A short note on an adaptive damped Newton method for strongly monotone and Lipschitz continuous operator equations

PASCAL HEID

Abstract. We consider the damped Newton method for strongly monotone and Lipschitz continuous operator equations in a variational setting. We provide a very accessible justification why the undamped Newton method performs better than its damped counterparts in a vicinity of a solution. Moreover, in the given setting, an adaptive step-size strategy be presented, which guarantees the global convergence and favours an undamped update if admissible.

Mathematics Subject Classification. 65J15, 49M15, 65M12.

Keywords. Strongly monotone problems, Damped Newton method, Adaptive step-size strategy.

1. Introduction. In these notes, we focus on the damped Newton method for strongly monotone and Lipschitz continuous operator equations on a real-valued Hilbert space $X$ with inner product $(\cdot, \cdot)_X$ and induced norm $\|\cdot\|_X$. Specifically, given a nonlinear operator $F : X \to X'$, we focus on the equation

$$u \in X : \quad F(u) = 0 \quad \text{in } X';$$

here, $X'$ denotes the dual space of $X$. In weak form, this problem is given by

$$u \in X : \quad \langle F(u), v \rangle = 0 \quad \text{for all } v \in X,$$

where $\langle \cdot, \cdot \rangle$ signifies the duality pairing in $X' \times X$. Throughout this work, we impose the following structural assumptions on the nonlinear operator $F : X \to X'$:

(F1) The operator $F$ is Lipschitz continuous; i.e., there exists a constant $L > 0$ such that

$$|\langle F(u) - F(v), w \rangle| \leq L \|u - v\|_X \|w\|_X \quad \text{for all } u, v, w \in X;$$
The operator $F$ is strongly monotone; i.e., there exists a constant $\nu > 0$ such that
\[ \langle F(u) - F(v), u - v \rangle \geq \nu \|u - v\|^2_X \quad \text{for all } u, v \in X. \] (4)
Under those conditions, it is well-known that the operator equation (1) has a unique solution $u^* \in X$; see, e.g., [11, Section 3.3] or [13, Section 25.4]. However, since $F$ is assumed to be nonlinear, it is in general not feasible to solve for $u^* \in X$. As a remedy, one may apply an iteration scheme to obtain an approximation $u^n$ of $u^*$. The most famous method to approximate iteratively a root of a nonlinear operator $F$ is the Newton method, which is defined as follows: For a given initial guess $u^0 \in X$, we define recursively
\[ u^{n+1} = u^n - F'(u^n)^{-1}F(u^n), \quad n \in \mathbb{N}, \] (5)
where $F'(u^n) : X \to X'$ denotes the Gateaux derivative of $F$ at $u^n \in X$.

The main advantage of the Newton method over other iterative linearisation schemes is the local quadratic convergence rate. However, in many cases, the Newton scheme requires an initial guess close to a solution of the operator equation. To improve the lack of global convergence, one may introduce a damping parameter. Then, the damped Newton method is given by
\[ u^{n+1} = u^n - \delta(u^n)F'(u^n)^{-1}F(u^n), \quad n \in \mathbb{N}, \] (6)
where $\delta(u^n) > 0$ is a damping parameter, which may depend on the given iterate. Often, it is not clear how to choose the damping parameter optimally. Moreover, choices that guarantee the global convergence a priori may often entail an inferior convergence rate compared to the undamped scheme (5). For an extensive overview of Newton’s method for nonlinear problems, we refer the interested reader to [4], especially [4, Section 3] for its damped counterpart, and the references therein. Further and more recent works on the development of adaptive step-size strategies for the damped Newton method may include [1,2,12].

The motivation of this work is to present an adaptive damped Newton method, which guarantees the global convergence, and, in a vicinity of a solution, exhibits the optimal quadratic convergence rate. By no means at all do we claim to present a state-of-the-art nonlinear solver for the problem considered herein. However, those superior solvers and adaptive step-size strategies are much more involved, whereas the presented scheme is easily accessible and fairly simple to implement. Finally, the algorithm presented herein can also be applied efficiently in combination with an adaptive finite element method. Indeed, the presented nonlinear solver fits into the framework of [6], and even with the adaptive damping strategy yields the optimal convergence rate with respect to the degrees of freedom and the overall computational costs; see [6, Thm. 4.3]

**Contribution.** The novelty of the present manuscript is to combine the convergence analysis of the Newton method from [8, Thm. 2.6] (see also [10, Cor. 2.7] for a closely related result) and the step-size strategies from [10, Section 3],
which were developed in the context of the Kačanov iteration scheme. In particular, transferring the step-size strategy from [10, Section 3.1] to the setting of the damped Newton scheme leads to a simple justification why the classical Newton method is optimal in a neighbourhood of a solution. On the other hand, a variant of the adaptive step-size strategy from [10, Section 3.2] yields the global convergence combined with the local quadratic convergence rate, provided that an undamped step is admissible close to the solution.

