Accepted for publication in **Enumath 2017 Proceedings**, 2018

An accelerated Newton method for nonlinear materials in structure mechanics and fluid mechanics

T. Richter *   C. Mehlmann †

We analyze a modified Newton method that was first introduced by Turek and coworkers. The basic idea of the acceleration technique is to split the Jacobian $A'(x)$ into a "good part" $A_1'(x)$ and into a troublesome part $A_2'(x)$. This second part is adaptively damped if the convergence rate is bad and fully taken into account close to the solution, such that the solver is a blend between a Picard iteration and the full Newton scheme. We will provide first steps in the analysis of this technique and discuss the effects that accelerate the convergence.

1 Introduction

On the ENUMATH 2015 conference in Ankara, Turek and coworkers [4] presented an accelerated Newton scheme, the *operator related damped Newton method*, for the solution of complex non-Newtonian flow problems. The basic idea is to split the Jacobian $A(x)$ into a "good part" $A_1(x)$ that stabilizes the solution and into the "bad part" $A_2(x)$. This splitting is mostly *ad hoc*. The Jacobian is replaced by $A_\delta(x) = A_1(x) + \delta A_2(x)$ and the parameter $\delta \in [0, 1]$ is adaptively tuned to the recent convergence history.

Since the first notion of the method it has been used by different authors in several application problems like granulate flow [5], fracture propagation [7] and also by ourselves in a study on a viscous-plastic rheology used in describing the dynamics of the sea ice layer on the arctic and antarctic ocean [6].

In most of these applications, the partitioning of the Jacobian matrix $A(x)$ is ad hoc. We discuss, as example, the Navier-Stokes equations: the Newton linearization of the convective term $N(\mathbf{v})(\phi) := (\mathbf{v} \cdot \nabla \mathbf{v}, \phi)$ in a search direction $\mathbf{w}$ is given by $(\mathbf{v} \cdot \nabla \mathbf{w}, \phi) + (\mathbf{w} \cdot \nabla \mathbf{v}, \phi)$. Here, a possible choice for the "good part" would by $(\mathbf{v} \cdot \nabla \mathbf{w}, \phi)$ due to definiteness, the remaining part $(\mathbf{w} \cdot \nabla \mathbf{v}, \phi)$ would be left due to its undetermined sign that could possibly cause problems. The optimal choice is less obvious for complex problems like fracture mechanics [7].

A second problem of the operator related damped Newton approach is the proper control of the parameter δ used for partial damping. In [5] the author computationally tested various techniques for adaptively controlling this parameter and derived a formula that performs very well in different

*Institute of Analysis and Numerics, Otto-von-Guericke University Magdeburg, thomas.richter@ovgu.de
†Institute of Analysis and Numerics, Otto-von-Guericke University Magdeburg, carolin.mehlmann@ovgu.de

applications [6, 7]. There is however no rigorous proof for the robustness of this strategy that relates the damping parameter to the convergence rate of the Newton scheme. The relevant mechanism of the operator related damped Newton scheme is not understood. Several applications show that a very small deviation of δ from 1 already shows a substantial effect.

Aim of this article is to give some better inside to the operator related Newton scheme. In section 2 we start by presenting some results from an application of this Newton strategy to a viscous-plastic flow model taken from [6]. Here we put the focus on the partitioning of the matrices and extend the findings to more general classes of problems. Next in section 4 we give first steps for an analysis of this Newton method. Finally, we conclude.

2 An application of the operator related Newton scheme to a viscous-plastic flow problem

The results presented in this section are mostly taken from [6] and have also been presented on the 2017 Enumath conference.

