https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

The Newton solver with step size control is faster than the Picard iteration in simulating ice flow (FEniCS-full-Stokes v1.1.0)

Niko Schmidt[1], Angelika Humbert[2,3], and Thomas Slawig[1,4]

[1]Dep. of Computer Science, Kiel University, Kiel, Germany
[2]Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Bremen, Germany
[3]Faculty of Geosciences, University of Bremen, Bremen, Germany
[4]Kiel Marine Science (KMS) – Centre for Interdisciplinary Marine Science

Correspondence: Niko Schmidt (nis@informatik.uni-kiel.de)

Abstract. Solving the momentum balance is the computationally expensive part of simulating the evolution of ice sheets. The momentum balance is described by the nonlinear full-Stokes equations. As a nonlinear problem, they are solved iteratively. We solve these equations with Newton's method. A step size control guarantees global superlinear convergence. For the step size control, we need a minimization problem. Minimizing a specific convex function is equivalent to solving the full-Stokes equations. We use two algorithms to determine the step size for Newton's method: Armijo and the exact step sizes. Additionally, we use the latter for the Picard iteration. Finally, we compare the Picard iteration and the variants of Newton's method in two benchmark experiments, called ISMIP-HOM experiments A and B. These experiments consist of a more realistic domain and are designed to test the quality of ice models. We obtain that the Picard iteration and Newton's method with exact step sizes greatly reduce the necessary number of iterations.

1 Introduction

Simulating the evolution of the ice sheets in Greenland and Antarctica in adequate physics and resolution is a challenging task. The dynamics of ice sheets is described as a fluid mechanical problem with the momentum balance reduced to a Stokes problem as acceleration and Coriolis forces are negligible. In the past computational constraints led to the reduction of the problem by approximating the momentum balance. If the spatial resolution cannot be choosen sufficiently large, the benefit from solving the Stokes problem is lost. Consequently, in practical terms, Stokes models are leading to large problems and thus efficient solvers are inevitable. This is what this study is focusing on.

The full-Stokes equations are nonlinear partial differential equations. These equations are also described as shear thinning, in which the viscosity depends nonlinear on the symmetric gradient. More precisely, we consider the stationary variant of these equations in the variational formulation. This formulation is needed to calculate a solution with finite elements. A standard method to calculate the solution of these equations is the Picard iteration, (see Colinge and Rappaz, 1999). The Picard iteration fixes the nonlinear viscosity, calculates a new velocity, and updates the viscosity.

Instead, we employ Newton's method by formulating the variational formulation as a root problem. If we start near the solution, Newton's method is superlinear convergent, (Hinze et al., 2009). Thus, the error between the approximation and

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

the real solution reduces faster than linear. In contrast, the Picard iteration converges only linearly (see Fraters et al., 2019). However, starting with an unsuitable initial velocity field for Newton's method could lead to a diverging velocity. A step size control guarantees convergence from every initial guess. This step size control is constructed by defining a function that we want to minimize. One variant is presented in Fraters et al. (2019). We consider another approach that only needs to calculate integrals. It allows us to use two different step size controls. Newton's method with one of these step sizes converges from every initial guess superlinear to the solution, (Schmidt, 2023). Additionally, we employ the exact step size for the Picard iteration to provide a possibility to reduce the necessary number of iterations without implementing Newton's method.

The computation of the step size is computationally cheap compared to solving the linear systems of equations in each iteration. The work of Habbal et al. (2017) considers different solvers to reduce the simulation time for solving the system of linear equations. Nonetheless, for all solvers, the system of linear equations is still the main computational effort. Our step size control reduces the computation time by reducing the necessary number of iterations.

As a test case, we use the ISMIP-HOM experiments A and B. These experiments are designed to test the quality of glaciological models. They reflect a sinusoidal bedrock topography, a large domain of the glaciers, and a large aspect ratio.

The manuscript has the following structure: In Sect. 2, we introduce the equations in the variational formulation and the Picard iteration. In the subsequent section, we formulate Newton's method. In Sect. 4, we introduce the new idea, the step size control that decreases the number of iterations and verifies convergence from every initial guess. In Sect. 5, we discuss our numerical solution of the ISMIP-HOM experiments A and B and compare them with the results in Pattyn et al. (2008). Finally, we give a summary in Sect. 6 and an outlook in Sect. 7.

2 The full-Stokes equations as a root problem

Let $\Omega \subseteq \mathbb{R}^N$ with $N \in \{2,3\}$. For describing the movement of ice, we need the second-order tensor $\boldsymbol{\sigma}$, the density ρ, and the gravitational acceleration $\boldsymbol{g}$. These quantities describe the full-Stokes equations, the most complex equations for simulating ice, by:

$$\begin{aligned} -\mathrm{div}\boldsymbol{\sigma} &= -\rho\boldsymbol{g}, \\ \mathrm{div}\boldsymbol{v} &= 0 \end{aligned} \tag{1}$$

on the domain Ω. We describe the stress tensor $\boldsymbol{\sigma}$ with the pressure p, the identity tensor (matrix) $\boldsymbol{I}$, the symmetric gradient D, the velocity $\boldsymbol{v}$, and the viscosity μ by $\boldsymbol{\sigma} := p\boldsymbol{I} - \mu D\boldsymbol{v}$. We define the nonlinear viscosity μ as

$$\mu = B\left(|D\boldsymbol{v}|^2 + \delta^2\right)^{\frac{1-n}{2n}}, \quad (D\boldsymbol{v})_{ij} = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right), \quad |D\boldsymbol{v}|^2 := D\boldsymbol{v} : D\boldsymbol{v} := \sum_{i,j=1}^{N} |(D\boldsymbol{v})_{ij}|^2 \tag{2}$$

with $n \in (1,\infty)$ and $B = B(x_1,x_2,x_3)$, $\delta > 0$. The constant $\delta > 0$ guarantees $\mu < \infty$. We choose $n = 3$, which is in the typical range for glacier simulations, which is $n \in [3,4]$. The boundary consists of the bedrock Γ_b, the surface Γ_s, and the lateral

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

boundary Γ_ℓ. Our boundary conditions are:

$$\begin{aligned} \boldsymbol{v} &= 0 \quad \text{on } \Gamma_b \cup \Gamma_\ell, \\ \boldsymbol{\sigma} \cdot \boldsymbol{n} &= 0 \quad \text{on } \Gamma_s \end{aligned} \tag{3}$$

with the outer normal vector $\boldsymbol{n}$. Here, $\boldsymbol{\sigma} \cdot \boldsymbol{n}$ is the inner tensor-product (matrix-vector multiplication).