2. Convergence of the damped Newton method. In many problems of scientific interest, $X$ is an infinite dimensional space, but numerical computations are carried out on finite dimensional subspaces, e.g., finite element spaces. We note, however, that our analysis equally applies to closed subspaces of $X$—infinite and finite dimensional ones. In order to guarantee the global convergence of the damped Newton method, we need to impose some further assumptions on $F$:

(F3) The operator $F$ is Gateaux differentiable. Moreover, $F'$ is uniformly coercive and bounded in the sense that, for any given $u \in X$,

$$
\langle F'(u)v, v \rangle \geq \alpha_{F'} \|v\|_X^2 \quad \text{for all } v \in X
$$

and

$$
\langle F'(u)v, w \rangle \leq \beta_{F'} \|v\|_X \|w\|_X \quad \text{for all } v, w \in X,
$$

where $\alpha_{F'}, \beta_{F'} > 0$ are independent of $u$;

(F4) There exists a Gateaux differentiable functional $G : X \to \mathbb{R}$ such that $G'(u) = F'(u)u$ in $X'$ for any $u \in X$, and $G' : X \to X'$ is continuous with respect to the weak topology in $X'$;

(F5) $F$ is a potential operator; i.e., there exists a Gateaux differentiable functional $H : X \to \mathbb{R}$ with $H' = F$.

Then, the following convergence result holds.

**Theorem 2.1** ([8, Thm. 2.6]). Assume (F1)–(F5) and that $\delta : X \to [\delta_{\min}, \delta_{\max}]$ for some $0 < \delta_{\min} \leq \delta_{\max} < 2\alpha_{F'}/L$. Then, the damped Newton method (6) converges to the unique solution $u^* \in X$ of (1).

In many cases, we have that $2\alpha_{F'}/L < 1$. This, however, excludes the choice $\delta(u^n) = 1$, which yields the local quadratic convergence rate. By examining the proof of [8, Thm. 2.6], one realises that the specific upper bound on the damping parameter was imposed to guarantee the existence of a constant $C_H > 0$ such that

$$
H(u^n) - H(u^{n+1}) \geq C_H \|u^n - u^{n+1}\|_X^2 \quad \text{for all } n \in \mathbb{N};
$$

we also refer to the closely related work [6], especially Eq. (2.9) in that reference. Indeed, in the proof of [8, Thm. 2.6], it is shown that

$$
H(u^n) - H(u^{n+1}) \geq \left( \frac{\alpha_{F'}}{\delta(u^n)} - \frac{L}{2} \right) \|u^n - u^{n+1}\|_X^2.
$$
Consequently, given that \( \delta(u^n) \leq \delta_{\text{max}} < 2\alpha_F/L \), we have

\[
H(u^n) - H(u^{n+1}) \geq \left( \frac{\alpha_F}{\delta_{\text{max}}} - \frac{L}{2} \right) \|u^n - u^{n+1}\|_X^2 = C_H \|u^n - u^{n+1}\|_X^2, \tag{11}
\]

where \( C_H := \left( \frac{\alpha_F}{\delta_{\text{max}}} - \frac{L}{2} \right) > 0 \). By no means, however, is \( \delta(u^n) \leq \delta_{\text{max}} < 2\alpha_F/L \) a necessary condition for (9). The following generalised convergence theorem remains true, which can be easily verified by following along the lines of the proof of [8, Thm. 2.6].

**Theorem 2.2.** Given (F1)–(F5), for any damping strategy that guarantees \( \delta(u^n) \in [\delta_{\text{min}}, \delta_{\text{max}}] \) and (9) for all \( n \in \mathbb{N} \), where \( 0 < \delta_{\text{min}} \leq \delta_{\text{max}} < \infty \) and \( C_H > 0 \) are given constants independent of \( n \), we have that \( u^n \to u^* \) strongly in \( X \) as \( n \to \infty \).

