
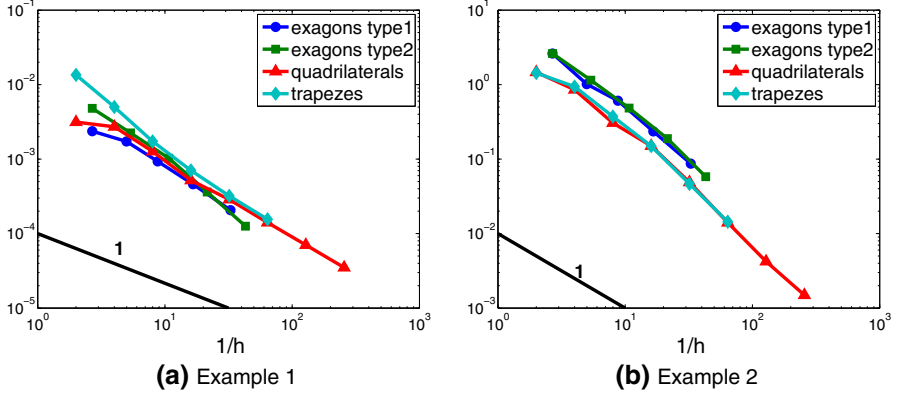
Fig. 2 Computed relative errors $\|u_{\mathtt{I}} - u_h\|_{1,h}/\|u_{\mathtt{I}}\|_{1,h}$ versus $1/h$ (loglog scale). Example 1 (*left*) and Example 2 (*right*)

6.2 Example 2

The second example is taken from [37]. Here, the source term f is selected so that $u(x, y) = x(1 - x)y(1 - y)(1 - 2y)\exp(-20((2x - 1)^2))$ is the analytical solution of problem (21). The numerical results are reported in Fig. 2b, where the computed relative errors in the discrete norm defined in (13) are shown (loglog scale). Also in this case, we observe that the error goes to zero linearly as the mesh is refined.

6.3 On Assumption (28)

Finally, we check numerically the validity of hypothesis (28) in Theorem 1. On the sequence of considered grids, we have computed the quantity

$$\left\| \kappa\left(\Pi^0 \, |\nabla u|^2\right) - \kappa(\mathcal{G}_h^E(u_{\mathtt{I}})) \right\|_{\infty,h},$$

where Π^0 denotes the projection onto the space of piecewise constant functions defined on Ω_h, $u_{\mathtt{I}}$ is the interpolation of the exact solutions of Examples 6.1 and 6.2, respectively, and

$$\|v_h\|_{\infty,h} := \sup_{\mathsf{v} \in \mathcal{N}_h} |v_h^{\mathsf{v}}| \forall v_h \in V_h.$$

The computed results are reported in Fig. 3. From the numerical results it seems that

$$\|\kappa\left(\Pi^0 \, |\nabla u|^2\right) - \kappa(\mathcal{G}_h^E(u_{\mathtt{I}}))\|_{\infty,h} \lesssim h.$$

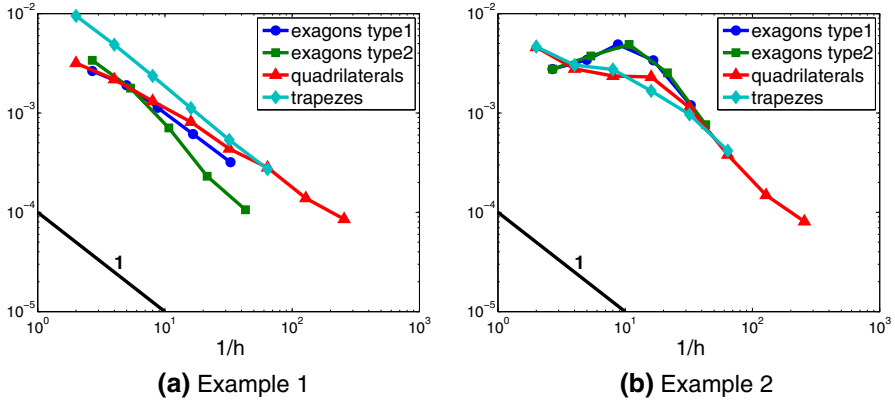This, together with standard interpolation error estimates suggests the validity of (28) with $\alpha = 1$.

Fig. 3 $\left\| \kappa \left(\Pi^0 \, |\nabla u|^2 \right) - \kappa (\mathcal{G}_h^E(u_{\mathcal{I}})) \right\|_{\infty,h}$ versus $1/h$ (loglog scale). Here u is the exact solution of Example 1 (left) and Example 2 (right)

7 Conclusions

We have proposed a MFD discretization of a quasilinear elliptic problem. Under a suitable approximation assumption, we have shown optimal a-priori error estimates in a mesh-dependent energy norm. A Kačanov iterative scheme has been proposed and analyzed to solve the resulting nonlinear discrete problem. All the theoretical results have been confirmed by numerical experiments. Future research will be devoted to extend the presented results to (free-boundary) quasi-newtonian flow problems governed by the non-linear Stokes equations by taking advantage of the Uzawa iterative algorithm, where at each step a quasilinear elliptic problem similar to (1) has to be solved.