We want to consider the variational formulation. We set $H := \{\boldsymbol{v} \in H^1(\Omega)^N;\, \boldsymbol{v}|_{\Gamma_b \cup \Gamma_l} = 0\}$, where $H^1(\Omega)^N$ is the space of vector-valued square integrable functions with a square integrable derivative. We set $L := \{p \in L^2(\Omega);\, \int_\Omega p\, dx = 0\}$ for the space of square integrable functions with zero integral. We determine the variational formulation by multiplying with test functions and using partial integration. We have a small diffusion term with $\mu_0 > 0$ as an aditional summand for mathematical properties. We want to find $\boldsymbol{v} \in H$ and $p \in L$ and test with all $\boldsymbol{\phi} \in H$ and all $\psi \in L$. We define an operator $G : H \times L \to H^* \times L^*$ by

$$\langle G(\boldsymbol{v}, p), (\boldsymbol{\phi}, \psi) \rangle = \int_\Omega B\left(|D\boldsymbol{v}|^2 + \delta^2\right)^{\frac{1-n}{2n}} D\boldsymbol{v} : \nabla \boldsymbol{\phi}\, dx + \mu_0 \int_\Omega \nabla \boldsymbol{v} : \nabla \boldsymbol{\phi}\, dx - \int_\Omega p \mathrm{div} \boldsymbol{\phi}\, dx - \int_\Omega \mathrm{div} \boldsymbol{v} \psi\, dx + \int_\Omega \rho \boldsymbol{g} \cdot \boldsymbol{\phi}\, dx, \tag{4}$$

where $\boldsymbol{v} \in H$ and $p \in L$ are the solution, and $\boldsymbol{\phi} \in H$ and $\psi \in L$ are test functions. The square brackets on the left-hand side of the equation are used because, formally, we have a function that maps to the dual space. The dual space is denoted by the star after the space, e.g. H^* and L^*. There exists a unique solution to that problem for $\mu_0 > 0$, (see Schmidt, 2023). We formulate the problem in infinite-dimensional spaces H and L. In these infinite-dimensional spaces, mathematical convergence properties are independent of the mesh resolution and the used finite elements, as long as the finite elements are a subspace of the infinite-dimensional spaces. Ice models often use finite elements. Moreover, the formulation in discretized spaces is identical, only the functions are from finite-dimensional spaces.

A standard method to solve the variational formulation of the full-Stokes equations in glaciological models is the Picard iteration, see Algorithm 1. We need $\delta > 0$ to guarantee well-posedness of the nonlinear viscosity. Furthermore, we need $\delta > 0$

Algorithm 1 Picard iteration

1: Let $\boldsymbol{v_0} \in H$ and $p_0 \in L$ be given.
2: **for** $k = 0, 1, \ldots$ **do**
3: Calculate $\boldsymbol{v_{k+1}} \in H$ and $p_{k+1} \in L$ with

$$\int_\Omega B\left(|D\boldsymbol{v_k}|^2 + \delta^2\right)^{\frac{1-n}{2n}} D\boldsymbol{v_{k+1}} : \nabla \boldsymbol{\phi}\, dx + \mu_0 \int_\Omega \nabla \boldsymbol{v_{k+1}} : \nabla \boldsymbol{\phi}\, dx - \int_\Omega p_{k+1} \mathrm{div} \boldsymbol{\phi}\, dx - \int_\Omega \mathrm{div} \boldsymbol{v_{k+1}} \psi\, dx = -\int_\Omega \rho \boldsymbol{g} \cdot \boldsymbol{\phi}\, dx$$

for all $\phi \in H$ and $\psi \in L$.
4: **end for**

to calculate derivatives. The variational formulation also exists for $\delta = 0$.

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

3 Newton's method

The Picard iteration converges slowly (see Fraters et al., 2019). Thus, it can be beneficial to consider faster converging algorithms. Newton's method is at least superlinear convergent, also in infinite dimensions, (see Hinze et al., 2009). For Newton's method, the calculation of the derivative is necessary. Due to the variational formulation, we can only express the derivative of G in terms of the direction and the test functions. The derivative of G in $(\boldsymbol{v},p)$ in direction $(\boldsymbol{w},q)$ is

$$\begin{aligned}\langle G'(\boldsymbol{v},p)(\boldsymbol{w},q),(\boldsymbol{\phi},\psi)\rangle = &\int_\Omega \frac{1-n}{n} B\left(|D\boldsymbol{v}|^2+\delta^2\right)^{\frac{1-3n}{2n}} (D\boldsymbol{v}:D\boldsymbol{w})\,(D\boldsymbol{v}:\nabla\boldsymbol{\phi})\,dx + \mu_0\int_\Omega \nabla\boldsymbol{w}:\nabla\boldsymbol{\phi}\,dx \\ &+\int_\Omega B\left(|D\boldsymbol{v}|^2+\delta^2\right)^{\frac{1-n}{2n}} D\boldsymbol{w}:\nabla\boldsymbol{\phi}\,dx - \int_\Omega q\,\mathrm{div}\boldsymbol{\phi}\,dx - \int_\Omega \mathrm{div}\boldsymbol{w}\psi\,dx. \end{aligned} \tag{5}$$

A mathematical proof that G is continuously differentiable is presented in Schmidt (2023). A more detailed deduction of the derivative is in Subsect. $A1$.

Newton's method can solve the full-Stokes equations by Algorithm 2. Because Newton's method is only locally convergent, we use a step size control in Algorithm 2. We explain the step size control in the next section.