### 3. Local optimality of the classical Newton scheme.

Before we introduce our step-size strategy, which is borrowed from [10, Section 3.1], we first remark that the unique solution of the operator equation (1) is equally the unique minimiser of the potential \( H \), see [13, Thm. 25.F]. Furthermore, the damped Newton method can be reformulated as

\[
u^{n+1} = u^n - \delta^n \rho^n, \tag{12}\]

where \( \delta^n \) is the damping parameter and \( \rho^n = F'(u^n)^{-1}F(u^n) \) is the undamped update at the given iterate. Equivalently, we have that

\[
u^{n+1} - u^n = -\delta^n \rho^n. \tag{13}\]

Since we want to obtain the unique minimiser \( u^* \in X \) of \( H \), a sensible strategy is to choose the step-size \( \delta^n \) in (12) in such a way that, for given \( u^n \in X \), the difference

\[
H(u^{n+1}) - H(u^n) \tag{14}
\]

is minimised. Indeed, this corresponds to a maximal decay of the potential at step \( n+1 \). We now follow along the lines of [10, Section 3.1], where the strategy was originally developed, to the best of our knowledge, as a possible choice of the damping parameter of a modified Kačanov scheme. Let us recall that \( H \) is the potential of \( F \), and thus \( H' = F \). Consequently, thanks to the fundamental theorem of calculus, we have that

\[
H(u^{n+1}) - H(u^n) = \int_0^1 \langle F(u^n + t(u^{n+1} - u^n)), u^{n+1} - u^n \rangle \, dt. \tag{15}\]

Let \( \psi(t) := \langle F(u^n + t(u^{n+1} - u^n)), u^{n+1} - u^n \rangle \) be the integrand of the right-hand side above. In view of (13) and the linearity of the dual product, the integrand \( \psi \) can be rewritten as

\[
\psi(t) = -\delta^n \langle F(u^n - t\delta^n \rho^n), \rho^n \rangle. \tag{16}\]

In order to minimise (15), we employ a first order Taylor approximation to (16) at \( t = 0 \). Specifically, a straightforward calculation reveals that

\[
\psi(t) \approx -\delta^n \langle F(u^n), \rho^n \rangle + t(\delta^n)^2 \langle F'(u^n)\rho^n, \rho^n \rangle. \tag{17}\]
Plugging (17) into (15) and, subsequently, integrating from zero to one yields
\[
H(u^{n+1}) - H(u^n) \approx \frac{1}{2}(\delta_n)^2 \langle F'(u^n)\rho^n, \rho^n \rangle - \delta_n \langle F(u^n), \rho^n \rangle. \tag{18}
\]
Then, since \(F'(u^n)\) is coercive, it follows immediately that the right-hand side of (18) is minimised for \(\delta_n = \frac{\langle F(u^n), \rho^n \rangle}{\langle F'(u^n)\rho^n, \rho^n \rangle} = 1; \tag{19}\)
here, we used that \(F'(u^n)\rho^n = F'(u^n)F(u^n)^{-1}F(u^n) = F(u^n)\). Furthermore, the approximations (17) and (18), respectively, become more accurate the smaller \(\|\rho^n\|_X\) gets, and thus especially in a neighbourhood of a solution. In particular, this means that \(\delta \equiv 1\) is indeed the optimal local damping parameter in a vicinity of the solution.

4. Adaptive damped Newton algorithm. Now we shall present our adaptive damped Newton method, cf. Algorithm 1, which guarantees the global convergence and favours the step-size \(\delta_n = 1\), if admissible, leading to local quadratic convergence rate. This algorithm is closely related to [10, Algorithm 1].

---

**Algorithm 1 Adaptive damped Newton method**

**Input:** Initial guess \(u^0 \in X\), correction factor \(\sigma \in (0, 1)\), and \(\theta \in (0, 0.5]\).

1: for \(n = 0, 1, 2, \ldots\) do
2: Solve the linear problem \(F'(u^n)\rho^n = F(u^n)\) for \(\rho^n \in X\).
3: Set \(\delta_n = 1\).
4: repeat
5: Compute \(u^{n+1} = u^n - \delta_n \rho^n\).
6: Update \(\delta_n \leftarrow \max\{\sigma\delta_n, \alpha\rho/L\}\).
7: until \(H(u^n) - H(u^{n+1}) \geq \theta \min\{\alpha, L\} \|u^n - u^{n+1}\|_X^2\).
8: end for

---

**Theorem 4.1.** Given the assumptions (F1)–(F5), the sequence \(\{u^n\}\) generated by Algorithm 1 converges strongly to the unique solution \(u^* \in X\) of the operator equation (1).

