ON THE FINITE ELEMENT APPROXIMATION OF A SEMICOERCIVE STOKES VARIATIONAL INEQUALITY ARISING IN GLACIOLOGY

GONZALO G. DE DIEGO*, PATRICK E. FARRELL*, AND IAN J. HEWITT*

Abstract. Stokes variational inequalities arise in the formulation of glaciological problems involving contact. We consider the problem of a two-dimensional marine ice sheet with a grounding line, although the analysis presented here is extendable to other contact problems in glaciology, such as that of subglacial cavitation. The analysis of this problem and its discretisation is complicated by the nonlinear rheology commonly used for modelling ice, the enforcement of a friction boundary condition given by a power law, and the presence of rigid modes in the velocity space, which render the variational inequality semicoercive. In this work, we consider a mixed formulation of this variational inequality involving a Lagrange multiplier and provide an analysis of its finite element approximation. Error estimates in the presence of rigid modes are obtained by means of a specially-built projection operator onto the subspace of rigid modes and a Korn-type inequality. These proofs rely on the fact that the subspace of rigid modes is at most one-dimensional, a property which is a consequence of the two-dimensionality of the domain. Numerical results are reported to validate the error estimates.

Key words. non-Newtonian Stokes, glaciology, variational inequality, semicoercive, convergence analysis

AMS subject classifications. 65N12, 65N15, 65N30, 86A40

1. Introduction.

We consider the problem of a marine ice sheet resting on a bedrock and sliding into the ocean, where it goes afloat. Such a configuration is found in Greenland and Antarctica, and the dynamics of the grounding line, the point where ice loses contact with the bedrock, is of crucial importance for predicting future sea level rise and comprehending large scale climate dynamics [44, 37, 35, 21]. This contact problem is modelled by coupling a Stokes problem for the ice flow with a time-dependent advection equation for the free surface. At each instant in time, the Stokes equation must be solved with contact boundary conditions that allow the detachment of the ice from the bedrock. These contact conditions transform the instantaneous Stokes problem into a variational inequality. Similar contact conditions appear in a related problem of subglacial cavitation (see Remark 2.1 below), which is also of fundamental importance in glaciology [18, 36, 45], and the results presented in this work are extendable to such problems [10].

Numerous finite element simulations of these equations have been carried out [13, 16, 41]. However, to the best of our knowledge, no formal analysis of these problems and their approximation exist in the mathematical literature. Moreover, we believe that the discretisations used in these computations can be improved upon, by exploiting the structure of the variational inequality. Although the Stokes variational inequality is superficially similar to the elastic contact problem, which has been widely studied [33, 27], the Stokes problem in the context of marine ice sheets with a grounding line includes three substantial difficulties that must be addressed carefully: the presence of rigid body modes in the space of admissible velocities, the nonlinear rheological law used to model ice as a viscous fluid, and the nonlinearity of the friction boundary condition.

In this work we analyse the instantaneous Stokes variational inequality and its approximation. The presence of rigid body modes renders this problem semicoercive.

*Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK (gonzalezdedi@maths.ox.ac.uk). PEF was supported by EPSRC grants EP/V001493/1 and EP/R029423/1.

Although semicoercive variational inequalities have been studied in the past [26, 40, 2], existing analyses use purely indirect arguments which give very limited information on how different meshes and finite elements affect the discretisation. Here, we present a novel constructive approach based on the use of a specially designed projection operator onto the subspace of rigid modes that satisfies a Korn type inequality. Error estimates are obtained for the rigid component of the velocity error by exploiting the fact that the dimension of the subspace of rigid modes is at most one. The nonlinear rheology and friction boundary condition complicate the estimation of errors for the discrete problem. Here, we use the techniques from [4, 29] to establish a convergence analysis.

We propose a mixed formulation of the Stokes variational inequality where a Lagrange multiplier is used to enforce the contact conditions. This formulation permits a structure-preserving discretisation that explicitly enforces a discrete version of the contact conditions, up to rounding errors. This allows for a precise distinction between regions where the ice detaches from the bed and those where it remains attached. This precision is extremely useful when coupling the Stokes variational inequality with the time-dependent advection equation for the free surface. Numerical results with this scheme in the context of subglacial cavitation can be found in [10].

1.1. Outline of the paper. In Section 2, the Stokes variational inequality and its mixed formulation are presented. We prove a Korn-type inequality involving a projection operator onto the subspace of rigid modes that will be used throughout the analysis, and we demonstrate that the mixed formulation is well posed. In Section 3, we analyse a family of finite element approximations of the mixed problem and present error estimates in terms of best approximation results for the velocity, pressure and Lagrange multiplier. Finally, in Section 4, a concrete finite element scheme involving quadratic elements for the velocity and piecewise constant elements for the pressure and the Lagrange multiplier is introduced. We then present error estimates for this scheme and we solve a problem with a manufactured solution to calculate convergence rates and compare these with our estimates.

1.2. Notation. Given two normed vector spaces X and Y and a bounded linear operator $T : X \to Y$, the dual of X is denoted by X' and the dual operator to T by $T' : Y' \to X'$. The range of T is denoted by $\operatorname{Ran} T$ and its kernel by $\operatorname{Ker} T$. The norm in X is denoted by $\|\cdot\|_X$ and the pairing between elements in the primal and dual spaces by $\langle f, x\rangle_X$ for $f \in X'$ and $x \in X$. We will work with the Lebesgue and Sobolev spaces $W^{m,r}(\Omega)$, where $m \geq 0$ and $r \geq 1$, defined as the set of functions with weak derivatives up to order m which are r-integrable. When $m = 0$ we write $L^r(\Omega)$. The space of polynomials of degree k over a simplex E (interval, triangle, tetrahedron) is denoted by $\mathcal{P}_k(E)$. The space of continuous functions over a domain Ω is given by $\mathcal{C}(\Omega)$. Vector-valued functions and vector-valued function spaces will be denoted with bold symbols, e.g. $\boldsymbol{u}$ and $\boldsymbol{W}^{m,r}(\Omega)$. We write $f \sim g$, $f \lesssim g$ and $f \gtrsim g$ if there exist generic constants $c, C > 0$ such that $cf \leq g \leq Cf$, $cf \leq g$, and $cf \geq g$, respectively. Throughout this work, we assume that these generic constants do not depend on the mesh size or on the continuous and discrete solutions of the problem.

2. Formulation of the problem. In this section we introduce the semicoercive variational inequality that arises in the study of marine ice sheets and present its formulation as a mixed problem with a Lagrange multiplier. We then analyse the existence and uniqueness of solutions for the mixed problem.

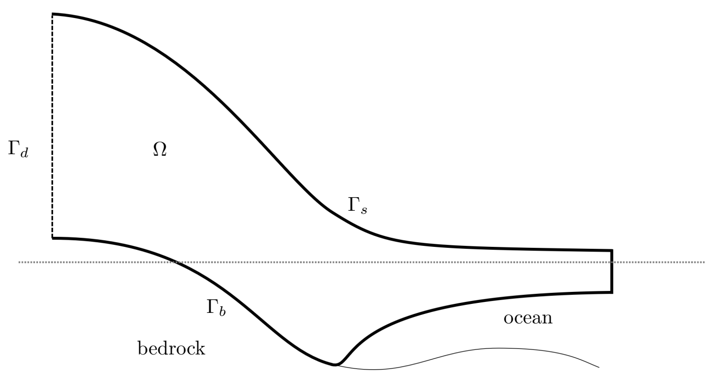

FIG. 1. *Geometry of the problem under consideration. The domain Ω represents a half of a symmetric marine ice sheet. The boundary of Ω is partitioned into Γ_b (the ice-bedrock interface), Γ_s (the ice-ocean and ice-atmosphere interface), and Γ_d (the symmetry axis). The horizontal dotted line represents the sea level.*

2.1. A model for ice flow. We consider a two-dimensional symmetrical marine ice sheet resting on a bedrock and sliding into the ocean. This is the most common configuration considered when studying marine ice sheets [38, 37, 13] and is generally used as a benchmark test case [34]. We denote by $\Omega \subset \mathbb{R}^2$ the domain which represents one half of the ice sheet and we assume it to be connected and polygonal. The latter assumption is made to simplify the analysis, but we expect the essential results presented here to extend to domains with smooth enough boundaries. Ice is generally modelled as a viscous incompressible flow whose motion is described by the Stokes equation [19]:

$$
\begin{aligned}
&\text{(2.1a)} \qquad & -\nabla \cdot (2\eta(|\mathbf{D}\boldsymbol{u}|)\mathbf{D}\boldsymbol{u}) + \nabla p = \boldsymbol{f} \qquad & \text{in } \Omega,\\
&\text{(2.1b)} \qquad & \nabla \cdot \boldsymbol{u} = 0 \qquad & \text{in } \Omega.
\end{aligned}
$$