Algorithm 2 Globalized Newton's method

1: Let $(\boldsymbol{v_0},p_0)$ be given.
2: **for** $k=0,1,\ldots$ **do**
3: Calculate $(\boldsymbol{w_k},q_k)$ with

$$\langle G'(\boldsymbol{v_k},p_k)(\boldsymbol{w_k},q_k),(\boldsymbol{\phi},\psi)\rangle = -\langle G(\boldsymbol{v_k},p_k),(\boldsymbol{\phi},\psi)\rangle \text{ for all } \boldsymbol{\phi}\in H,\psi\in L.$$

4: Set $\boldsymbol{v_{k+1}} := \boldsymbol{v_k}+\alpha_k\boldsymbol{w_k}$ and $p_{k+1} := p_k+\alpha_k q_k$ with a suitable $\alpha_k>0$.
5: **end for**

4 Step size control

In this section, we derive a global convergent Newton method by using a step size control: We have the actual velocity field $\boldsymbol{v}$ and the direction $\boldsymbol{w}$. Instead of setting our new field $\tilde{\boldsymbol{v}} := \boldsymbol{v}+\boldsymbol{w}$, we choose $\alpha>0$ with $\tilde{\boldsymbol{v}} := \boldsymbol{v}+\alpha\boldsymbol{w}$. We want an algorithm for choosing this α. Classical approaches for determining the step size α check, if the euclidean norm $\|G(\boldsymbol{v_{k+1}},p_{k+1})\|_2$ reduces enough compared to $\|G(\boldsymbol{v_k},p_k)\|_2$. What enough reduction means is, for example, discussed in Hinze et al. (2009). However, we use an alternative approach. Solving $G(\boldsymbol{v},p)=0$ is equivalent to minimizing $J: H\to\mathbb{R}$

$$J(\boldsymbol{v}) = \int_\Omega \frac{n}{1+n} B\left(|D\boldsymbol{v}|^2+\delta^2\right)^{\frac{1+n}{2n}} dx + \frac{\mu_0}{2}\int_\Omega |\nabla\boldsymbol{v}|^2\,dx + \int_\Omega \rho\boldsymbol{g}\cdot\boldsymbol{v}\,dx, \tag{6}$$

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

see Schmidt (2023). Such types of convex functions were also used in Hirn (2013) for $\mu_0 = 0$ with Dirichlet boundary conditions, and Chen et al. (2013) for $\delta = 0$ and $\mu_0 = 0$ with more realistic boundary conditions. The equivalence is clear because the minimizer of the function and the root of the derivative are at the same point for strict convex functions. At first glance, the equivalence seems only true in the divergence-free space without the pressure term. However, with a divergence-free velocity and a derivative in a divergence-free direction the pressure terms do not change the solution. Thus, we have

$$J'(\boldsymbol{v})\boldsymbol{w} = \langle G(\boldsymbol{v}, p), (\boldsymbol{w}, q)\rangle \tag{7}$$

for all divergence-free $\boldsymbol{v}, \boldsymbol{w} \in H$ and $p, q \in L$. Because minimizing J and finding the root of G are equivalent, a natural choice for a step size tries to minimize the problem

$$\min_{\alpha\in(0,\infty)} J_k(\alpha) := \min_{\alpha\in(0,\infty)} J(\boldsymbol{v_k} + \alpha \boldsymbol{w_k}) \tag{8}$$

in each iteration. The function J_k is strictly convex, which implies that a simple bisection, see Algorithm 3, calculates the minimum of J_k. In practice, we calculate the exact step size as precisely as desired. Thus, we denote the approximate exact step size as the exact step size.

Algorithm 3 Exact step size

1: Set $a, b \in [0, \infty)$ with $a < b$.
2: **for** $i = 0, 1, \ldots$ **do**
3: **if** $J_k'((a+b)/2) > 0$ **then**
4: Set $b := (a+b)/2$.
5: **else**
6: Set $a := (a+b)/2$.
7: **end if**
8: **end for**
9: **return** $\alpha := (a+b)/2$

An alternative classical approach is an Armijo step size as in Hinze et al. (2009), see Algorithm 4. This approach does not use the convexity of J_k. We modify the Picard iteration, see Algorithm 1, by a relaxation: We set

$$\tilde{\boldsymbol{v}}_{k+1} := (1-\alpha_k)\boldsymbol{v_k} + \alpha_k \boldsymbol{v_{k+1}}, \quad \tilde{p}_{k+1} := (1-\alpha_k)p_k + \alpha_k p_{k+1} \tag{9}$$

and choose α_k as α in Algorithm 3.

5 Numerical experiments

We analyze the four algorithms we introduced: The classical Picard iteration as a reference, exact step sizes for the Picard iteration and Newton's method, and Armijo step sizes for Newton's method. We implemented all these algorithms in *FEniCS*

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

Algorithm 4 Armijo step size

1: Let $\gamma \in (0, 1/2)$.
2: **for** $i = 0, 1, \ldots$ **do**
3: Set $\alpha := 1$.
4: **while** $J(\boldsymbol{v} + \alpha \boldsymbol{w}) - J(\boldsymbol{v}) > \alpha\gamma J'(\boldsymbol{v})\boldsymbol{w}$ **do**
5: Set $\alpha := 0.5\alpha$.
6: **end while**
7: **end for**
8: **return** α

version 2019.1.0, (see Logg et al., 2012). *FEniCS* is a library that allows to implement variational formulations easily. Hence, it allows fast testing of algorithms without implementing them in complex codes. We determine the performance of these algorithms by comparing each iteration step with a reference solution for the experiments ISMIP-HOM A and B, (see Pattyn et al., 2008). The reference solution is calculated with 80 prescribed Picard iterations. In Pattyn et al. (2008), the authors described the ISMIP-HOM experiments to analyze the quality of ice models. Moreover, they compared simulation results.

We set the physical variables according to Pattyn et al. (2008): $B := 0.5 \cdot (10^{-16})^{-1/3}$ (Pa)$^{-3}$ a^{-1}, $\rho := 910$ kg m^{-3}, and $\boldsymbol{g} := (0, 9.81)$ m s^{-2}.