2.1 A brief description of a viscous-plastic sea ice model

A common and widely used sea ice model was introduced by Hibler [2] and is based on the following modeling assumptions:
1) The ice layer on the ocean is described as a two dimensional vector field denoting the velocity $\mathbf{v}:\Omega\to\mathbb{R}^2$, by the scalar ice height $H:\Omega\to\mathbb{R}_+$ and the scalar ice concentration $A:\Omega\to\mathbb{R}\cap[0,1]$ that indicates whether a certain area is covered by stiff ice ($A=1$) or by open water ($A=0$). H and A are advected with the ice velocity $\mathbf{v}$.
2) Ice mass is conserved by linearized conservation laws

$$\rho_{ice}\partial_t\mathbf{v}-\operatorname{div}\boldsymbol{\sigma}=\rho_{ice}\mathbf{f}, \tag{1}$$

where $\mathbf{f}$ describes forcing by ocean and atmosphere current as well as all gravitational or rotational effects (here these effects are simplified and we refer to Hibler [2] and [6] for a complete description of the model.
3) Ice is considered as a viscous-plastic material following a normal flow. The yield stress (the ice strength P) is given by

$$P=P^{\star}H\exp\big(-20(1-A)\big), \tag{2}$$

where $P^{\star}=2.75\cdot10^4\mathrm{N/m}^3$ is a constant. Below this critical value a linear viscous behavior is assumed. The stresses consist with an elliptical (eccentricity $e=2$) yield curve giving the stress-strain relationship

$$\boldsymbol{\sigma}=2e^{-2}\zeta\dot{\boldsymbol{\epsilon}}'+\zeta\,\mathrm{tr}(\dot{\boldsymbol{\epsilon}})I-\frac{P}{2}I \tag{3}$$

with $\epsilon=\frac{1}{2}(\nabla\mathbf{v}+\nabla\mathbf{v}^T)$, $\dot{\boldsymbol{\epsilon}}'=\dot{\boldsymbol{\epsilon}}-\frac{1}{2}\,\mathrm{tr}(\dot{\boldsymbol{\epsilon}})I$ and the viscosity ζ.
4) In [6] we have derived the following formulation to express the viscosity

$$\zeta=\frac{P}{2\Delta(\dot{\boldsymbol{\epsilon}})},\quad \Delta(\dot{\boldsymbol{\epsilon}})=\big(\Delta_{\min}^2+2e^{-2}\dot{\boldsymbol{\epsilon}}':\dot{\boldsymbol{\epsilon}}'+\mathrm{tr}(\dot{\boldsymbol{\epsilon}})^2\big)^{\frac{1}{2}}, \tag{4}$$

The constant $\Delta_{\min}=2\cdot10^{-9}\mathrm{s}$ determines the critical strain for a transition to the viscous case. $\Delta(\dot{\boldsymbol{\epsilon}})$ contains a smooth transition between viscous and plasti'c behavior to ensure differentiability.

Besides problems of data uncertainty, size and variability of geometry and parameters the great challenge in sea ice simulations is the nonlinearity coming from the plastic material behavior. The viscosities given by (4) can greatly vary. For ice at rest, their maximum is bound by $P^\star \Delta_{min}^{-1} \approx 10^{13}\text{Ns/m}^2$ while for typical strain rates of fast moving ice (usually still less than 50cm/s) the viscosities can go down to values as small as 10^2Ns/m^2. Currently there exists no satisfying nonlinear solver for accurate solutions to this nonlinear sea ice models. The most advanced and reliable techniques are based on a Newton method with various globalization techniques, see Lemieux, Losch et al. [3].

2.2 Linearization by Newton

We only discuss the linearization of the critical stress terms. The complete derivation is given in [6]. The stresses $A(\mathbf{v})(\phi) = (\boldsymbol{\sigma}, \nabla\phi)$ can be written as

$$A(\mathbf{v})(\phi) := \left(\frac{P\boldsymbol{\tau}(\mathbf{v}), \boldsymbol{\tau}(\phi)}{\left(\Delta_{\min}^2 + \boldsymbol{\tau}(\mathbf{v}) : \boldsymbol{\tau}(\mathbf{v})\right)^{\frac{1}{2}}} \right)_\Omega, \ \boldsymbol{\tau}(\phi) = \sqrt{2}e^{-1}\dot{\boldsymbol{\epsilon}}'(\phi) + \frac{1}{\sqrt{2}}\operatorname{tr}(\dot{\boldsymbol{\epsilon}}(\phi))I. \tag{5}$$