In the equations above, $\boldsymbol{u} : \Omega \to \mathbb{R}^2$ represents the ice velocity, $p : \Omega \to \mathbb{R}$ the pressure and $\boldsymbol{f} : \Omega \to \mathbb{R}^2$ is a prescribed body force, generally due to gravitational forces. The tensor $\mathbf{D}\boldsymbol{u}$ is the symmetric part of the velocity gradient, that is,

$$\mathbf{D}\boldsymbol{u} = \frac{1}{2}\left(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^\top\right).$$

The coefficient $\eta(|\mathbf{D}\boldsymbol{u}|)$ is the effective viscosity of ice, which relates the stress and strain rates. A power law, usually called Glen's law [23], is the most common choice of rheological law for ice:

$$\eta(|\mathbf{D}\boldsymbol{u}|) = \frac{1}{2}\mathcal{A}^{-1/n}\left(\frac{1}{2}|\mathbf{D}\boldsymbol{u}|^2\right)^{\frac{1-n}{2n}}. \tag{2.2}$$

Here, $|\cdot|$ represents the Frobenius norm of a matrix: for $B \in \mathbb{R}^{m\times m}$ with components B_{ij} we have $|B|^2 = \sum_{ij} B_{ij}$. The field $\mathcal{A} \in L^\infty(\Omega)$ is a prescribed function for which $\operatorname{ess\,inf}\mathcal{A} > 0$. The parameter n is constant and is usually set to $n = 3$; for $n = 1$ we recover the standard linear Stokes flow. From now on, we simply write

$$\eta(|\mathbf{D}\boldsymbol{u}|) = \frac{1}{2}\alpha|\mathbf{D}\boldsymbol{u}|^{r-2}, \tag{2.3}$$

where $\alpha = (1/2)^{(r-2)/2}\mathcal{A}^{1-r}$ is in $L^\infty(\Omega)$ and satisfies $\alpha \geq \alpha_0$ a.e. on Ω for some $\alpha_0 > 0$. Moreover, $r = 1 + 1/n$ is in $(1, 2]$ for $n \geq 1$. This expression for η reveals the r-Stokes nature of the problem when considered as a variational problem in the setting of Sobolev spaces.

2.2. Boundary conditions. For a given velocity and pressure field, we define the stress tensor $\sigma = \sigma(\boldsymbol{u}, p)$ by

$$\sigma = \alpha|\mathbf{D}\boldsymbol{u}|^{r-2}\mathbf{D}\boldsymbol{u} - pI,$$

where $I : \Omega \to \mathbb{R}^{2\times 2}$ is the identity tensor field. Let $\boldsymbol{n}$ denote the unit outward-pointing normal vector to the boundary $\partial\Omega$ and $\mathbf{T} = \mathbf{I} - \boldsymbol{n}\boldsymbol{n}^\top$ the orthogonal projection onto the tangential component to the boundary. We define the normal and tangential stresses at the boundary as

$$\sigma_{nn} = (\sigma\boldsymbol{n}) \cdot \boldsymbol{n} \quad \text{and} \quad \boldsymbol{\sigma}_{nt} = \mathbf{T}\sigma\boldsymbol{n}.$$

The boundary $\partial\Omega$ is partitioned into three disjoint open sets Γ_s, Γ_b and Γ_d of positive measure, see Figure 1. The subset Γ_s represents the part of the boundary in contact with the atmosphere and the ocean. Here we enforce

$$\sigma_{nn} = p_s \quad \text{and} \quad \boldsymbol{\sigma}_{nt} = 0 \quad \text{on } \Gamma_s, \tag{2.4}$$

where $p_s : \Gamma_s \to \mathbb{R}$ represents a prescribed surface traction force. On Γ_b the ice is in contact with the bedrock. Here, we enforce the contact conditions which allow the ice to detach from but not penetrate the bedrock. In particular, detachment can occur if the normal stress equals the subglacial water pressure, which is defined everywhere along a thin lubrication layer in between the ice and the bedrock. We also assume that the ice slides along the bedrock according to a power law. Then, the boundary conditions on Γ_b are given by

$$\boldsymbol{u} \cdot \boldsymbol{n} \leq 0, \quad \sigma_{nn} \leq -p_w \quad \text{and} \quad (\boldsymbol{u} \cdot \boldsymbol{n})\,(\sigma_{nn} + p_w) = 0 \quad \text{on } \Gamma_b, \tag{2.5a}$$

$$\boldsymbol{\sigma}_{nt} = -\tau|\mathbf{T}\boldsymbol{u}|^{r-2}\mathbf{T}\boldsymbol{u} \quad \text{on } \Gamma_b, \tag{2.5b}$$

where $p_w : \Gamma_b \to \mathbb{R}$ is the water pressure at the ice-bedrock interface and $\tau > 0$ a constant. A power-law boundary condition as in (2.5b) was first proposed by Weertman [43] and has since become a popular model for glacier sliding [25, 19].

Finally, Γ_d represents the ice divide of the ice sheet, which is essentially its symmetry axis. As such, it is a vertical surface on which we enforce the symmetry conditions

$$\boldsymbol{u} \cdot \boldsymbol{n} = 0 \quad \text{and} \quad \boldsymbol{\sigma}_{nt} = 0 \quad \text{on } \Gamma_d. \tag{2.6}$$

2.3. The mixed formulation. We now present the mixed formulation whose analysis and approximation is the focus of this work. To do so, we first write (2.1) with boundary conditions (2.4)-(2.6) as a variational inequality. Then, we introduce the mixed formulation by defining a Lagrange multiplier which enforces a constraint that arises due to the contact boundary conditions (2.5a). In Appendix A we specify and prove the sense in which these different formulations are equivalent.

In order to build a weak formulation of (2.1) and (2.4)-(2.6), we must first define suitable function spaces in which to seek the velocity and the pressure. For $r' = 1/(r-1)$, we write

$$V = \left\{\boldsymbol{v} \in \boldsymbol{W}^{1,r}(\Omega) : \boldsymbol{v} \cdot \boldsymbol{n} = 0 \quad \text{on } \Gamma_d\right\}, \quad Q = L^{r'}(\Omega).$$

We denote by $\gamma_n : V \to L^r(\Gamma_b)$ the normal trace operator onto Γ_b. This operator is built by extending to V the operator $\boldsymbol{v} \mapsto \boldsymbol{v} \cdot \boldsymbol{n}$ on Γ_b, defined on smooth functions. The closed convex subset K of V is then defined by

$$K = \{\boldsymbol{v} \in V : \gamma_n \boldsymbol{v} \leq 0 \quad \text{a.e. on } \Gamma_b\}.$$

We also introduce the operators $A : V \to V'$, $G : V \to V'$ and $B : Q \to V'$, defined by

$$\langle A\boldsymbol{u}, \boldsymbol{v}\rangle_V = \int_\Omega \alpha |\mathbf{D}\boldsymbol{u}|^{r-2} (\mathbf{D}\boldsymbol{u} : \mathbf{D}\boldsymbol{v}) \, \mathrm{d}x, \tag{2.7}$$

$$\langle G\boldsymbol{u}, \boldsymbol{v}\rangle_V = \int_{\Gamma_b} \tau |\mathbf{T}\boldsymbol{u}|^{r-2} (\mathbf{T}\boldsymbol{u} \cdot \mathbf{T}\boldsymbol{v}) \, \mathrm{d}s, \tag{2.8}$$

$$\langle Bq, \boldsymbol{v}\rangle = \int_\Omega (\nabla \cdot \boldsymbol{v}) \, q \, \mathrm{d}x. \tag{2.9}$$

Moreover, the action of the applied body and surface forces on the domain Ω is expressed via the function $F \in V'$, defined as

$$\langle F, \boldsymbol{v}\rangle_V = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} \, \mathrm{d}x + \int_{\Gamma_s} p_s (\boldsymbol{v} \cdot \boldsymbol{n}) \, \mathrm{d}s - \int_{\Gamma_b} p_w (\boldsymbol{v} \cdot \boldsymbol{n}) \, \mathrm{d}s. \tag{2.10}$$

In order for (2.10) to make sense, we require $\boldsymbol{f} \in L^{r'}(\Omega)$, $p_s \in L^{r'}(\Gamma_s)$ and $p_w \in L^{r'}(\Gamma_b)$. Then, (2.1) with boundary conditions (2.4)-(2.6) can be reformulated as the variational inequality: find $(\boldsymbol{u}, p) \in K \times Q$ such that

$$\langle A\boldsymbol{u} + G\boldsymbol{u} - Bp - F, \boldsymbol{v} - \boldsymbol{u}\rangle_V + \langle Bq, \boldsymbol{u}\rangle_V \geq 0 \quad \forall (\boldsymbol{v}, q) \in K \times Q. \tag{2.11}$$