We set the constant $\delta := 10^{-12}$a^{-1} and $\mu_0 := 10^{-17}$ kg a m^{-1} s^{-2}. We derive the unit for μ_0 by $[\mu_0 |\nabla \boldsymbol{v}|^2] = [\rho \boldsymbol{g} \cdot \boldsymbol{v}]$. In the experimental design, the nonlinear term is $2B(0.5|D\boldsymbol{v}|^2 + \delta^2)^{(1-n)/(2n)}$ instead of $B(|D\boldsymbol{v}|^2 + \delta^2)^{(1-n)/(2n)}$, (see Pattyn et al., 2008). We choose the constant δ such that δ is smaller than the typical magnitude of $D\boldsymbol{v}$, $3 \cdot 10^{-4}$ a^{-1} and 3 a^{-1}, multiplied with the machine precision eps:

$$\delta < eps \sqrt{0.5} |D\boldsymbol{v}|. \tag{10}$$

We choose μ_0 such that

$$\mu_0 \int_\Omega |\nabla \boldsymbol{v}|^2 \, dx < eps \int_\Omega B\big(0.5|D\boldsymbol{v}|^2 + \delta^2\big)^{\frac{1-n}{2n}} |D\boldsymbol{v}|^2 \, dx \tag{11}$$

for typical values of $|D\boldsymbol{v}|$. This choice verifies that δ and the diffusion term $\mu_0 \nabla \boldsymbol{v} : \nabla \boldsymbol{\phi}$ do not influence the result for typical values of $|D\boldsymbol{v}|$ as they are smaller than the machine precision. Nonetheless, they guarantee that Newton's method with Armijo step sizes converges, (see Schmidt, 2023). A more detailed derivation of δ and μ_0 is in Schmidt (2023).

In all methods, we calculate the initial velocity by replacing $(0.5|D\boldsymbol{v}|^2 + \delta^2)^{(1-n)/(2n)}$ with 10^6 and solving this linear problem.

5.1 The original experiment ISMIP-HOM B

In this subsection, we introduce details from Pattyn et al. (2008) that are specific to the experiment ISMIP-HOM B. This experiment has a domain with a sinusoidal, slightly tilted $(0.5°)$, bottom. The boundaries at the left and right are vertical, and

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

the boundary at the top has a linear slope of $0.5°$. Furthermore, periodic boundary conditions are used at Γ_ℓ. The experiment prescribes Dirichlet zero boundary conditions, $\boldsymbol{v} = 0$ on Γ_b and $\boldsymbol{\sigma} \cdot \boldsymbol{n} = 0$ on Γ_s.

The length $L := 5$ km is the horizontal extent. The angle $\beta := 0.5°$ describes a slight decline at the surface and the bottom by

$$z_s(x) = -x \tan(\beta), \quad z_b(x) = z_s(x) - 1000 + 500 \cdot \sin(\omega x) \tag{12}$$

with $\omega := 2\pi/L$.

5.2 Modifications to the experiment ISMIP-HOM *B*

Formulating the convex function J, see Eq. (6), that corresponds to periodic boundary conditions is complicated. Thus, we use the alternative introduced in the supplement of Pattyn et al. (2008), by copying the glacier to the right and the left. We have three copies to the right and the left, see Fig. 1. At the lateral boundaries Γ_ℓ, we impose Dirichlet zero boundary conditions. Also, the resolution at the outer copies is lower than for the original domain. This reduces for the two-dimensional experiment the number of elements by 30 % and in three dimensions by 51 %. Nevertheless, the three-dimensional experiment was performed on a high performance computer. In two dimensions, the local refinement has no relevant impact on the solution. Also, one can simulate the two-dimensional experiment on a laptop.

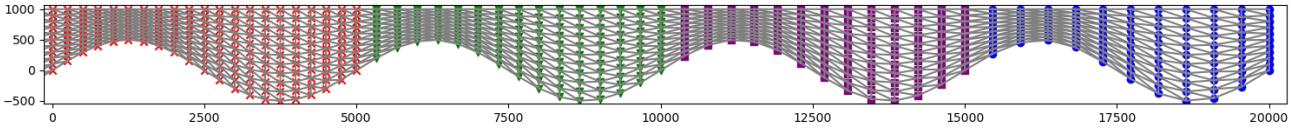

Figure 1. The domain with a grid with red dots and three copies to the right with green, purple, and blue dots.

Instead of the slope, we rotate the gravity. Thus, we should rotate the lateral boundaries Γ_ℓ of the domain. We neglect this and stick to vertical boundaries at the left and the right.

5.3 Results for experiment ISMIP-HOM *B*

Our velocity fields at the surface (see Fig. 2) are close to the mean of the simulations in Pattyn et al. (2008). Also, all our methods produce very similar velocity fields at the surface, as displayed in Fig. 3. Next, we compare how many iterations are necessary to reduce the relative difference and relative local difference compared to the reference solution. We calculate the relative difference and the relative local difference for a velocity $\boldsymbol{v}$ and the reference solution $\boldsymbol{v}_{\mathbf{ref}}$ by

$$\sqrt{\frac{\int_\Omega |\boldsymbol{v} - \boldsymbol{v}_{\mathbf{ref}}|^2 \, dx}{\int_\Omega |\boldsymbol{v}_{\mathbf{ref}}|^2 \, dx}} \text{ and } \sqrt{\int_\Omega \frac{|\boldsymbol{v} - \boldsymbol{v}_{\mathbf{ref}}|^2}{\max(|\boldsymbol{v}_{\mathbf{ref}}|^2, c^2)} \, dx} \tag{13}$$

with $c = 1$ mm a^{-1}. We use two error measurements because one method could be better for one purpose and the other for another. The local relative difference reflects that regions with small velocities should also be represented with a small rela-

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

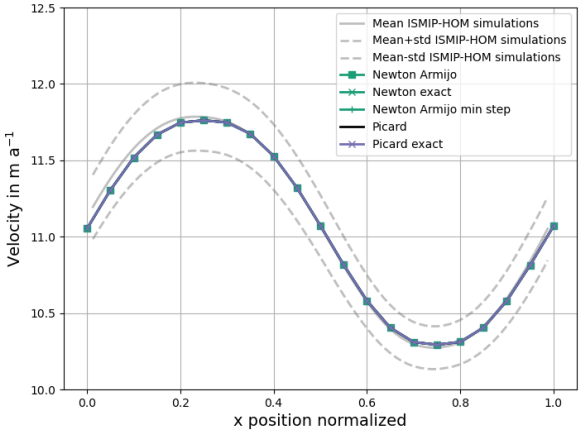

Figure 2. Simulated surface velocity for different solvers for ISMIP-HOM B. All our calculated velocity fields overlap each other. In grey are plotted the mean and the standard deviation (std) from Pattyn et al. (2008) with 6 models. The mean and standard deviation have no values at $x = 0$ and $x = 1$ due to missing values.