This allows for a simple expression of the Jacobian as

$$A'(\mathbf{v})(\mathbf{w},\phi) = \underbrace{\left(\frac{P\boldsymbol{\tau}(\mathbf{w}), \boldsymbol{\tau}(\phi)}{\left(\Delta_{\min}^2 + \boldsymbol{\tau}(\mathbf{v}) : \boldsymbol{\tau}(\mathbf{v})\right)^{\frac{1}{2}}} \right)_\Omega}_{=:A_1'(\mathbf{v})(\mathbf{w},\phi)} - \underbrace{\left(\frac{P(\boldsymbol{\tau}(\mathbf{v}) : \boldsymbol{\tau}(\phi))\boldsymbol{\tau}(\mathbf{w}), \boldsymbol{\tau}(\mathbf{v})}{\left(\Delta_{\min}^2 + \boldsymbol{\tau}(\mathbf{v}) : \boldsymbol{\tau}(\mathbf{v})\right)^{\frac{3}{2}}} \right)_\Omega}_{=:A_2'(\mathbf{v})(\mathbf{w},\phi)}.$$

Theorem 3 (Jacobian of the sea ice problem) *Let $\boldsymbol{\tau}(\cdot)$ be continuous with a Korn's like inequality $c_1\|\mathbf{w}\|_{H^1(\Omega)} \le \|\tau(\mathbf{w})\| \le c_2\|\mathbf{w}\|_{H^1(\Omega)}$. Then, the Jacobian is positive definite satisfying the estimates*

$$c_1 \min_\Omega \left\{ \frac{\Delta_{min}^2}{\Delta^2} \right\} \|\sqrt{\zeta}\mathbf{w}\|_{H^1(\Omega)}^2 \le A'(\mathbf{v})(\mathbf{w},\mathbf{w}) \le c_2\|\sqrt{\zeta}\mathbf{w}\|_{H^1(\Omega)}^2$$
$$c_1\|\sqrt{\zeta}\mathbf{w}\|_{H^1(\Omega)}^2 \le A_1'(\mathbf{v})(\mathbf{w},\mathbf{w}) \le c_2\|\sqrt{\zeta}\mathbf{w}\|_{H^1(\Omega)}^2$$
$$-c_1 \min_\Omega \left\{ 1 - \frac{\Delta_{min}^2}{\Delta^2} \right\} \|\sqrt{\zeta}\mathbf{w}\|_{H^1(\Omega)}^2 \le A_2'(\mathbf{v})(\mathbf{w},\mathbf{w}) \le 0.$$

Proof: The proof is given in [6]. From the analysis of the Jacobian we derive the following consequences: as $A'(\mathbf{v})$ is positive definite, a properly damped Newton scheme should converge globally. The Jacobian's lower bound can get very small for high strain rates (such that $\Delta_{\min}/\Delta \approx \Delta_{min} = 2\cdot 10^{-9}$). The first part of the Jacobian $A_1'(\mathbf{v})$ is positive definite and bounded with constants that are properly scaled. This explains the choice as "good term". We add arguments in section 4. Finally, the second part $A_2'(\mathbf{v})$ is negative semidefinite and the reason for the critical lower bound in the full Jacobian. Hence, this part is selected for damping.

3.1 The operator related damped Newton method for the simulation of the sea ice dynamics

We apply the operator related damped Newton scheme to the partitioning $A_\delta'(\mathbf{v}) = A_1'(\mathbf{v}) + \delta A_2'(\mathbf{v})$ and using the adaptive control introduced in [5]. In figure (1) we show the progression of the Newton scheme with and without control of the parameter δ. In addition, both settings are completed with a standard line search scheme. While standard Newton does not give convergence, that operator related damping helps. We observe, that the damping parameter δ is never decreased to values below $\delta \approx 0.98$. Nevertheless, the effect is immense. See [6] for further details.