**Proof.** Let \(C_H := \theta \min\{\alpha, L\}\). In the following, we shall distinguish two cases:

(I) If \(\alpha/L > 1\), then the algorithm accepts, for every \(n \in \mathbb{N}\), the initial choice \(\delta_n = 1\). Indeed, we have that \((\alpha/L - L/2) \geq L/2 \geq \theta \min\{\alpha, L\}\), wherefore the statement in line 7 of Algorithm 1 is satisfied, thanks to (10), without any correction. In turn, the convergence follows from Theorem 2.1.

(II) Now let us presume that \(0 < \delta_{\min} := \alpha/L \leq 1 =: \delta_{\max}\). Thanks to Theorem 2.2, it only remains to prove (9). By line 7 of Algorithm 1, this amounts to verify that the repeat-loop (lines 4–7) in Algorithm 1 terminates for each \(n \in \mathbb{N}\). By contradiction, assume that this was not
the case; i.e., there exists \( n \in \mathbb{N} \) such that the statement in line 7 is never satisfied. For this specific \( n \in \mathbb{N} \), we have after finitely many steps that \( \delta^n = \alpha_F / L \). Recalling (10), we find that

\[
H(u^n) - H(u^{n+1}) \geq \left( \frac{\alpha_F}{\delta^n} - \frac{L}{2} \right) \| u^n - u^{n+1} \|^2_X
= \frac{L}{2} \| u^n - u^{n+1} \|^2_X
\geq C_H \| u^n - u^{n+1} \|^2_X;
\]

i.e., the repeat-loop terminates, which yields the desired contradiction.

\( \square \)

**Remark 4.2.** Of course, Algorithm 1 could be further improved by taking into account the accepted damping parameter from the previous step. However, for simplicity of the presentation, we stick to our current algorithm.

**Remark 4.3.** By modus operandi of Algorithm 1, we have that

\[
H(u^n) - H(u^{n+1}) \geq C_H \| u^n - u^{n+1} \|^2_X
\]

for all \( n \in \mathbb{N} \), where \( C_H := \theta \min\{\alpha_F, L\} > 0 \), and thus the proposed scheme fits into the framework of the iterative nonlinear solver of [6, Section 2]. Moreover, recalling the proof of Theorem 4.1, it is apparent that the number of repeat-loops (lines 4–7) in Algorithm 1 is uniformly bounded for \( n \in \mathbb{N} \). Consequently, we may replace the single iterative linearisation step in [6, Algorithm 1 (line 4)] with our nonlinear solver from Algorithm 1, and the convergence results of [6, Section 4] still remain valid. Most importantly, our suggested iterative nonlinear solver in combination with an adaptive finite element method may lead to a computational scheme, which converges to a solution of the original problem (1) with optimal rate with respect to the degrees of freedom as well as the overall computational time.

**Remark 4.4.** It shall be emphasised that, for the problem considered herein, there exist superior numerical schemes as well as more elaborated adaptive step-size strategies from nonlinear optimisation, such as the Armijo or the Wolfe-Powell rule; we refer the interested reader, e.g., to [5]. However, as mentioned at the beginning, the goal of this work is to present an easily accessible damped Newton method with guaranteed convergence and local quadratic convergence rate, which in addition is fairly simple to implement.

### 5. Application to quasilinear elliptic diffusion models.

In the following, let \( X := H^1_0(\Omega) \) denote the standard Sobolev space of \( H^1 \)-functions on \( \Omega \) with zero trace along the boundary \( \Gamma := \partial \Omega \), where \( \Omega \subset \mathbb{R}^d \), \( d \in \mathbb{N} \), is an open and bounded domain with Lipschitz boundary. The inner product and norm on \( X \) are defined by \((u,v)_X := \int_{\Omega} \nabla u \cdot \nabla v \ dx \) and \( \| u \|_X^2 = \int_{\Omega} |\nabla u|^2 \ dx \), respectively.