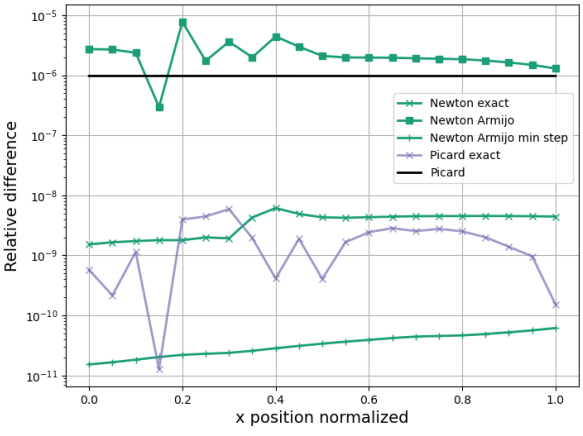

Figure 3. Relative difference of $|\boldsymbol{v} - \boldsymbol{v}_{\mathbf{ref}}|/\boldsymbol{v}_{\mathbf{ref}}$ for each grid point at the surface. The reference solution is the solution from 80 Picard iterations.

tive error. Both error measurements consider the velocity field for the whole domain of the glacier. In contrast, the original experiment (Pattyn et al., 2008) only considers the velocity field at the surface.

Figure 4 displays the relative difference over the iteration number. The classic Picard iteration has a slow convergence rate. It needs 39 iterations to obtain a reduction to 10^{-6}. Newton's method using Armijo step sizes obtains this reduction after only 7 iterations. This reduces the necessary number of iterations by 82 %. We see this even better if we consider just the

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

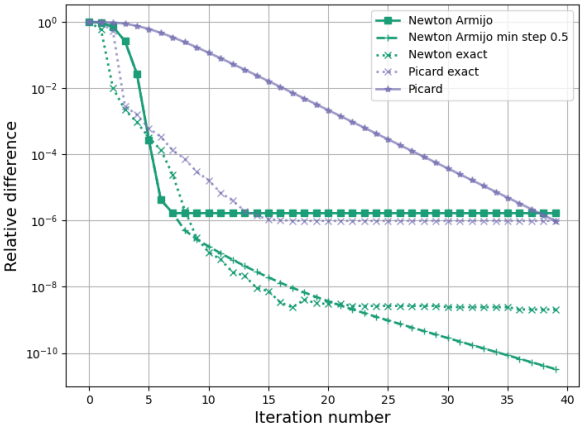

Figure 4. Relative difference compared to the reference solution for ISMIP-HOM B.

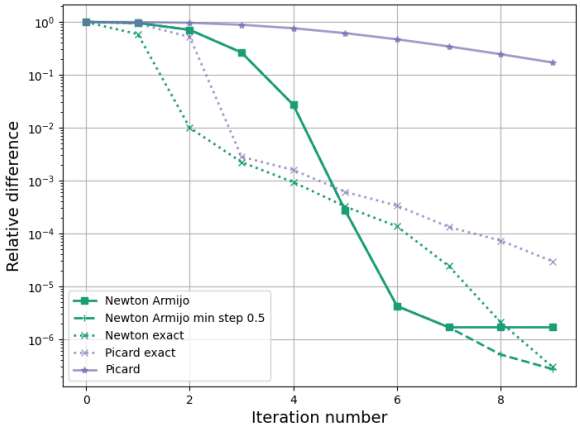

Figure 5. Relative difference compared to the reference solution for ISMIP-HOM B for the first 9 iterations.

first 9 iterations (Fig. 5). After a few iterations, Newton's method does not reduce the relative difference anymore. Imposing a minimal step size of 0.5 helps to circumvent this problem. Then Newton's method reduces the relative difference up to iteration 39. Newton's method with exact step sizes is of similar quality. It has the advantage that the error reduces even more without using a minimal step size. Thus, one less parameter needs to be selected. Even the Picard iteration with exact step sizes is much better than the Picard iteration. It only needs 15 iterations to obtain the accuracy, for which the Picard iteration needs 39 iterations. That corresponds to a reduction of 62 %. The latter approach also has the advantage that there is no need to implement a new method to solve the problem. Only the relatively simple calculation of the step sizes needs to be implemented.

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

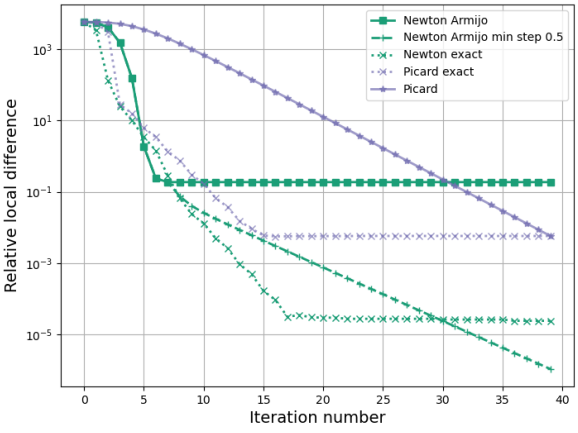

Figure 6. Relative local difference compared to the reference solution for ISMIP-HOM B.