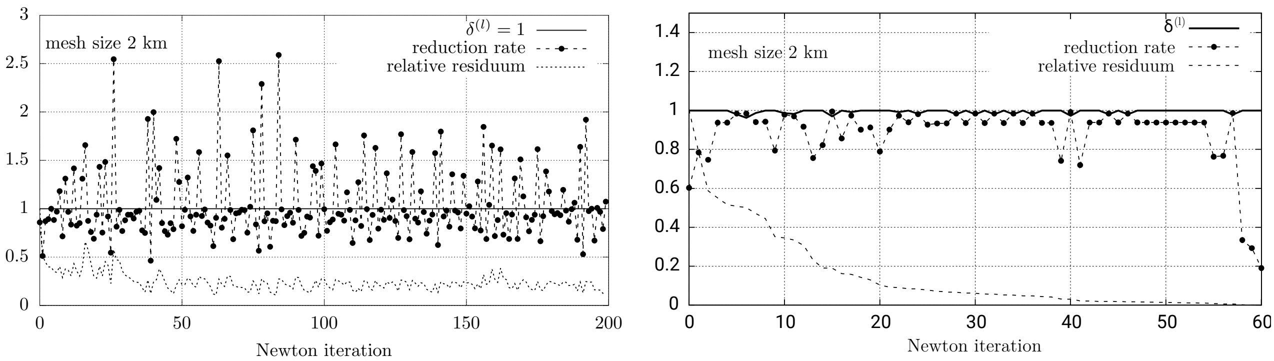

Figure 1: Sea ice simulation on a mesh of 2km mesh size (which is considered as very accurate within the community). Left: simulation with standard Newton with line-search. No convergence is observed. Right: the operator related damped Newton gives convergence. It is observed that the adaptive control required only vary small deviations of the damping parameter δ from 1.

3.2 Related results for various viscosity models

The stress formulation of the sea ice problem (5) can be easily altered to fit various other applications. The p-Laplacian could be interpreted by choosing the tensor $\boldsymbol{\tau}(\mathbf{v}) = \nabla \mathbf{v}$ and by relaxing the exponent to $p \in (1, \infty)$

$$A(\mathbf{v})(\phi) = \int_\Omega \frac{\boldsymbol{\tau}(\mathbf{v}) \cdot \boldsymbol{\tau}(\phi)}{\left(\epsilon^2 + \boldsymbol{\tau}(\mathbf{v}) \cdot \boldsymbol{\tau}(\mathbf{v})\right)^{\frac{2-p}{2}}} \, \mathrm{d}x,$$

where $\epsilon > 0$ is a regularization parameter. For the Jacobian we get the same results as indicated in Theorem 3. The limit case $p = 1$ is the parabolic minimal surface problem that also falls into our class of possible formulations. Power-law fluid models like the Carreau fluid can be represented by using the tensor $\boldsymbol{\tau}(\mathbf{v}) = \dot{\boldsymbol{\epsilon}}'(\mathbf{v})$ and the variational formulation

$$A(\mathbf{v})(\phi) = \int_\Omega \left(\nu_\infty + \frac{\nu_0 - \nu_\infty}{\left(1 + \lambda^2 \boldsymbol{\tau}(\mathbf{v}) : \boldsymbol{\tau}(\mathbf{v})\right)^{\frac{1-n}{2}}} \right) \boldsymbol{\tau}(\mathbf{v}) : \boldsymbol{\tau}(\phi) \, \mathrm{d}x$$

with limiting viscosities ν_0 and ν_∞, the power-index n and the relaxation time λ. For all these cases we can show similar bounds to theorem 3. The complete Jacobian - in all these cases - is positive definite with a lower bound that is possibly close to zero and that depends on the regularization parameter ϵ and the limiting viscosities ν_0, ν_∞. We observe a similar structure. The derivative of the linear part is positive definite with robust bounds, while the remaining part (the derivative with respect to the viscosity) is negative semidefinite. This allows a clear splitting of the Jacobian into an easy part A_1' and into a destabilizing part A_2'.

4 First analysis of the operator related Newton scheme

The effect of the operator related Newton scheme is an acceleration of the convergence. It is not strictly a globalization method as we apply it to problems with positive definite Jacobian that already allow for global convergence. The problem in numerical realization is the identification of optimal line search parameters. To start the exposition we give a variant of a well-known result concerning globalization of Newton schemes (compare Lemma 3.5 in [1]).