We consider the quasilinear elliptic partial differential equation

\[
u \in X : \quad F(u) := -\nabla \cdot \left\{ \mu \left( |\nabla u|^2 \right) \nabla u \right\} - g = 0 \quad \text{in } X'; \quad (20)
\]
here, $\mu : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a diffusion coefficient and $g \in H^{-1}(\Omega) = X'$ a given source function. Models of the form (20) are widely applied in physics, for instance in plasticity and elasticity, as well as in hydro- and gas-dynamics. Moreover, they also served as our model problems in our previous and closely related works [6,8–10].

To guarantee that the properties (F1)–(F5) are satisfied in the given setting, we need to impose the following assumptions on the diffusion coefficient:

(M1) The diffusion coefficient $\mu : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is continuously differentiable and monotonically decreasing; i.e., $\mu'(t) \leq 0$ for all $t \geq 0$.

(M2) There exist positive constants $0 < m_\mu < M_\mu < \infty$ such that

$$m_\mu(t - s) \leq \mu(t^2)t - \mu(s^2)s \leq M_\mu(t - s) \quad \text{for all } t \geq s \geq 0. \quad (21)$$

The assumption (M2) implies (F1) and (F2) with $L = 3M_\mu$ and $\nu = m_\mu$, respectively; see, e.g. [13, Proposition 25.26]. Furthermore, $F$ is Gateaux differentiable and the derivative satisfies (7) and (8) with $\alpha_F = m_\mu$ and $\beta_F = 2M_\mu - m_\mu$, respectively; we refer to the proof of [8, Proposition 5.3]. Note that, in the given setting, $2\alpha_F/L = 2m_\mu/3M_\mu < 1$, and thus the convergence of the classical Newton method is not guaranteed a priori. For the verification of (F4), we refer to [8, Lemma 5.4] and the proof of [8, Proposition 5.3]. Finally, $F$ has the potential

$$H(u) := \int_\Omega \psi(|\nabla u|^2) \, dx - \langle g, u \rangle, \quad (22)$$

where $\psi(s) := \frac{1}{2} \int_0^s \mu(t) \, dt$.

5.1. Numerical experiments. Now we run two experiments to test our adaptive algorithm. For that purpose, we discretise the continuous problem with a conforming $P1$-finite element method. Moreover, we compute an accurate approximation of the discrete solution with the Kačanov iteration scheme, see, e.g. [11, Section 4.5] or [13, Section 25.14], which is guaranteed to converge given the assumptions (M1) and (M2). In Algorithm 1, we set the parameters $\sigma = 0.8$ and $\theta = 0.1$ for both of our experiments.

**Experiment 5.1.** In our first experiment, we consider the L-shaped domain $\Omega := (-1,1)^2 \setminus ([0,1] \times [0,1])$ and the diffusion coefficient $\mu(t) = (t + 1)^{-1} + \frac{1}{2}$. It is straightforward to verify that $\mu$ satisfies (M1) and (M2); specifically, it can be shown that $m_\mu = 3/8$ and $M_\mu = 3/2$. The source function $g$ is chosen in such a way that the exact solution of the continuous problem is given by $u^*(x,y) = \sin(\pi x) \sin(\pi y)$, where $(x,y) \in \mathbb{R}^2$ denote the Euclidean coordinates. Finally, we consider the constant null function as our initial guess.

In this experiment, we compare the performance of Algorithm 1 and the damped Newton method with fixed step-size $\delta = \frac{\alpha_F}{L} = \frac{m_\mu}{3M_\mu}$. We can observe in Fig. 1 that, in the given setting, $\delta \equiv 1$ is an admissible step-size and leads to a quadratic decay rate of the error $\|u^n - u^*\|_X$. In contrast, the fixed damping parameter $\delta = \frac{m_\mu}{3M_\mu}$, which is chosen in accordance to Theorem 2.1, leads to a very poor convergence rate.
Experiment 5.2. In our second numerical test, we consider the square domain \( \Omega := (0,1)^2 \). Furthermore, our diffusion coefficient is given by

\[
\mu(t) := \frac{\gamma}{\sqrt{t + k}} + 2\zeta
\]

for some constants \( k, \gamma, \zeta > 0 \); this function is based on the Bercovier–Engelman regularisation of the viscosity coefficient of a Bingham fluid, cf. [3]. Clearly, \( \mu \) is differentiable and decreasing. Furthermore, the assumption (M2) is satisfied with \( m_\mu = 2\zeta \) and \( M_\mu = 2\zeta + k\gamma \), see [7, Lemma 4.1]. For our experiment, we set \( \gamma = 0.3, \zeta = 1 \), and \( k = 100 \). For simplicity, we consider the same source function \( g \) as in the previous experiment and set \( u^0(x,y) = \sin(\pi x)\sin(\pi y) \), or more precisely its piecewise linear interpolation in the nodes of the mesh.