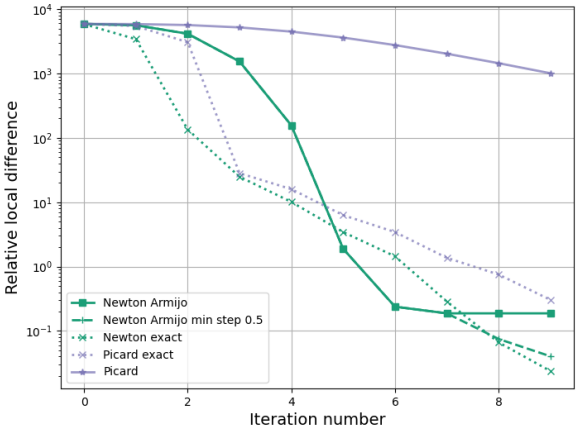

Figure 7. Relative local difference compared to the reference solution for ISMIP-HOM B for the first 9 iterations.

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

The results are really similar for our second measure of the accuracy, the relative local error, see Fig. 6. All our algorithms are better than the classic Picard iteration in this measurement. The reduction with Newton's approach with both step size controls is 77 % now. The fast convergence is again impressive, especially for the first 9 iterations, see Fig. 7.

5.4 The experiment ISMIP-HOM *A*

Because real-world applications are three-dimensional, we consider experiment ISMIP-HOM A. This experiment extends ISMIP-HOM B to three dimensions. All chosen constants are the same as in the experiment ISMIP-HOM B. The experiment ISMIP-HOM A has a sinusoidal bottom in both horizontal dimensions. Again, we have three copies of the glacier in both horizontal directions. Thus, we have in total 48 copies. We describe the surface and bottom by

$$z_s(x,y) = -x\tan(\beta), \quad z_b(x,y) = z_s(x,y) - 1000 + 500 \cdot \sin(\omega x)\sin(\omega y). \tag{14}$$

5.5 Results for experiment ISMIP-HOM *A*

All our methods produce very similar results and are overlapping, see Fig. 8 and Fig. 9. Our simulations reproduce the surface velocity at $y = L/4$ from Pattyn et al. (2008) for the majority of the glacier. But they produce higher velocity values than the mean plus the standard deviation around $x = L/3$. Nonetheless, the maximum relative difference is less than 0.02, see Fig. 8.

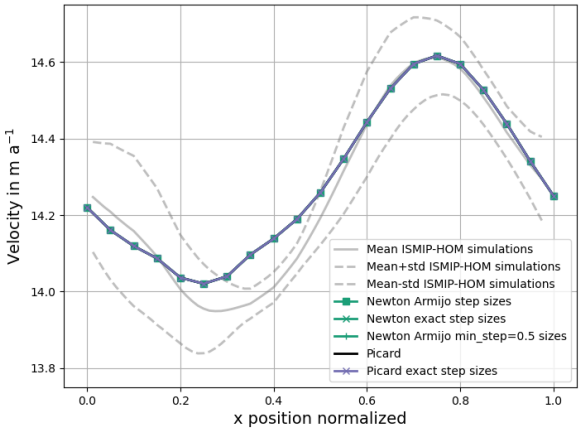

Figure 8. Simulated surface velocity at $y = L/4$ in meters per year for different solvers for ISMIP-HOM A. All our calculated velocity fields overlap each other. In grey are plotted the mean and the standard deviation from Pattyn et al. (2008) with 5 models. The mean and standard deviation have no values at $x = 0$ and $x = 1$ due to missing values.

The general convergence behavior for the three-dimensional experiment is similar to the two-dimensional experiment. However, the Armijo step sizes are even better for Newton's method in three dimensions, see Fig. 10. Again zooming to the first few iterations states the benefit from Newton's method and the step size control more impressing, see Fig. 11. The Picard

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

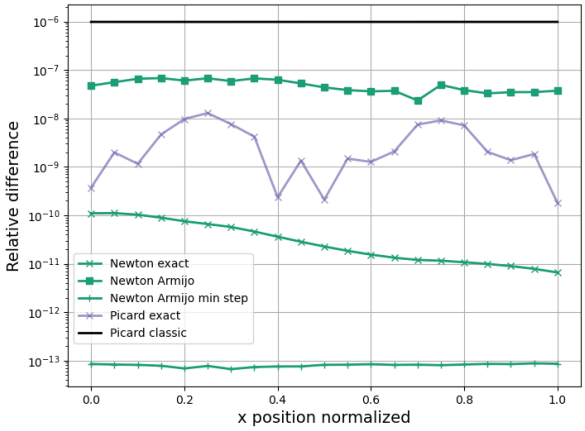

Figure 9. Relative difference of $|v - v_{\mathrm{ref}}|/v_{\mathrm{ref}}$ for each grid point at the surface at $y = L/4$ for ISMIP-HOM A. The reference solution is the solution from 80 Picard iterations.

iteration needs 39 iterations to have the same accuracy as Newton's method using Armijo step sizes after 6 iterations. Thus, the necessary number of iterations is reduced by more than 85 %. Again, a minimum step size of $\alpha = 0.5$ helps to reduce the relative difference after a few iterations. The exact step sizes for Newton's method are even better. They decrease the relative difference, see Fig. 10, and the relative local difference, see Fig. 12, further than the Armijo step sizes. Also, exact step sizes improve the Picard iteration. Again it is interesting to consider the relative local difference for a few iterations, see Fig. 13. This figure emphasizes that the Picard iteration converges slowly compared to the other methods.

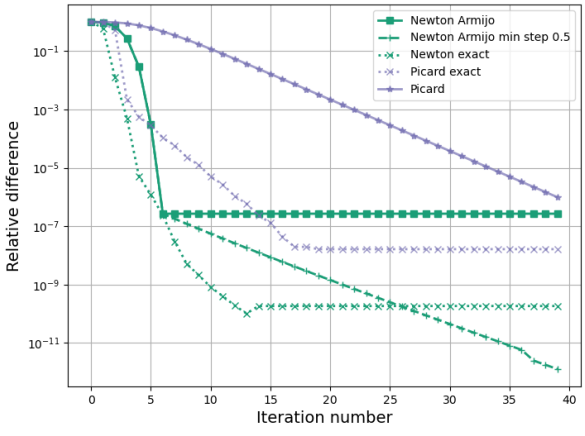

Figure 10. Relative difference compared to the reference solution for ISMIP-HOM A.