In the mixed formulation, the constraint $\boldsymbol{v} \cdot \boldsymbol{n} \leq 0$ on Γ_b is enforced via a Lagrange multiplier. We denote the range of γ_n by Σ and equip this space with the $W^{1-1/r,r}(\Gamma_b)$ norm. We assume the geometry of Ω and Γ_b to be sufficiently regular for this space to be a Banach space, see [33, Section 5], [27, Chapter III] and [1, Chapter 7] for discussions on normal traces and trace spaces. The Lagrange multiplier is sought in the convex cone of multipliers

$$\Lambda = \{\mu \in \Sigma' : \langle \mu, \zeta\rangle_\Sigma \geq 0 \quad \forall \zeta \in \Sigma \text{ s.t. } \zeta \leq 0\}.$$

The equivalent mixed formulation of (2.11) is: find $(\boldsymbol{u}, p, \lambda) \in V \times Q \times \Lambda$ such that

$$\langle A\boldsymbol{u} + G\boldsymbol{u} - Bp - F, \boldsymbol{v}\rangle_V - \langle \lambda, \gamma_n \boldsymbol{v}\rangle_\Sigma = 0 \qquad \forall \boldsymbol{v} \in V, \tag{2.12a}$$

$$\langle Bq, \boldsymbol{u}\rangle_V = 0 \qquad \forall q \in Q, \tag{2.12b}$$

$$\langle \mu - \lambda, \gamma_n \boldsymbol{u}\rangle_\Sigma \geq 0 \qquad \forall \mu \in \Lambda. \tag{2.12c}$$

The Lagrange multiplier λ essentially coincides with $\sigma_{nn} + p_w$ on Γ_b. Indeed, if the solution to (2.12) is sufficiently smooth for integration by parts to hold, we arrive at $\lambda = \sigma_{nn} + p_w$ on Γ_b. Moreover, the conditions $\lambda \in \Lambda$ and (2.12c) are equivalent to

$$\langle \mu, \gamma_n \boldsymbol{u}\rangle_\Sigma \geq 0 \quad \forall \mu \in \Lambda, \quad \lambda \in \Lambda \quad \text{and} \quad \langle \lambda, \gamma_n \boldsymbol{u}\rangle_\Sigma = 0, \tag{2.13}$$

which is a weak representation of the contact boundary conditions (2.5a).

2.4. Well posedness of the mixed formulation. Questions on the existence and uniqueness of solutions of the mixed system (2.12) can be answered by studying an equivalent minimisation problem. This equivalence depends on the so-called inf-sup property holding for the operators B and γ_n. Let

$$V_b = \{\boldsymbol{v} \in V : \gamma_n \boldsymbol{v} = 0 \quad \text{a.e. on } \Gamma_b\}\,. \tag{2.14}$$

These inf-sup conditions can be stated as

$$\sup_{\boldsymbol{v}\in V_b} \frac{\langle Bq, \boldsymbol{v}\rangle_V}{\|\boldsymbol{v}\|_V} \gtrsim \|q\|_Q \quad \forall q \in Q, \tag{2.15}$$

$$\sup_{\boldsymbol{v}\in V} \frac{\langle \mu, \gamma_n \boldsymbol{v}\rangle_\Sigma}{\|\boldsymbol{v}\|_V} \gtrsim \|\mu\|_{\Sigma'} \quad \forall \mu \in \Sigma'. \tag{2.16}$$

Condition (2.15) is proved in [31, Lemma 3.2.7] and (2.16) follows from the inverse mapping theorem because γ_n is surjective onto the Banach space Σ. We also define the space of divergence-free functions $\mathring{V}$ and the convex set $\mathring{K}$ as

$$\mathring{V} = \{\boldsymbol{v} \in V : \nabla\cdot\boldsymbol{v} = 0\} \quad \text{and} \quad \mathring{K} = \mathring{V} \cap K.$$

Then, (2.12) is equivalent to the minimisation of the functional

$$\mathcal{J}(\boldsymbol{v}) = \frac{1}{r}\int_\Omega \alpha|\mathbf{D}\boldsymbol{v}|^r \,\mathrm{d}x + \frac{1}{r}\int_{\Gamma_b} \tau|\mathbf{T}\boldsymbol{v}|^r \,\mathrm{d}s - \langle F, \boldsymbol{v}\rangle_V \tag{2.17}$$

over $\mathring{K}$, see Appendix A. The existence of minimisers of $\mathcal{J}$ over $\mathring{K}$ hinges on whether the set

$$R_V = \left\{\boldsymbol{r} \in V : \int_\Omega |\mathbf{D}\boldsymbol{r}|^r \,\mathrm{d}x + \int_{\Gamma_b} |\mathbf{T}\boldsymbol{r}|^r \,\mathrm{d}s = 0\right\}$$

is equal to or larger than the trivial set $\{0\}$. As shown in [33, Lemma 6.1], the kernel of $\mathbf{D}$ coincides with the set of rigid modes in Ω, defined by

$$R = \left\{\boldsymbol{r} \in \boldsymbol{H}^1(\Omega) : \boldsymbol{r}(x,y) = \begin{pmatrix} a \\ b \end{pmatrix} + \omega \begin{pmatrix} -y \\ x \end{pmatrix}, \quad (a,b,\omega) \in \mathbb{R}^3\right\}.$$

Hence, R_V is the set of rigid modes $\boldsymbol{r} \in R$ satisfying $\mathbf{T}\boldsymbol{r} = 0$ on Γ_b and $\boldsymbol{r}\cdot\boldsymbol{n} = 0$ on Γ_d. For this reason, the dimension of R_V can be at most 1 whenever Γ_b is a flat surface perpendicular to Γ_d. In this case, R_V is given by purely vertical translations.

Remark 2.1. *Although a flat bedrock may appear to be unrealistic, these are considered in many theoretical studies of marine ice sheets [38, 34, 41]. One-dimensional subspaces of rigid modes in V also arise in marine ice sheets which can slide freely ($\tau = 0$) and in the subglacial cavity problem considered in [20, 10] (whenever a horizontal velocity boundary condition is imposed at the top boundary). An analysis of these two problems and their discretisation can be completed using the techniques and steps presented in this paper.*

We define the projection operator $\mathbb{P} : V \to R_V$ by

$$\mathbb{P}(\boldsymbol{v}) = \begin{cases} \dfrac{\int_{\Gamma_b} \boldsymbol{v}\cdot\boldsymbol{n}\,\mathrm{d}s}{\int_{\Gamma_b} \boldsymbol{e}_R\cdot\boldsymbol{n}\,\mathrm{d}s}\boldsymbol{e}_R & \text{if } \dim R_V = 1, \\ 0 & \text{if } \dim R_V = 0, \end{cases}$$

where $\boldsymbol{e}_R \in R_V$ is a basis function that spans R_V when $\dim R_V = 1$. We choose this projection operator because it satisfies $\mathbb{P}(K) \subset K$. The operator $\mathbb{Q} = \mathbb{I} - \mathbb{P}$ then maps elements in V onto a closed subspace whose intersection with R_V is $\{0\}$. As a result, we have the following variation of Korn's inequality:

LEMMA 2.2. *The inequality*

$$\|\mathbb{Q}\boldsymbol{v}\|_V \lesssim \|\mathbf{D}\boldsymbol{v}\|_{L^r(\Omega)} + \|\mathbf{T}\boldsymbol{v}\|_{L^r(\Gamma_b)} \tag{2.18}$$

holds uniformly for all $\boldsymbol{v} \in V$.