Theorem 5 *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ with $f \in C^1(D)$ and Lipschitz continuous Jacobian $f' = f_1' + f_2'$ (Lipschitz constant being $L > 0$). For $\delta \in [0,1]$ define $f_\delta' := f_1' + \delta f_2'$ such that*

$$\|f_\delta'(x)\| \geq \gamma_\delta \|x\|, \quad \gamma_\delta > 0.$$

Then, let $\delta \in [0,1]$ and $f_\delta' := f_1' + \delta f_2'$ the following estimate holds for the operator related damped Newton iterate $x_\omega = x - \omega f_\delta'(x)^{-1} f(x)$

$$\|f(x_\omega)\| \leq \left(1 - \omega + (1-\delta)\omega \|f_2'(x) f_\delta'(x)^{-1}\| + \frac{L\omega^2}{2\gamma_\delta} \|f_\delta'(x)^{-1} f(x)\|\right) \|f(x)\|.$$

Proof: Let $x \in D$ be the current iterate. The proof is standard and based on an intermediate Newton step

$$x_s := x - f_\delta'(x)^{-1} f(x) \;\Rightarrow\; f(x_\omega) = \left[I - \int_0^\omega f'(x_s) f_\delta'(x)^{-1} f(x)\right] f(x)$$

Insertion of $\pm f_\delta'(x)$ and estimation gives the result. We first discuss the standard case $\delta = 1$ corresponding to the full Newton scheme. The convergence rate ρ and the optimal damping parameter ω_{opt} with a resulting optimal bound for the convergence rate is given by

$$\rho(\omega) = 1 - \omega + \frac{L\omega^2}{2\gamma} \quad \Rightarrow \quad \omega_{\text{opt}} = \frac{\gamma}{L} \quad \Rightarrow \quad \rho(\omega_{\text{opt}}) = 1 - \frac{\gamma}{2L}, \tag{6}$$

where $\gamma := \gamma_1$ is the lower bound for the full Jacobian from theorem 3. The sea ice problem yields $\gamma \approx 10^{-14}$ (this value is taken from simulations and is actually less worse than the estimate $\gamma \sim \Delta_{\min}^2 = 4 \cdot 10^{-18}$ might suggest). This explains the very slow convergence of standard Newton schemes. Furthermore, with ω_{opt} being defined by (6) it is difficult to numerically identify a correct parameter ω in this tiny range. To understand the effect of the operator related damping we give bounds for the Jacobian A_δ' comparable to those indicated in theorem 3. Combining $A_1'(\mathbf{v})$ and $A_2'(\mathbf{v})$ directly gives

$$c_1 \max\left\{1 - \delta, \delta \min_\Omega \left\{\frac{\Delta_{\min}^2}{\Delta^2}\right\}\right\} \|\sqrt{\zeta}\mathbf{w}\|_{H^1(\Omega)}^2 \leq A_\delta'(\mathbf{v})(\mathbf{w},\mathbf{w}) \leq c_2 \|\sqrt{\zeta}\mathbf{w}\|_{H^1(\Omega)}^2$$

such that the lower bound is shifted to $\gamma_\delta \geq 1 - \delta$. Compared to $\gamma \sim 10^{-14}$ even very small deviations of δ from 1 result in a substantial improvement of this lower bound. In light of (6) the improved bound *would* give a prediction of $\omega_{\text{opt},\delta}$ and $\rho(\omega_{\text{opt},\delta})$ shows a great improvement

$$\omega_{\text{opt},\delta} = \frac{\gamma_\delta}{L} \quad \Rightarrow \quad \rho(\omega_{\text{opt},\delta}) = 1 - \frac{\gamma_\delta}{2L}.$$

However, this result is only part of the truth, as the complete estimate given in theorem 5 for $\delta < 1$ includes the term (skipped here)

$$(1-\delta)\omega \|f_2'(x) f_\delta'(x)^{-1}\|$$

that counteracts the acceleration by an improved bound for $\gamma_\delta > 0$. A complete analysis cannot be performed without further insight to the differential operators.