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

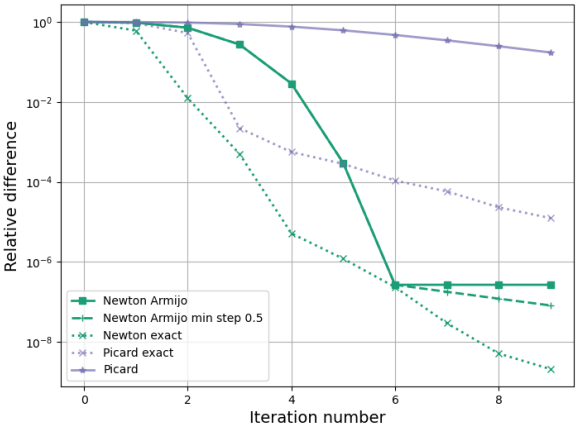

Figure 11. Relative difference compared to the reference solution for ISMIP-HOM A for the first 9 iterations.

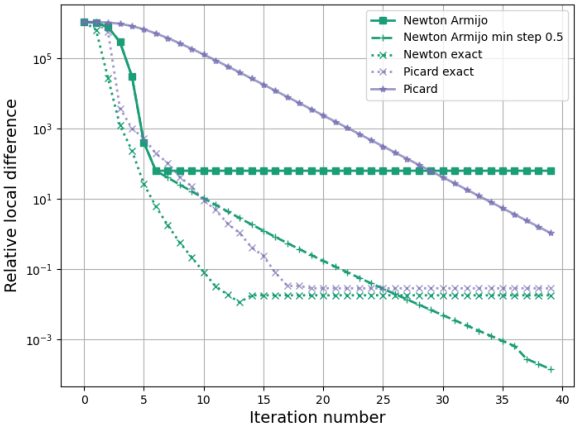

Figure 12. Relative local difference compared to the reference solution for ISMIP-HOM A.

6 Summary and conclusion

We conclude for the ISMIP-HOM experiments A and B: Our simulations are similar to the results in Pattyn et al. (2008) for two and three dimensions. The Picard iteration converges considerably slower than Newton's method. A good choice of step sizes guarantees that Newton's method always converges. The exact step sizes improve the Picard iteration.

The effort to implement the algorithms above is relatively low. For every additional boundary condition to those above, one has to check if a convex functional exists. One only needs to implement these convex functionals, the directional derivatives,

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

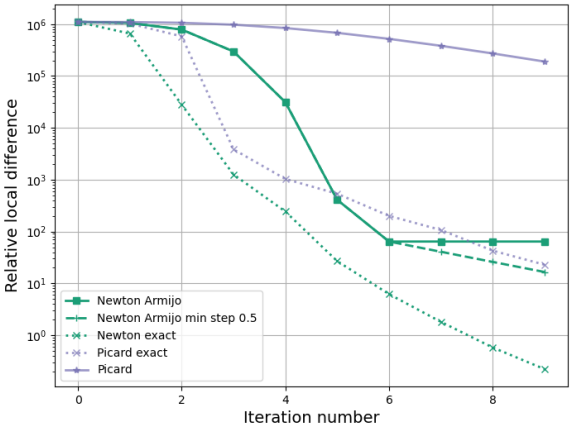

Figure 13. Relative local difference compared to the reference solution for ISMIP-HOM A for the first 9 iterations.

and the Armijo and exact step sizes, respectively. The Picard iteration or Newton's method should already be implemented for solving the full-Stokes equations.

7 Outlook

The next step could be to test the step size control in a real ice model. Also, one could include sliding boundary conditions to consider step by step all physical boundary conditions that apply to a glacier in the real world. The mathematical theory for a nonlinear sliding boundary condition is discussed in Schmidt (2023). Furthermore, the step size control might be used to reduce the number of iterations for the Higher-Order equations. Solving those equations is also equivalent to finding the minimum of a convex function, (see Schoof, 2010).

Code and data availability. The model is available at https://doi.org/10.5281/zenodo.8154332. The latest version of the source code is available at https://github.com/Niko-ich/FEniCS-full-Stokes.

Appendix A: Mathematical derivations

A1 The directional derivative of G

In this appendix, we compute the derivative of G at the velocity $\boldsymbol{v} \in H$ and pressure $p \in L$ in the direction $\boldsymbol{w} \in H$ and $q \in L$ with the diffusion $\mu_0 > 0$. Because we have a variational formulation, we can only interpret this derivative for test functions

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

$\boldsymbol{\phi} \in H$ and $\psi \in L$. We calculate

$$
\begin{aligned}
&\langle G'(\boldsymbol{v},p)(\boldsymbol{w},q),(\boldsymbol{\phi},\psi)\rangle \\
&= \lim_{t\to 0} \frac{\langle G(\boldsymbol{v}+t\boldsymbol{w},p+tq),(\boldsymbol{\phi},\psi)\rangle - \langle G(\boldsymbol{v},p),(\boldsymbol{\phi},\psi\rangle}{t} \\
&= \lim_{t\to 0} \int_\Omega \frac{B}{t}\Big(\big(|D(\boldsymbol{v}+t\boldsymbol{w})|^2+\delta^2\big)^{\frac{1-n}{2n}} D\boldsymbol{v} - \big(|D\boldsymbol{v}|^2+\delta^2\big)^{\frac{1-n}{2n}} D\boldsymbol{v}\Big) : \nabla\boldsymbol{\phi}\, dx \\
&\quad + \lim_{t\to 0} \int_\Omega \frac{B}{t}\Big(\big(|D(\boldsymbol{v}+t\boldsymbol{w})|^2+\delta^2\big)^{(1-n)/(2n)} tD\boldsymbol{w}\Big) : \nabla\boldsymbol{\phi}\, dx \\
&\quad + \lim_{t\to 0} \mu_0 \int_\Omega \nabla\left(\frac{\boldsymbol{v}+t\boldsymbol{w}-\boldsymbol{v}}{t}\right) : \nabla\boldsymbol{\phi}\, dx - \int_\Omega \frac{p+tq-p}{t}\mathrm{div}\boldsymbol{\phi}\, dx - \int_\Omega \mathrm{div}\left(\frac{\boldsymbol{v}+t\boldsymbol{w}-\boldsymbol{v}}{t}\right)\psi\, dx.
\end{aligned} \tag{A1}
$$