Proof. Following the proof of [8, Lemma 3], we first notice that (2.18) follows from

$$\int_\Omega |\boldsymbol{v}|^r \,\mathrm{d}x \lesssim \int_\Omega |\mathbf{D}\boldsymbol{v}|^r \,\mathrm{d}x + \int_{\Gamma_b} |\mathbf{T}\boldsymbol{v}|^r \,\mathrm{d}s \quad \forall \boldsymbol{v} \in \operatorname{Ran}\mathbb{Q} \tag{2.19}$$

due to the generalised Korn inequality [8, Lemma 2]. Since $\operatorname{Ran}\mathbb{Q} \cap R_V = \{0\}$, the proof of (2.18) is completed by assuming (2.19) to be false and mimicking the steps in the proof of [8, Lemma 3]. ∎

Whenever $R_V \neq \{0\}$, (2.12) is semicoercive in the sense that the operator $A + G$ has a nontrivial kernel. In Theorem 2.3 below, we show that a consequence of semicoercivity is that (2.12) will have a solution only when the following compatibility condition holds:

$$\langle F, \boldsymbol{r}\rangle_V < 0 \quad \forall \boldsymbol{r} \in (R_V \cap K) \setminus \{0\}. \tag{2.20}$$

Condition (2.20) allows us to establish the well-posedness of (2.12) and the error estimates, because the restriction of the map $\boldsymbol{r} \mapsto \langle F, \boldsymbol{r}\rangle_V$ to the boundary of the unit ball in $K \cap R_V$ is a continuous map defined over a compact set. Therefore, whenever (2.20) holds, we have the inequality

$$\delta \|\boldsymbol{r}\|_V \leq -\langle F, \boldsymbol{r}\rangle_V \quad \forall \boldsymbol{r} \in R_V \cap K, \tag{2.21}$$

where

$$\delta = \min_{\substack{\boldsymbol{r} \in R_V \cap K,\\ \|\boldsymbol{r}\|_V = 1}} -\langle F, \boldsymbol{r}\rangle_V.$$

Inequality (2.21) is used to prove that the solutions to the continuous and discrete problems are bounded from above in Theorems 2.3 and 3.1 below, respectively. It is also used in the proof of Lemma 3.2 to obtain error estimates for the rigid component of the velocity error.

The importance of the compatibility condition (2.20) is well-known in the study of semicoercive variational inequalities, see [30, 40, 33] in the context of general variational inequalities and [39, 8] in a glaciological setting. The compatibility condition has the geometrical interpretation that the applied force F should have an obtuse angle with the directions of escape of the body given by $R_V \cap K$, which in this case correspond with vertical upward movements whenever Γ_b is flat.

THEOREM 2.3. *If* $R_V = \{0\}$*, then a solution to* (2.12) *exists and is unique. If* $R_V \neq \{0\}$*, then there is a unique solution to* (2.12) *provided the compatibility condition* (2.20) *holds. Conversely, if* $R_V \neq \{0\}$ *and a solution exists, we have that*

$$\langle F, \boldsymbol{r}\rangle_V \leq 0 \quad \forall \boldsymbol{r} \in R_V \cap K. \tag{2.22}$$

Moreover, a solution $(\boldsymbol{u}, p, \lambda) \in V \times Q \times \Lambda$ *of* (2.12) *is bounded from above, i.e.*

$$\|\boldsymbol{u}\|_V + \|p\|_Q + \|\lambda\|_{\Sigma'} \lesssim 1, \tag{2.23}$$

if (2.20) *holds when* $R_V \neq \{0\}$.

Proof. If $R_V = \{0\}$, then $\mathbb{Q} = \mathbb{I}$ and we can establish the coercivity of $\mathcal{J}$ with (2.18). Existence then follows from [15, Theorem 2, Section 8.2] due to the convexity of $\mathcal{J}$ [8]. We may prove existence when $R_V \neq \{0\}$ and (2.20) holds with the method used in [17, 1.II]. If a minimising sequence $(\boldsymbol{u}_n)$ in $\mathring{K}$ contains a bounded subsequence, then we may extract a subsequence that converges weakly to a point $\boldsymbol{u}$. By the weak lower semicontinuity of $\mathcal{J}$, we can then show that $\boldsymbol{u}$ minimises $\mathcal{J}$. We must therefore show that a minimising sequence $(\boldsymbol{u}_n)$ has a bounded subsequence. Assume it does not; we must then have that $\|\boldsymbol{u}_n\|_V \to \infty$. From inequality (2.18) we deduce that

$$\|\mathbb{Q}\boldsymbol{u}_n\|_V^r \lesssim \mathcal{J}(\boldsymbol{u}_n) + \langle F, \boldsymbol{u}_n \rangle_V. \tag{2.24}$$

We define $\boldsymbol{w}_n = \boldsymbol{u}_n / \|\boldsymbol{u}_n\|_V$ and deduce from (2.24) that

$$\|\mathbb{Q}\boldsymbol{w}_n\|_V^r \lesssim \frac{1}{\|\boldsymbol{u}_n\|_V^r}\mathcal{J}(\boldsymbol{u}_n) + \frac{\|F\|_{V^*}}{\|\boldsymbol{u}_n\|_V^{r-1}}.$$

Hence, $\mathbb{Q}\boldsymbol{w}_n \to 0$ as $n \to \infty$. Also, since $\|\boldsymbol{w}\|_V = 1$, we have that $\mathbb{P}\boldsymbol{w}_n$ is bounded in R_V, so there is a subsequence, which we also denote by $(\boldsymbol{w}_n)$, that converges to a $\boldsymbol{r} \in R_V$. In fact, since $\mathbb{Q}\boldsymbol{w}_n \to 0$, we have that $\boldsymbol{w}_n \to \boldsymbol{r}$ and therefore $\boldsymbol{r} \in (\mathring{K} \cap R_V) \setminus \{0\}$. We reach a contradiction when we write (2.24) as

$$\|\boldsymbol{u}_n\|_V^{r-1} \|\mathbb{Q}\boldsymbol{w}_n\|_V^r \lesssim \frac{1}{\|\boldsymbol{u}_n\|_V} (\mathcal{J}(\boldsymbol{u}_n) - \mathcal{J}(0)) + \langle F, \boldsymbol{w}_n \rangle_V.$$

and observe that the lim sup of the left-hand side is strictly positive, while the lim inf of the right-hand side is strictly negative due to the compatibility condition (2.20).

Regarding the uniqueness of solutions, if $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are two minimisers for $\mathcal{J}$ in $\mathring{K}$, it follows that

$$\langle A\boldsymbol{u}_1 - A\boldsymbol{u}_2, \boldsymbol{u}_1 - \boldsymbol{u}_2 \rangle_V + \langle G\boldsymbol{u}_1 - G\boldsymbol{u}_2, \boldsymbol{u}_1 - \boldsymbol{u}_2 \rangle_V \leq 0.$$

Since the operator $A + G$ is monotone, the inequality above must be an equality and using (2.18) we deduce that $\|\mathbb{Q}(\boldsymbol{u}_1 - \boldsymbol{u}_2)\|_V = 0$. Therefore, if $\boldsymbol{u} \in \mathring{K}$ minimises $\mathcal{J}$, any other minimiser must be of the form $\boldsymbol{u} + \boldsymbol{r}$ with $\boldsymbol{r} \in R_V$. If $R_V = \{0\}$, we see that any solution must be unique. On the other hand, if $R_V \neq \{0\}$, since $\mathcal{J}(\boldsymbol{u}+\boldsymbol{r}) = \mathcal{J}(\boldsymbol{u}) + F(\boldsymbol{r})$, the function $\boldsymbol{r} \in R_V$ must also satisfy $F(\boldsymbol{r}) = 0$. Moreover, whenever (2.20) holds, we have that $F(\boldsymbol{r}) = 0$ if and only if $\boldsymbol{r} = 0$ because $\dim R_V = 1$. As a result, $\boldsymbol{r} = 0$ and the solution is unique.

For the converse statement, assume $R_V \neq \{0\}$ and let $(\boldsymbol{u}, p, \lambda)$ solve (2.12). Then, from (2.12a) we deduce that

$$\langle F, \boldsymbol{r} \rangle_V = -\langle \lambda, \gamma_n \boldsymbol{r} \rangle_\Sigma \leq 0 \quad \forall \boldsymbol{r} \in K \cap R_V.$$

To prove (2.23), we first note that (2.12a) and (2.18) lead to

$$\|\mathbb{Q}\boldsymbol{u}\|_V^r \lesssim \langle A\boldsymbol{u} + G\boldsymbol{u}, \boldsymbol{u} \rangle_V = \langle F, \boldsymbol{u} \rangle_V. \tag{2.25}$$

If $R_V = \{0\}$, we have that $\|\boldsymbol{u}\|_V^{r-1} \lesssim \|F\|_{V^*}$. If $R_V \neq \{0\}$ and (2.20) holds, then $\langle F, \mathbb{P}\boldsymbol{u}\rangle_V \leq 0$ and we find that $\|\mathbb{Q}\boldsymbol{u}\|_V^{r-1} \lesssim \|F\|_{V^*}$. By using the inf-sup conditions and Hölder's inequality, we can establish the bounds

$$\|\lambda\|_{\Sigma'} + \|p\|_Q \lesssim \|\mathbb{Q}\boldsymbol{u}\|_V^{r-1} + \|F\|_{V^*} \lesssim \|F\|_{V^*}\,. \tag{2.26}$$

We then use (2.21) to show that

$$\|\mathbb{P}\boldsymbol{u}\|_V \lesssim -\langle F, \mathbb{P}\boldsymbol{u}\rangle_V = \langle \lambda, \gamma_n(\mathbb{Q}\boldsymbol{u})\rangle_\Sigma \lesssim \|F\|_{V^*}^{r'}\,. \tag{2.27}$$

We finally establish the bound (2.23) by putting together (2.25), (2.26), and (2.27), and noting that $\|\boldsymbol{u}\|_V \leq \|\mathbb{P}\boldsymbol{u}\|_V + \|\mathbb{Q}\boldsymbol{u}\|_V$. ∎

3. Abstract discretisation. In this section we propose an abstract discretisation of the mixed system (2.12) built in terms of a collection of finite dimensional spaces satisfying certain key properties. We can then introduce a discrete system analogous to (2.12) and investigate the conditions under which we have a unique solution. Then, we prove Lemmas 3.2, 3.4, and 3.6, which establish upper bounds for the errors of the discrete solutions.