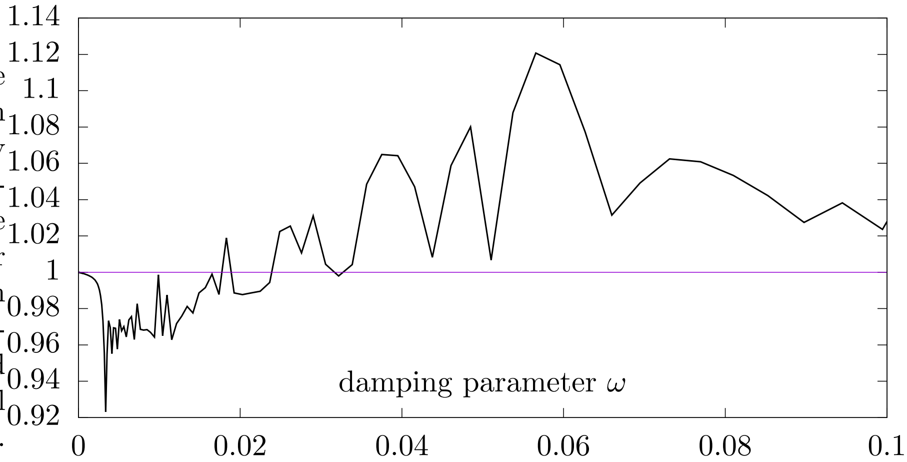

Figure 2: Convergence of the standard Newton scheme in the very first step. Variation of the relative residual change for different linesearch parameters. Convergence is obtained for only a very small range of parameters.

5.1 Numerical analysis of the operator related Newton method

To validate our findings and the role of the damping parameter δ we analyze as simpler problem (compared to the sea ice problem) the limit-case of the p-Laplacian, given by the variational formulation

$$A(u,\phi) = \big(\zeta(u)\nabla u, \nabla\phi\big)_\Omega, \quad \zeta(u) := \big(\varepsilon + |\nabla u|^2\big)^{-\frac{1}{2}}.$$

This is a variant of the minimal surface problem which is obtained for $\varepsilon = 1$. The Jacobian is given by

$$A'(u)(w,\phi) = \underbrace{(\zeta(u)\nabla w, \nabla\phi)_\Omega}_{=:A_1'(u)(w,\phi)} \underbrace{- \big(\zeta(u)^3(\nabla u\cdot\nabla w)\nabla u, \nabla\phi\big)_\Omega}_{=:A_2'(u)(w,\phi)}$$

and it directly suggests a splitting. It is easy to check that the partially damped Jacobian $A_\delta'(u)(w,\phi) := A_1'(u)(w,\phi) + \delta A_2'(u)(w,\phi)$ is positive definite for all $\delta \in [0,1]$ and allows for the bounds

$$c_1 \max\left\{1-\delta, \min_\Omega\left\{\varepsilon^2\zeta(u)^2\right\}\right\} \|\sqrt{\zeta}\nabla w\|^2 A_\delta'(u)(w,w) \le c_2\|\sqrt{\zeta}\nabla w\|^2.$$

Given $\varepsilon \ll 1$ even small deviations of δ from 1 have a large effect.

We numerically study this problem for $\varepsilon^2 = 10^{-4}$ using a standard finite element discretization with piecewise linear finite elements on a quadrilateral mesh of the domain $\Omega = (0,1)^2$ that consists of 64×64 elements. We study the homogeneous problem with Dirichlet $u(x,y) = \sin(\pi xy) + \sin(\pi(2x+y))$ on $\partial\Omega$. The initial solution $u_h^{(0)}$ is set to zero within the domain and it complies with the Dirichlet data on the boundary. This set of parameters is sufficiently difficult to pose great challenges to a standard Newton scheme that - by theory - should globally converge, as the Jacobian is positive definite. It is however very difficult to numerically find suitable line search parameters ω. In figure 2 we show the convergence rates that would result from different choices of the damping parameter ω in the very first step of the Newton iteration. Here, convergence is only surely obtained for $\omega \in (0, 0.01)$.