The limits for the second and third lines on the right-hand side of the last equality are clear. For the first line, we use the Taylor expansion. Therefore, we define the function $f_x : [0,\infty) \to \mathbb{R}$,

$$
f_x(t) = \big(|D(\boldsymbol{v}(x)+t\boldsymbol{w}(x))|^2+\delta^2\big)^{\frac{1-n}{2n}}. \tag{A2}
$$

Its derivative is

$$
f_x'(t) = \frac{1-n}{n}\big(|D(\boldsymbol{v}(x)+t\boldsymbol{w}(x))|^2+\delta^2\big)^{\frac{1-3n}{2n}} (D\boldsymbol{v}(x) : D\boldsymbol{w}(x) + t|D\boldsymbol{w}(x)|^2). \tag{A3}
$$

We calculate the derivative by assuming we can draw the limes into the integral. A detailed explanation of why we can do this is in Schmidt (2023). We obtain with $\xi : \Omega \to [0,t]$ for the Taylor expansion

$$
\begin{aligned}
&\int_\Omega \lim_{t\to 0} \frac{B}{t}\Big(\big(|D(\boldsymbol{v}+t\boldsymbol{w})|^2+\delta^2\big)^{\frac{1-n}{2n}} D\boldsymbol{v} - \big(|D\boldsymbol{v}|^2+\delta^2\big)^{\frac{1-n}{2n}} D\boldsymbol{v}\Big) : \nabla\boldsymbol{\phi}\, dx \\
&= \int_\Omega \lim_{t\to 0} \frac{B}{t}\big(f_x(t)-f_x(0)\big) D\boldsymbol{v} : \nabla\boldsymbol{\phi}\, dx \\
&= \int_\Omega \lim_{t\to 0} \frac{B}{t} f_x'(\xi(x)) tD\boldsymbol{v} : \nabla\boldsymbol{\phi}\, dx \\
&= \int_\Omega \lim_{t\to 0} B\frac{1-n}{n}\big(|D\boldsymbol{v}(x)+\xi(x)\boldsymbol{w}(x)|^2+\delta^2\big)^{\frac{1-3n}{2n}} (D\boldsymbol{v}(x) : D\boldsymbol{w}(x) + \xi(x)|D\boldsymbol{w}(x)|^2) D\boldsymbol{v}(x) : \nabla\boldsymbol{\phi}(x)\, dx \\
&= \int_\Omega B\frac{1-n}{n}\big(|D\boldsymbol{v}|^2+\delta^2\big)^{\frac{1-3n}{2n}} (D\boldsymbol{v} : D\boldsymbol{w})\,(D\boldsymbol{v} : \nabla\boldsymbol{\phi})\, dx.
\end{aligned} \tag{A4}
$$

Author contributions.

The manuscript is written by NS with contributions and suggestions from TS and AH. The code is implemented by NS with support from TS and AH.

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors appreciate helpful explanations of the ISMIP-HOM experiments from Dr. Martin Rückamp from the Bavarian Academy of Sciences and Humanities, and Dr. Thomas Kleiner from Alfred-Wegener-Institut in Bremerhaven.

https://doi.org/10.5194/egusphere-2023-1569
Preprint. Discussion started: 29 August 2023

References

Chen, Q., Gunzburger, M., and Perego, M.: Well-Posedness Results for a Nonlinear Stokes Problem Arising in Glaciology, SIAM Journal on Mathematical Analysis, 45, 2710–2733, https://doi.org/10.1137/110848694, 2013.

Colinge, J. and Rappaz, J.: A strongly nonlinear problem arising in glaciology, ESAIM: Mathematical Modelling and Numerical Analysis, 33, 395–406, https://doi.org/10.1051/m2an:1999122, 1999.

Fraters, M. R. T., Bangerth, W., Thieulot, C., Glerum, A. C., and Spakman, W.: Efficient and practical Newton solvers for non-linear Stokes systems in geodynamic problems, Geophysical Journal International, 218, 873–894, https://doi.org/10.1093/gji/ggz183, 2019.

Habbal, F., Larour, E., Morlighem, M., Seroussi, H., Borstad, C. P., and Rignot, E.: Optimal numerical solvers for transient simulations of ice flow using the Ice Sheet System Model (ISSM versions 4.2.5 and 4.11), Geoscientific Model Development, 10, 155–168, https://doi.org/10.5194/gmd-10-155-2017, 2017.

Hinze, M., Pinnau, R., Ulbrich, M., and Ulbrich, S.: Optimization with PDE Constraints, Springer Netherlands, https://doi.org/10.1007/978-1-4020-8839-1, 2009.

Hirn, A.: Finite element approximation of singular power-law systems, Mathematics of Computation, 82, 1247–1268, http://www.jstor.org/stable/42002697, 2013.

Logg, A., Mardal, K.-A., and Wells, G., eds.: Automated Solution of Differential Equations by the Finite Element Method, Springer Berlin Heidelberg, https://doi.org/10.1007/978-3-642-23099-8, 2012.

Pattyn, F., Perichon, L., Aschwanden, A., Breuer, B., de Smedt, B., Gagliardini, O., Gudmundsson, G. H., Hindmarsh, R. C. A., Hubbard, A., Johnson, J. V., Kleiner, T., Konovalov, Y., Martin, C., Payne, A. J., Pollard, D., Price, S., Rückamp, M., Saito, F., Souček, O., Sugiyama, S., and Zwinger, T.: Benchmark experiments for higher-order and full-Stokes ice sheet models (ISMIP–HOM), The Cryosphere, 2, 95–108, https://doi.org/10.5194/tc-2-95-2008, 2008.

Schmidt, N.: Global q-superlinear convergence of the infinite-dimensional Newton's method for the regularized p-Stokes equations, preprint, https://arxiv.org/abs/2307.02930, 2023.

Schoof, C.: Coloumb friction and other sliding laws in a higher-order glacier flow model, Mathematical Models and Methods in Applied Sciences, 20, 157–189, https://doi.org/10.1142/s0218202510004180, 2010.