3.1. The discrete mixed formulation. For each parameter $h > 0$, let $V_h \subset V$, $Q_h \subset Q$ and $\Sigma_h \subset L^2(\Gamma_b)$ be finite dimensional subspaces. We also assume that $R_V \subset V_h$ to avoid the need of introducing discrete compatibility conditions. We define the discrete convex sets

$$\Lambda_h = \{\mu_h \in \Sigma_h : \mu_h \leq 0 \text{ on } \Gamma_b\}\,,$$

and

$$K_h = \{\boldsymbol{v}_h \in V_h : \langle \mu_h, \gamma_n \boldsymbol{v}_h\rangle_\Sigma \geq 0 \quad \forall \mu_h \in \Lambda_h\}\,.$$

An immediate consequence of the definitions of Λ_h and K_h is that $\Lambda_h \subset \Lambda$ but $K_h \not\subset K$ unless $\gamma_n(V_h) \subset \Sigma_h$. By the assumption $R_V \subset V_h$ and the fact that R_V is given by purely vertical translations whenever $R_V \neq \{0\}$, we have that $K \cap R_V = K_h \cap R_V$.

The discrete analogue of the variational inequality (2.11) is: find $(\boldsymbol{u}_h, p_h) \in K_h \times Q_h$ such that

$$\langle A\boldsymbol{u}_h + G\boldsymbol{u}_h - Bp_h - F, \boldsymbol{v}_h - \boldsymbol{u}_h\rangle_V + \langle Bq_h, \boldsymbol{u}_h\rangle_V \geq 0 \quad \forall (\boldsymbol{v}_h, q_h) \in K_h \times Q_h. \tag{3.1}$$

This discrete variational inequality can be written as a mixed problem by introducing a Lagrange multiplier. This results in the discrete mixed formulation that is the counterpart of (2.12): find $(\boldsymbol{u}_h, p_h, \lambda_h) \in V_h \times Q_h \times \Lambda_h$ such that

$$\langle A\boldsymbol{u}_h + G\boldsymbol{u}_h - Bp_h - F, \boldsymbol{v}_h\rangle_V - \langle \lambda_h, \gamma_n \boldsymbol{v}_h\rangle_\Sigma = 0 \qquad \forall \boldsymbol{v}_h \in V_h, \tag{3.2a}$$

$$\langle Bq_h, \boldsymbol{u}_h\rangle_V = 0 \qquad \forall q_h \in Q_h, \tag{3.2b}$$

$$\langle \mu_h - \lambda_h, \gamma_n \boldsymbol{u}_h\rangle_\Sigma \geq 0 \qquad \forall \mu_h \in \Sigma_h. \tag{3.2c}$$

An advantage of using a mixed formulation at the discrete level is that we explicitly enforce a discrete version of the contact conditions (2.5a). Just as in (2.13), it is possible to show that the conditions $\lambda_h \in \Lambda_h$ and (3.2c) are equivalent to

$$\langle \mu_h, \gamma_n \boldsymbol{u}_h\rangle_\Sigma \geq 0 \quad \forall \mu_h \in \Lambda_h, \quad \lambda_h \in \Lambda_h \quad \text{and} \quad \langle \lambda_h, \gamma_n \boldsymbol{u}_h\rangle_\Sigma = 0. \tag{3.3}$$

In order to state a minimisation problem equivalent to (3.2), we must introduce the subspace of V_h of discretely divergence-free functions and the discrete convex set $\mathring{K}_h$:

$$\mathring{V}_h = \{\boldsymbol{v}_h \in V_h : b(\boldsymbol{v}_h, q_h) = 0 \quad \forall q_h \in Q_h\} \quad \text{and} \quad \mathring{K}_h = \mathring{V}_h \cap K_h.$$

Then, the discrete mixed problem (3.2) is equivalent to the minimisation over $\mathring{K}_h$ of the functional $\mathcal{J} : V \to \mathbb{R}$ defined in (2.17), provided that two discrete inf-sup conditions hold. For $V_{b,h} = V_h \cap V_b$, these discrete conditions can be stated as

$$\sup_{\boldsymbol{v}_h \in V_{b,h}} \frac{\langle Bq_h, \boldsymbol{v}_h\rangle_V}{\|\boldsymbol{v}_h\|_V} \gtrsim \|q_h\|_Q \quad \forall q_h \in Q_h, \tag{3.4}$$

$$\sup_{\boldsymbol{v}_h \in V_h} \frac{\langle \mu_h, \gamma_n \boldsymbol{v}_h\rangle_\Sigma}{\|\boldsymbol{v}_h\|_V} \gtrsim \|\mu_h\|_\Sigma \quad \forall \mu_h \in \Sigma_h. \tag{3.5}$$

When the conditions (3.4) and (3.5) hold, then (3.1), (3.2) and the minimisation of $\mathcal{J}$ over $\mathring{K}$ are equivalent problems. The proofs for such equivalences require the same arguments as the proofs presented in Appendix A. If $\mathcal{J}$ admits a unique minimiser over $\mathring{K}_h$, the discrete inf-sup conditions guarantee a unique solution for (3.2) and set constraints on the choice of spaces V_h, Q_h and Σ_h used when approximating solutions of (2.12). As in the continuous case, the well-posedness of (3.2) requires the compatibility condition (2.20) to hold. The theorem below can be proved in the same way as Theorem 2.3.

Theorem 3.1. *Assume that the discrete inf-sup conditions* (3.4) *and* (3.5) *hold. If $R_V = \{0\}$, then a solution to* (3.2) *exists and is unique. If $R_V \neq \{0\}$, then there is a unique solution to* (3.2) *if the compatibility condition* (2.20) *holds. Conversely, if $R_V \neq \{0\}$ and a solution exists,* (2.22) *must hold. The solution of* (2.12) *is bounded from above independently of h, provided* (2.20) *holds when $R_V \neq \{0\}$.*

3.2. Upper bounds for the velocity error. An important tool presented in [4, 29] for establishing error estimates for non-Newtonian flows is the use of the function $\mathbf{F}$. Here, for ease of notation, we denote by $\mathbf{F}$ an operator that acts on both $\mathbb{R}^{2\times 2}$ and $\mathbb{R}^2$ by

$$\mathbf{F}(\mathrm{A}) = |\mathrm{A}|^{\frac{r-2}{2}}\mathrm{A} \quad \text{for } \mathrm{A} \in \mathbb{R}^{2\times 2} \text{ or } \mathrm{A} \in \mathbb{R}^2. \tag{3.6}$$

This operator is closely related to the operators A and G. Let the operator $\mathrm{E} : V \times V \to \mathbb{R}$ be given by

$$\mathrm{E}(\boldsymbol{u}, \boldsymbol{v}) = \|\mathbf{F}(\mathbf{D}\boldsymbol{u}) - \mathbf{F}(\mathbf{D}\boldsymbol{v})\|^2_{L^2(\Omega)} + \|\mathbf{F}(\mathbf{T}\boldsymbol{u}) - \mathbf{F}(\mathbf{T}\boldsymbol{v})\|^2_{L^2(\Gamma_b)}.$$