Finally figure 3 shows the effect of controlling the parameter $\delta \in [0,1]$ for a damping of the negative semidefinite part $A_2'(u)$. We show the results for the very first Newton step and after some iterations when the initial residual was already reduced by two orders of magnitude. These results show two effects of the operator related damped Newton scheme: first, the optimal convergence rates are strongly improved by choosing a proper value for the damping parameter δ. Second - and even more important - a good choice of $\delta < 1$ enlarges the interval where a suitable linesearch parameter ω might be found. This eases the construction of numerically robust solvers. In the first step, given $\delta = 0$, convergence is given for all $\omega \in (0,1]$ insteat of $\omega \in (0, 0.01)$ in the undamped case.

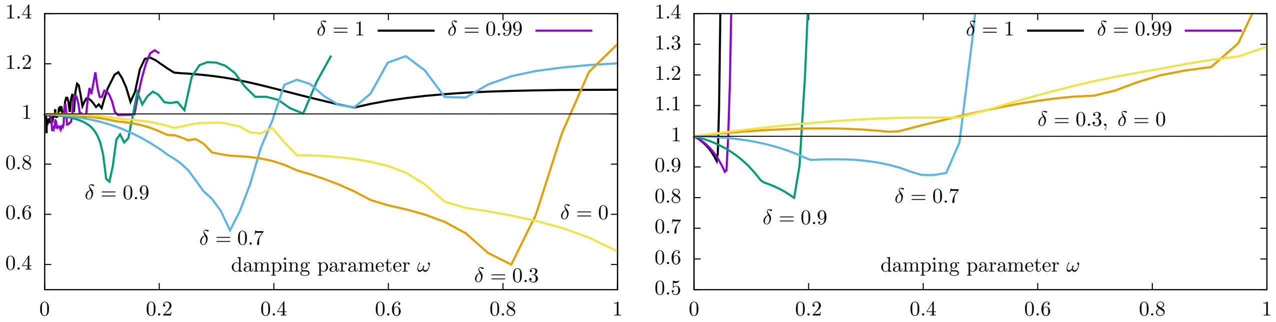

Figure 3: Convergence of the operator related damped Newton scheme for different damping parameters δ. Each figure shows the relative change of the residual after one Newton step for different linesearch parameters.

6 Conclusion

We conclude with two findings: the operator related damped Newton scheme that has been introduced by Mandal, Ouazzi and Turek [4] is highly effective. In different applications it has proven to successfully accelerate Newton solvers. The effects are partially due to an enlarging of the range of linesearch parameters that give convergence. This allows for efficient numerical schemes with few residual evaluations during the linesearch procedure and good convergence rates. Second by correct damping we can improve the convergence rates.

A rigorous analysis of the operator related Newton scheme is still open. In particular a robust adaptive procedure for the choice of the damping parameter that leads to a convergent solver is missing.

References

[1] P. Deuflhard, *Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms*, Computational Mathematics **35**, Springer (2011)

[2] W.D. Hibler, *A dynamic thermodynamic sea ice model*, J. Phys. Oceanogr. **9** (1979), 815–846

[3] M. Losch, A Fuchs, J.F. Lemieux and A. Vanselow, *A parallel Jacobian-free Newton-Krylov solver for a coupled sea ice-ocean model*, J. of Comp. Physics **257** (2014), 901–911

[4] S. Mandal, A. Ouazzi and S. Turek, *Modified Newton solver for yield stress fluids*, Proceedings of the ENUMATH 2015, Springer, 2016.

[5] S. Mandal, *Efficient FEM solver for quasi-Newtonian problems with application to granular materials*, Dissertation, Technical University Dortmund (2016)

[6] C. Mehlmann and T. Richter, *A modified global Newton solver for viscous-plastic sea ice models*, Ocean Modelling **116** (2017), 96–107.

[7] T. Wick, *Modified Newton methods for solving fully monolithic phase-field quasi-static brittle fracture propagation*, CMAME **325** (2017), 577–611