References

1. Andreianov, B., Boyer, F., Hubert, F.: Discrete duality finite volume schemes for Leray-Lions-type elliptic problems on general 2D meshes. Numer. Methods Partial Differ. Equ. **23**(1), 145–195 (2007)
2. Antonietti, P.F., Beirão da Veiga, L., Lovadina, C., Verani, M.: Hiercarchical a posteriori error estimators for the mimetic discretization of elliptic problems. SIAM J. Numer. Anal. **51**, 654–675 (2013)
3. Antonietti, P.F., Beirão da Veiga, L., Verani, M.: A mimetic discretization of elliptic obstacle problems. Math. Comp. **82**(283), 1379–1400 (2013)
4. Antonietti, P.F., Bigoni, N., Verani, M.: Mimetic discretizations of elliptic control problems. J. Sci. Comput. **56**(1), 14–27 (2013)
5. Antonietti, P.F., Fadel, N.A., Verani, M.: Modelling and numerical simulation of the polymeric extrusion process in textile products. Commun. Appl. Ind. Math. **1**(2), 1–13 (2010)
6. Bao, W.: An economical finite element approximation of generalized Newtonian flows. Comput. Methods Appl. Mech. Eng. **191**(33), 3637–3648 (2002)
7. Bao, W., Barrett, J.W.: A priori and a posteriori error bounds for a nonconforming linear finite element approximation of a non-Newtonian flow. RAIRO Modél. Math. Anal. Numér. **32**(7), 843–858 (1998)
8. Baranger, J., El Amri, H.: Estimateurs a posteriori d'erreur pour le calcul adaptatif d'écoulements quasi-newtonians. RAIRO Modél. Math. Anal. Numér. **25**(1), 31–47 (1991)
9. Baranger, J., Najib, K.: Analyse numérique des écoulements quasi-newtoniens dont la viscosité obéit à la loi puissance ou la loi de carreau. Numer. Math. **58**(1), 35–49 (1990)