We then have that

$$\mathrm{E}(\boldsymbol{u}, \boldsymbol{v}) \sim \langle A\boldsymbol{u} - A\boldsymbol{v}, \boldsymbol{u} - \boldsymbol{v}\rangle_V + \langle G\boldsymbol{u} - G\boldsymbol{v}, \boldsymbol{u} - \boldsymbol{v}\rangle_V \tag{3.7}$$

for all $\boldsymbol{u}, \boldsymbol{v} \in V$. The following variation of Young's inequality,

$$\langle A\boldsymbol{u} - A\boldsymbol{v}, \boldsymbol{u} - \boldsymbol{w}\rangle_V + \langle G\boldsymbol{u} - G\boldsymbol{v}, \boldsymbol{u} - \boldsymbol{w}\rangle_V \leq \varepsilon \mathrm{E}(\boldsymbol{u}, \boldsymbol{v}) + c_\varepsilon \mathrm{E}(\boldsymbol{u}, \boldsymbol{w}), \tag{3.8}$$

is valid for any $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ and $\varepsilon > 0$, with the constant $c_\varepsilon > 0$ depending on ε. Additionally, the inequalities

$$\|\mathbf{D}\boldsymbol{v} - \mathbf{D}\boldsymbol{w}\|^2_{L^r(\Omega)} \lesssim \|\mathbf{F}(\mathbf{D}\boldsymbol{v}) - \mathbf{F}(\mathbf{D}\boldsymbol{w})\|^2_{L^2(\Omega)} \||\mathbf{D}\boldsymbol{v}| + |\mathbf{D}\boldsymbol{w}|\|^{2-r}_{L^r(\Omega)}, \tag{3.9}$$

$$\|\mathbf{T}\boldsymbol{v} - \mathbf{T}\boldsymbol{w}\|^2_{L^r(\Gamma_b)} \lesssim \|\mathbf{F}(\mathbf{T}\boldsymbol{v}) - \mathbf{F}(\mathbf{T}\boldsymbol{w})\|^2_{L^2(\Gamma_b)} \||\mathbf{T}\boldsymbol{v}| + |\mathbf{T}\boldsymbol{w}|\|^{2-r}_{L^r(\Gamma_b)}, \tag{3.10}$$

hold for any $\boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{W}^{1,r}(\Omega)$. A proof for inequalities (3.7) and (3.9) can be found in [29, Lemmas 2.3, 2.4], and [4, Lemma 2.7] for (3.8), for the case without friction. The presence of the operator G requires a version of [29, Lemmas 2.3, 2.4] and [4, Lemma 2.7] stated in terms of vectors in $\mathbb{R}^d$. Since these results are based on algebraic inequalities for matrices, the extension to vectors in $\mathbb{R}^d$ can be proved by considering diagonal matrices.

By applying the triangle inequality, (2.18) and (3.9)-(3.10), the velocity error can be decomposed into two components as

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_V \lesssim \|\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h)\|_V + \mathrm{E}(\boldsymbol{u}, \boldsymbol{u}_h). \tag{3.11}$$

For the first term on the right of (3.11), which represents the rigid component of the error, we present the following result:

Lemma 3.2. *Assume that $R_V \neq \{0\}$ and that the compatibility condition (2.20) holds. Let $(\boldsymbol{u}, p, \lambda) \in V \times Q \times \Lambda$ be the solution to (2.12) and $(\boldsymbol{u}_h, p_h, \lambda_h) \in V_h \times Q_h \times \Lambda_h$ to (3.2). Then,*

$$\|\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h)\|_V \lesssim \mathrm{E}(\boldsymbol{u}, \boldsymbol{u}_h) + \|\lambda - \mu_h\|_{\Sigma'} \quad \forall\, \mu_h \in \Lambda_h. \tag{3.12}$$

Proof. Under the assumption that $R_V \neq \{0\}$, we either have that $\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h) \in R_V \cap K$ or $-\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h) \in R_V \cap K$. If $\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h) \in R_V \cap K$, then inequality (2.21) and the continuous mixed system (2.12) allow us to write

$$\|\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h)\|_V \lesssim -\langle F, \mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h)\rangle_V = \langle \lambda, \gamma_n(\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h))\rangle_\Sigma,$$

where the equality follows from $\langle A\boldsymbol{u} + G\boldsymbol{u} - Bp, \mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h)\rangle_V = 0$. Then, by noting that $\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h) = -\mathbb{Q}(\boldsymbol{u} - \boldsymbol{u}_h) + \boldsymbol{u} - \boldsymbol{u}_h$ and $\langle \lambda, \gamma_n \boldsymbol{u}\rangle_\Sigma = 0$, using inequalities (2.18) and (3.9)-(3.10), and using the uniform in h boundedness of solutions to (3.2) (see Theorem 3.1), we arrive at

$$\|\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h)\|_V \lesssim \mathrm{E}(\boldsymbol{u}, \boldsymbol{u}_h) - \langle \lambda, \gamma_n \boldsymbol{u}_h\rangle_\Sigma. \tag{3.13}$$

On the other hand, if $-\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h) \in R_V \cap K$, then, by appealing to (2.20) and the discrete mixed system (3.2),

$$\|\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h)\|_V \lesssim \langle F, \mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h)\rangle_V = -\langle \lambda_h, \gamma_n(\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h))\rangle_\Sigma.$$

Following the same steps as before, we deduce that

$$\|\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h)\|_V \lesssim \mathrm{E}(\boldsymbol{u}, \boldsymbol{u}_h) - \langle \lambda_h, \gamma_n \boldsymbol{u}\rangle_\Sigma \leq \mathrm{E}(\boldsymbol{u}, \boldsymbol{u}_h), \tag{3.14}$$

where the final inequality follows from the fact that $\lambda_h \leq 0$ on Γ_b by the definition of Λ_h. As a result of (3.13) and (3.14), we have that

$$\|\mathbb{P}(\boldsymbol{u} - \boldsymbol{u}_h)\|_V \lesssim \mathrm{E}(\boldsymbol{u}, \boldsymbol{u}_h) + \max\{0, -\langle \lambda, \gamma_n \boldsymbol{u}_h\rangle_\Sigma\}$$

in all cases. Finally, given a $\mu_h \in \Lambda_h$, we have

$$-\langle \lambda, \gamma_n \boldsymbol{u}_h\rangle_\Sigma \leq \langle \mu_h - \lambda, \gamma_n \boldsymbol{u}_h\rangle \lesssim \|\lambda - \mu_h\|_{\Sigma'}$$

because $\langle \mu_h, \gamma_n \boldsymbol{u}_h\rangle_\Sigma \geq 0$. ∎

REMARK 3.3. *As mentioned in the introduction, previous analyses of finite element approximations of semicoercive variational inequalities either only consider the error in a seminorm [30] or use indirect arguments to prove the convergence of the approximate solution in the complete norm [42, 26, 40, 2, 7]. In these cases, arguments by contradiction involving a sequence of triangulations are used. In Lemma 3.2, on the other hand, we provide a fully constructive proof for bounding the rigid component of the velocity error from above. This result is a key ingredient in obtaining the error estimates for the finite element scheme presented in the next section. The proof of Lemma 3.2 relies on* $\dim R_V \leq 1$, *which holds for almost all Stokes variational inequalities considered in glaciology [13, 34, 16, 41, 10].*

The second term on the right of (3.11) can be bounded from above by using the properties of the operator E.

LEMMA 3.4. *Let the triples* $(\boldsymbol{u}, p, \lambda) \in V \times Q \times \Lambda$ *and* $(\boldsymbol{u}_h, p_h, \lambda_h) \in V_h \times Q_h \times \Lambda_h$ *be solutions to* (2.12) *and* (3.2)*, respectively. Then*

$$\mathrm{E}(\boldsymbol{u}, \boldsymbol{u}_h) \lesssim \mathrm{E}(\boldsymbol{u}, \boldsymbol{v}_h) + \|p - q_h\|_Q^2 + \langle \lambda - \lambda_h, \gamma_n(\boldsymbol{v}_h - \boldsymbol{u}_h)\rangle_\Sigma \tag{3.15}$$

holds for all $(\boldsymbol{v}_h, q_h) \in \mathring{V}_h \times Q_h$.