Here, we consider the adaptive damped Newton method from Algorithm 1, the damped Newton iteration with fixed step-size \( \delta = 0.2 \) (which, however, does not guarantee the convergence a priori), and the classical Newton scheme. As can be observed in Fig. 2, the undamped scheme does not converge in the given setting. In contrast, and in accordance with the theory, the adaptive step-size strategy from Algorithm 1 leads to the convergence of the damped Newton method. Moreover, after a first phase of reduced step-sizes, and in turn of inferior decay rate, the damping parameter finally becomes one, which then leads to the local quadratic convergence rate. As a consequence, the adaptive damping scheme outperforms the Newton method with fixed step-size \( \delta = 0.2 \), which converges at a linear rate.

6. Conclusion. We have presented an adaptively damped Newton method, which converges, under suitable assumptions, for any initial guess to the unique solution in the context of strongly monotone and Lipschitz continuous operator equations. Moreover, our numerical experiments highlighted that this algorithm may indeed converge in situations where the classical Newton scheme does not.
Figure 2. Experiment 5.2: Comparison of the adaptive and fixed damped Newton methods and the classical Newton scheme.

fails, and, nonetheless, exhibits the favourable local quadratic convergence rate of the undamped method. Finally, we recall that the proposed nonlinear solver can be applied in combination with an adaptive finite element method to obtain a computational scheme, which converges to the continuous solution with optimal rate with respect to the overall computational time.

Funding Information Open Access funding enabled and organized by Projekt DEAL.

Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[1] Amrein, M., Wihler, T.P.: An adaptive Newton-method based on a dynamical systems approach. Commun. Nonlinear Sci. Numer. Simul. 19(9), 2958–2973 (2014)

[2] Amrein, M., Wihler, T.P.: Fully adaptive Newton–Galerkin methods for semi-linear elliptic partial differential equations. SIAM J. Sci. Comput. 37(4), A1637–A1657 (2015)
[3] Bercovier, M., Engelman, M.: A finite element method for incompressible non-Newtonian flows. J. Comput. Phys. 36(3), 313–326 (1980)

[4] Deuflhard, P.: Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms. Springer Series in Computational Mathematics, vol. 35. Springer, Berlin (2004)

[5] Geiger, C., Kanzow, C.: Theorie und Numerik restringierter Optimierungsaufgaben. Springer, Berlin (2002)

[6] Heid, P., Praetorius, D., Wihler, T.P.: Energy contraction and optimal convergence of adaptive iterative linearized finite element methods. Comput. Methods Appl. Math. 21(2), 407–422 (2021)

[7] Heid, P., Süli, E.: An adaptive iterative linearised finite element method for implicitly constituted incompressible fluid flow problems and its application to Bingham fluids. Appl. Numer. Math. 181, 364–387 (2022)

[8] Heid, P., Wihler, T.P.: Adaptive iterative linearization Galerkin methods for nonlinear problems. Math. Comp. 89(326), 2707–2734 (2020)

[9] Heid, P., Wihler, T.P.: On the convergence of adaptive iterative linearized Galerkin methods. Calcolo 57(3), Paper No. 24, 23 pp. (2020)

[10] Heid, P., Wihler, T.P.: A modified Kačanov iteration scheme with application to quasilinear diffusion models. ESAIM Math. Model. Numer. Anal. 56(2), 433–450 (2022)

[11] Nečas, J.: Introduction to the Theory of Nonlinear Elliptic Equations. John Wiley and Sons, New Jersey (1986)

[12] Polyak, B., Tremba, A.: New versions of Newton method: step-size choice, convergence domain and under-determined equations. Optim. Methods Softw. 35(6), 1272–1303 (2020)

[13] Zeidler, E.: Nonlinear Functional Analysis and Its Applications. II/B. Nonlinear Monotone Operators. Springer, New York (1990)

PASCAL HEID
Department of Mathematics
Technical University of Munich
Boltzmannstr. 3
85748 Garching bei München
Germany

and

Munich Center for Machine Learning (MCML)
Munich
Germany
e-mail: pascal.heid@ma.tum.de

Received: 28 January 2023
Revised: 16 March 2023

Accepted: 30 March 2023