10. Barrett, J.W., Liu, W.B.: Finite element error analysis of a quasi-newtonian flow obeying the Carreau or power law. Numer. Math. **64**, 433–453 (1993)
11. Beirão Da Veiga, L.: A mimetic discretization method for linear elasticity. M2AN Math. Model. Numer. Anal. **44**(2), 231–250 (2010)
12. Beirão da Veiga, L., Brezzi, F., Cangiani, A., Manzini, G., Marini, L.D., Russo, A.: Basic principles of virtual element methods. Math. Models Methods Appl. Sci. **23**(1), 199–214 (2013)
13. Beirão da Veiga, L., Gyrya, V., Lipnikov, K., Manzini, G.: Mimetic finite difference method for the Stokes problem on polygonal meshes. J. Comput. Phys. **228**(19), 7215–7232 (2009)
14. Beirão da Veiga, L., Lipnikov, K.: A mimetic discretization of the Stokes problem with selected edge bubbles. SIAM J. Sci. Comput. **32**(2), 875–893 (2010)
15. Beirão da Veiga, L., Lipnikov, K., Manzini, G.: Convergence analysis of the high-order mimetic finite difference method. Numer. Math. **113**(3), 325–356 (2009)
16. Beirão da Veiga, L., Lipnikov, K., Manzini, G.: Arbitrary-order nodal mimetic discretizations of elliptic problems on polygonal meshes. SIAM J. Numer. Anal. **49**(5), 1737–1760 (2011)
17. Beirão da Veiga, L., Manzini, G.: An a posteriori error estimator for the mimetic finite difference approximation of elliptic problems. Int. J. Numer. Methods Eng. **76**(11), 1696–1723 (2008)
18. Berrone, S., Süli, E.: Two-sided a posteriori error bounds for incompressible quasi-Newtonian flows. IMA J. Numer. Anal. **28**(2), 382–421 (2008)
19. Boyer, F., Hubert, F.: Finite volume method for 2D linear and nonlinear elliptic problems with discontinuities. SIAM J. Numer. Anal. **46**(6), 3032–3070 (2008)
20. Brenner, S., Scott, L.: The mathematical theory of finite element methods. Springer, Berlin/Heidelberg (1994)
21. Brezzi, F., Buffa, A., Lipnikov, K.: Mimetic finite differences for elliptic problems. M2AN Math. Model. Numer. Anal. **43**(2), 277–295 (2009)
22. Brezzi, F., Lipnikov, K., Shashkov, M.: Convergence of the mimetic finite difference method for diffusion problems on polyhedral meshes. SIAM J. Numer. Anal. **43**(5), 1872–1896 (2005). (electronic)
23. Brezzi, F., Lipnikov, K., Shashkov, M.: Convergence of mimetic finite difference method for diffusion problems on polyhedral meshes with curved faces. Math. Models Methods Appl. Sci. **16**(2), 275–297 (2006)
24. Brezzi, F., Lipnikov, K., Simoncini, V.: A family of mimetic finite difference methods on polygonal and polyhedral meshes. Math. Models Methods Appl. Sci. **15**(10), 1533–1551 (2005)
25. Brezzi, F., Marini, L.D.: Virtual element methods for plate bending problems. Comput. Methods Appl. Mech. Eng. **253**, 455–462 (2013)
26. Cangiani, A., Gardini, F., Manzini, G.: Convergence of the mimetic finite difference method for eigenvalue problems in mixed form. Comput. Methods Appl. Mech. Eng. **200**(9–12), 1150–1160 (2011)
27. Chow, S.-S.: Finite element error estimates for nonlinear elliptic equations of monotone type. Numer. Math. **54**(4), 373–393 (1989)
28. Congreve, S., Houston, P., Suli, E., Wihler, T.P.: Discontinuous Galerkin finite element approximation of quasilinear elliptic boundary value problems ii: strongly monotone quasi-newtonian flows. Technical Report NA-11/20. The University of Notthigham (2011)
29. da Veiga, L.B., Brezzi, F., Marini, L.D.: Virtual elements for linear elasticity problems. SIAM J. Numer. Anal. **51**(2), 794–812 (2013)
30. Diening, L., Kreuzer, C.: Linear convergence of an adaptive finite element method for the p-Laplacian equation. SIAM J. Numer. Anal. **46**(2), 614–638 (2008)
31. Ekeland, I., Temam, R.: Convex analysis and variational problems. North-Holland Publishing Co., Amsterdam (1976). (Translated from the French, Studies in Mathematics and its Applications, Vol. 1)
32. El Alaoui, L., Ern, A., Vohralík, M.: Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems. Comput. Methods Appl. Mech. Eng. **200**(37–40), 2782–2795 (2011)
33. Evans, L.C.: Partial differential equations, volume 19 of graduate studies in mathematics, 2nd edn. American Mathematical Society, Providence (2010)
34. Garau, E.M., Morin, P., Zuppa, C.: Convergence of an adaptive Kačanov FEM for quasi-linear problems. Appl. Numer. Math. **61**(4), 512–529 (2011)
35. Han, W.: A posteriori error analysis via duality theory, volume 8 of advances in mechanics and mathematics. Springer, New York (2005)
36. Houston, P., Robson, J., Süli, E.: Discontinuous Galerkin finite element approximation of quasilinear elliptic boundary value problems. I. The scalar case. IMA J. Numer. Anal. **25**(4), 726–749 (2005)

37. Houston, P., Süli, E., Wihler, T.P.: A posteriori error analysis of hp-version discontinuous Galerkin finite-element methods for second-order quasi-linear elliptic PDEs. IMA J. Numer. Anal. **28**(2), 245–273 (2008)

38. Hyman, J., Shashkov, M.: Mimetic discretizations for Maxwell's equations and the equations of magnetic diffusion. Prog. Electromagn. Res. **32**, 89–121 (2001)

39. Kreuzer, C.: Analysis of an adaptive Uzawa finite element method for the nonlinear Stokes problem. Math. Comp. **81**(277), 21–55 (2012)

40. Lipnikov, K., Manzini, G., Brezzi, F., Buffa, A.: The mimetic finite difference method for the 3D magnetostatic field problems on polyhedral meshes. J. Comput. Phys. **230**(2), 305–328 (2011)

41. Lipnikov, K., Morel, J., Shashkov, M.: Mimetic finite difference methods for diffusion equations on non-orthogonal non-conformal meshes. J. Comput. Phys. **199**(2), 589–597 (2004)

42. Lipnikov, K., Moulton, J., Svyatskiy, D.: A multilevel multiscale mimetic (M^3) method for two-phase flows in porous media. J. Comp. Phys. **227**, 6727–6753 (2008)

43. Mitsoulis, E.: Three-dimensional non-Newtonian computations of extrudate swell with the finite element method. Comput. Methods Appl. Mech. Eng. **180**(34), 333–344 (1999)

44. Padra, C.: A posteriori error estimators for nonconforming approximation of some quasi-Newtonian flows. SIAM J. Numer. Anal. **34**(4), 1600–1615 (1997)

45. Veeser, A.: Convergent adaptive finite elements for the nonlinear Laplacian. Numer. Math. **92**(4), 743–770 (2002)

46. Zeidler, E.: Nonlinear functional analysis and its applications. III. Variational methods and optimization. Springer, New York (1985). (Translated from the German by Leo F, Boron)

47. Zeidler, E.: Nonlinear functional analysis and its applications. II, B. Nonlinear monotone operators. Springer, New York (1990). (Translated from the German by the author and Leo F, Boron)