Proof. From (2.12a) and (3.2a), we see that, for any $(\boldsymbol{v}_h, q_h) \in \mathring{V}_h \times Q_h$, we have

$$\begin{aligned} &\langle A\boldsymbol{u} - A\boldsymbol{u}_h, \boldsymbol{u} - \boldsymbol{u}_h\rangle_V + \langle G\boldsymbol{u} - G\boldsymbol{u}_h, \boldsymbol{u} - \boldsymbol{u}_h\rangle_V = \\ &\qquad \langle A\boldsymbol{u} - A\boldsymbol{u}_h, \boldsymbol{u} - \boldsymbol{v}_h\rangle_V + \langle G\boldsymbol{u} - G\boldsymbol{u}_h, \boldsymbol{u} - \boldsymbol{v}_h\rangle_V \\ &\quad + \langle B(p - q_h), \boldsymbol{v}_h - \boldsymbol{u}_h\rangle_V + \langle \lambda - \lambda_h, \gamma_n(\boldsymbol{v}_h - \boldsymbol{u}_h)\rangle_\Sigma \end{aligned}$$

Using (3.7) and (3.8)

$$\begin{aligned} \mathrm{E}(\boldsymbol{u}, \boldsymbol{u}_h) \lesssim\, & \varepsilon_1 \mathrm{E}(\boldsymbol{u}, \boldsymbol{u}_h) + c_{\varepsilon_1}\mathrm{E}(\boldsymbol{v}, \boldsymbol{v}_h) \\ & + \langle B(p - q_h), \boldsymbol{v}_h - \boldsymbol{u}_h\rangle_V + \langle \lambda - \lambda_h, \gamma_n(\boldsymbol{v}_h - \boldsymbol{u}_h)\rangle_\Sigma \end{aligned}$$

for an arbitrary $\varepsilon_1 > 0$. Additionally, by using Young's inequality,

$$\langle B(p - q_h), \boldsymbol{v}_h - \boldsymbol{u}_h\rangle_V \lesssim c_{\varepsilon_2} \|p - q_h\|_Q^2 + \varepsilon_2 \|\mathbf{D}(\boldsymbol{u}_h - \boldsymbol{v}_h)\|_{L^r(\Omega)}^2$$

for any $\varepsilon_2 > 0$. Then, via (3.9), and by setting ε_1 and ε_2 sufficiently small, inequality (3.15) is established. ∎

REMARK 3.5. *If the pair* $V_h \times Q_h$ *is divergence free in the sense that* $\langle Bq_h, \boldsymbol{w}_h\rangle_V = 0$ *for all* $q_h \in Q_h$ *implies that* $\nabla \cdot \boldsymbol{w}_h = 0$*, then the term* $\|p - q_h\|_Q^2$ *in inequality* (3.15) *can be removed.*

3.3. Upper bounds for the pressure and Lagrange multiplier errors. We finalise the analysis of the abstract discretisation by bounding the errors for the pressure and the Lagrange multiplier from above.

LEMMA 3.6. *Assume that the discrete inf-sup conditions* (3.4) *and* (3.5) *hold. Let* $(\boldsymbol{u}, p, \lambda) \in V \times Q \times \Lambda$ *be the solution of* (2.12) *and* $(\boldsymbol{u}_h, p_h, \lambda_h) \in V_h \times Q_h \times \Lambda_h$ *of* (3.2)*. Then*

$$\|p - p_h\|_Q \lesssim \mathrm{E}(\boldsymbol{u}, \boldsymbol{u}_h)^{1/r'} + \|p - q_h\|_Q, \tag{3.16}$$

$$\|\lambda - \lambda_h\|_{\Sigma'} \lesssim \mathrm{E}(\boldsymbol{u}, \boldsymbol{u}_h)^{1/r'} + \|p - q_h\|_Q + \|\lambda - \mu_h\|_{\Sigma'}, \tag{3.17}$$

for all $q_h \in Q_h$ *and* $\mu_h \in \Sigma_h$.

Proof. Since Q_h and Σ_h are subsets of Q and Σ respectively, we can obtain the following equality from (2.12a) and (3.2a):

$$\begin{aligned} \langle A\boldsymbol{u} - A\boldsymbol{u}_h, \boldsymbol{v}_h\rangle_V + \langle G\boldsymbol{u} - G\boldsymbol{u}_h, \boldsymbol{v}_h\rangle_V &= \\ \langle B(p - p_h), \boldsymbol{v}_h\rangle_V + \langle \lambda - \lambda_h, \gamma_n \boldsymbol{v}_h\rangle_\Sigma &\quad \forall \boldsymbol{v}_h \in V_h. \end{aligned} \tag{3.18}$$

The inf-sup condition (3.4) for the pressure space holds over the space $V_{b,h} \subset V_h$ of vector fields with a normal component vanishing on Γ_b. For $\boldsymbol{v}_h \in V_{b,h}$, from equation (3.18) we derive

$$\langle B(p_h - q_h), \boldsymbol{v}_h\rangle_V = \langle A\boldsymbol{u} - A\boldsymbol{u}_h, \boldsymbol{v}_h\rangle_V + \langle B(p - q_h), \boldsymbol{v}_h\rangle_V.$$

From the inf-sup condition (3.4) it follows that

$$\|p_h - q_h\|_Q \lesssim \sup_{\boldsymbol{v}_h \in V_{b,h}} \left(\frac{\langle A\boldsymbol{u} - A\boldsymbol{u}_h, \boldsymbol{v}_h\rangle_V}{\|\boldsymbol{v}_h\|_V} \right) + \|p - q_h\|_Q . \tag{3.19}$$

Then, by Hölder's inequality and [29, Lemma 2.4], we have that

$$\sup_{\boldsymbol{v}_h \in V_{b,h}} \left(\frac{\langle A\boldsymbol{u} - A\boldsymbol{u}_h, \boldsymbol{v}_h\rangle_V}{\|\boldsymbol{v}_h\|_V} \right) \lesssim \|\mathbf{F}(\mathbf{D}\boldsymbol{u}) - \mathbf{F}(\mathbf{D}\boldsymbol{u}_h)\|^{2/r'}_{L^2(\Omega)} . \tag{3.20}$$

Finally, (3.16) follows by applying the triangle inequality to $\|p - p_h\|_Q$ and using (3.19) and (3.20). The bound (3.17) follows in the same way. However, in this case, the inf-sup condition (3.5) is set over the whole space V_h, so the friction term in (3.18) does not vanish. Following the argument used in [29, Lemma 2.4], we can show that

$$\left\| |\mathbf{T}\boldsymbol{u}|^{r-2}\mathbf{T}\boldsymbol{u} - |\mathbf{T}\boldsymbol{u}_h|^{r-2}\mathbf{T}\boldsymbol{u}_h \right\|_{L^2(\Gamma_b)} \lesssim \|\mathbf{F}(\mathbf{T}\boldsymbol{u}) - \mathbf{F}(\mathbf{T}\boldsymbol{u}_h)\|^{2/r'}_{L^2(\Gamma_b)}$$

and therefore

$$\sup_{\boldsymbol{v}_h \in V_h} \left(\frac{\langle G\boldsymbol{u} - G\boldsymbol{u}_h, \boldsymbol{v}_h\rangle_V}{\|\boldsymbol{v}_h\|_V} \right) \lesssim \|\mathbf{F}(\mathbf{T}\boldsymbol{u}) - \mathbf{F}(\mathbf{T}\boldsymbol{u}_h)\|^{2/r'}_{L^2(\Gamma_b)} . \qquad \square$$

Lemmas 3.2, 3.4, and 3.6 give discretisation error estimates in terms of best approximation results. To derive a convergence result, we require bounds on these best approximations. We discuss this in the context of a finite element discretisation in the next section.

4. A finite element scheme.

We now consider a particular finite element discretisation of the mixed problem (2.12). We introduce a non-degenerate (in the sense of [6, Definition 4.4.13]) sequence of triangulations $\mathcal{T}_h$, where $h > 0$ denotes the maximum cell diameter in $\mathcal{T}_h$. The set of edges in $\mathcal{T}_h$ is denoted by $\mathcal{E}_h$; we assume that every edge $e \in \mathcal{E}_h$ in $\partial\Omega$ is either in $\overline{\Gamma}_s$, $\overline{\Gamma}_b$ or $\overline{\Gamma}_d$. We write $\mathcal{E}_h(\Gamma_s)$ and $\mathcal{E}_h(\Gamma_b)$ to denote the edges in $\overline{\Gamma}_s$ and $\overline{\Gamma}_b$ respectively. Associated to each $\mathcal{T}_h$ are the finite element spaces V_h, Q_h and Σ_h, defined by

$$V_h = \{\boldsymbol{v}_h \in \boldsymbol{\mathcal{C}}(\Omega) : \boldsymbol{v}_h|_c \in \boldsymbol{\mathcal{P}}_2(c) \quad \forall c \in \mathcal{T}, \quad \boldsymbol{v}_h \cdot \boldsymbol{n} = 0 \quad \text{on } \Gamma_d\}, \tag{4.1a}$$

$$Q_h = \left\{q_h \in L^2(\Omega) : q_h|_c \in \mathcal{P}_0(c) \quad \forall c \in \mathcal{T}\right\}, \tag{4.1b}$$

$$\Sigma_h = \left\{\mu_h \in L^2(\Gamma_b) : \mu_h|_e \in \mathcal{P}_0(e) \quad \forall e \in \mathcal{E}(\Gamma_b)\right\}. \tag{4.1c}$$

4.1. Analysis of the scheme. The first step in analysing this discretisation is to investigate whether the discrete mixed problem (3.2) is well-posed for this choice of $V_h \times Q_h \times \Sigma_h$, subject to the compatibility condition (2.20). Specifically, we must verify the discrete inf-sup conditions (3.4) and (3.5). The pair $V_h \times Q_h$ is well-known to satisfy (3.4), see [5, Proposition 8.4.3] for the case of $r = 2$; the general case $r \in [1, \infty]$ follows from the same arguments by using the interpolation operator π_V discussed in Appendix B.2. A proof for (3.5) is presented below using a similar argument to the one presented in [9, Proposition 3.3].

LEMMA 4.1. *The finite element pair V_h and Σ_h defined in (4.1) is inf-sup stable in the sense of (3.5).*

Proof. Let $\mu_h \in \Sigma_h$. By the Hahn-Banach theorem, there is a $\psi \in \Sigma$ such that $\langle \mu_h, \psi \rangle_\Sigma = \|\mu_h\|_{\Sigma'}$ and $\|\psi\|_\Sigma = 1$. In Appendix B.3 we construct an extension operator $\Phi : \Sigma \to V_h$ which is bounded uniformly with respect to h, i.e. $\|\Phi\psi\|_V \lesssim \|\psi\|_\Sigma$, and

$$\langle \mu_h, \gamma_n(\Phi\psi) \rangle_\Sigma = \langle \mu_h, \psi \rangle_\Sigma \quad \forall \mu_h \in \Sigma_h,$$

for all $\psi \in \Sigma$. Then

$$\|\mu_h\|_{\Sigma'} = \frac{\langle \mu_h, \psi \rangle_\Sigma}{\|\psi\|_\Sigma} \lesssim \frac{\langle \mu_h, \gamma_n(\Phi\psi) \rangle_\Sigma}{\|\Phi\psi\|_V} \leq \sup_{\boldsymbol{v}_h \in V_h} \frac{\langle \mu_h, \gamma_n \boldsymbol{v}_h \rangle_\Sigma}{\|\boldsymbol{v}_h\|_V}$$

and the result follows. ◻

We end this section with a discussion on the approximability of the mixed system (2.12). We show that the approximate solutions $(\boldsymbol{u}_h, p_h, \lambda_h)$ converge to the exact solutions of (2.12) as $h \to 0$ under a regularity condition, and we establish a rate of convergence for these approximations.

THEOREM 4.2. *Assume that the compatibility condition (2.20) holds whenever $R_V \neq \{0\}$. Let the triple $(\boldsymbol{u}, p, \lambda) \in V \times Q \times \Lambda$ be the solution to (2.12) and $(\boldsymbol{u}_h, p_h, \lambda_h) \in V_h \times Q_h \times \Lambda_h$ to (3.2). Additionally, assume that $(\boldsymbol{u}, p, \lambda) \in W^{2,r}(\Omega) \times W^{1,r'}(\Omega) \times W^{1-1/r',r'}(\Gamma_b)$ and $\mathbf{F}(\mathbf{D}\boldsymbol{u}) \in \boldsymbol{W}^{1,2}(\Omega)$ and $\mathbf{F}(\mathbf{T}\boldsymbol{u}) \in \boldsymbol{W}^{1,2}(\Gamma_b)$. Then*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_V \lesssim h, \tag{4.2a}$$

$$\|p - p_h\|_Q + \|\lambda - \lambda_h\|_{\Sigma'} \lesssim h^{2/r'} + h. \tag{4.2b}$$

Proof. We proceed by first finding a suitable upper bound for the term involving the Lagrange multiplier in (3.15). Since $\langle \lambda_h, \gamma_n \boldsymbol{u}_h \rangle_\Sigma = 0$ and $\langle \mu_h, \gamma_n \boldsymbol{v}_h \rangle_\Sigma \geq 0$ for all $(\boldsymbol{v}_h, \mu_h) \in K_h \times \Lambda_h$, one can show that

$$\begin{aligned} \langle \lambda - \lambda_h, \gamma_n(\boldsymbol{v}_h - \boldsymbol{u}_h) \rangle_\Sigma \leq{}& \langle \lambda - \mu_h, \gamma_n(\boldsymbol{v}_h - \boldsymbol{u}) \rangle_\Sigma \\ &+ \langle \lambda - \mu_h, \gamma_n(\boldsymbol{u} - \boldsymbol{u}_h) \rangle_\Sigma + \langle \mu_h, \gamma_n \boldsymbol{v}_h \rangle_\Sigma \end{aligned} \tag{4.3}$$

for all $(\boldsymbol{v}_h, \mu_h) \in K_h \times \Lambda_h$. By applying Young's inequality in (4.3) and using inequalities (3.11), (3.12), and (3.15), we arrive at

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_V \lesssim \mathrm{E}(\boldsymbol{u}, \boldsymbol{v}_h)^{1/2} + \|p - q_h\|_Q + \|\lambda - \mu_h\|_{\Sigma'} + \langle \mu_h, \gamma_n \boldsymbol{v}_h \rangle_\Sigma \tag{4.4}$$

for all $(\boldsymbol{v}_h, q_h, \mu_h) \in \mathring{K}_h \times Q_h \times \Lambda_h$.

Let $\pi_V : V \to V_h$ by the interpolation operator introduced in Appendix B.2. Additionally, let $\pi_Q : Q \to Q_h$ and $\pi_\Sigma : L^2(\Gamma_b) \to \Sigma_h$ be standard interpolation operators onto the space of piecewise-constant functions. We refer to [14] and the results in Appendix B.1 for proofs of optimal interpolation error estimates in the Q and Σ' norms. From the properties of these interpolation operators it follows that $\pi_V \boldsymbol{u} \in \mathring{K}_h$ and $\pi_\Sigma \lambda \in \Lambda_h$. Additionally, we have that $\int_e \boldsymbol{u}\cdot\boldsymbol{n}\,\mathrm{d}s = \int_e (\pi_V \boldsymbol{u})\cdot\boldsymbol{n}\,\mathrm{d}s$ for all $e \in \mathcal{E}$, so

$$\langle \pi_\Sigma \lambda, \gamma_n \pi_V \boldsymbol{u}\rangle_\Sigma = \langle \lambda, \pi_\Sigma(\gamma_n \boldsymbol{u})\rangle_\Sigma = \langle \lambda, \pi_\Sigma(\gamma_n \boldsymbol{u}) - \gamma_n \boldsymbol{u}\rangle_\Sigma.$$

Since $\langle \pi_\Sigma\lambda, \pi_\Sigma(\gamma_n \boldsymbol{u}) - \gamma_n\boldsymbol{u}\rangle_\Sigma = 0$, we have that

$$\langle \pi_\Sigma \lambda, \gamma_n \pi_V \boldsymbol{u}\rangle_\Sigma = \langle \lambda - \pi_\Sigma\lambda, \pi_\Sigma(\gamma_n \boldsymbol{u}) - \gamma_n \boldsymbol{u}\rangle_\Sigma. \tag{4.5}$$

Therefore, by setting $\boldsymbol{v}_h = \pi_V \boldsymbol{u}$, $q_h = \pi_Q p$, and $\mu_h = \pi_\Sigma \lambda$ in (4.4) and using (4.5), we can show that

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_V \lesssim \mathrm{E}(\boldsymbol{u}, \pi_V \boldsymbol{u})^{1/2} + \|p - \pi_Q p\|_Q + \|\lambda - \pi_\Sigma \lambda\|_{\Sigma'} + \|\gamma_n \boldsymbol{u} - \pi_\Sigma(\gamma_n \boldsymbol{u})\|_\Sigma . \tag{4.6}$$

We then establish (4.2a) with the approximation properties of the interpolation operators presented in Appendices B.1 and B.2 for π_Σ and π_V respectively, and [14] for π_Q. The estimate (4.2b) then follows from Lemma 3.6. ◻

REMARK 4.3. *The velocity and pressure error estimates coincide with those obtained in [4, Theorem 2.14] and in [29, Theorem 3.1] for the r-Stokes system without contact or friction boundary conditions. This indicates that these boundary conditions and the Lagrange multiplier do not reduce the order of convergence. This may be due to the use of piecewise constant elements for Σ_h. In [9], a proof with non-optimal convergence rates is presented for the case when continuous piecewise quadratic polynomials are used for the Lagrange multiplier.*

4.2. Discrete algebraic formulation. We now present an algebraic counterpart of (3.2) using the finite element spaces specified in (4.1) in terms of matrices and vectors. Let $V_h = \operatorname{span}\{\boldsymbol{v}_i\}_{i=1}^{N_v}$, $Q_h = \operatorname{span}\{q_j\}_{j=1}^{N_q}$ and $\Sigma_h = \operatorname{span}\{\mu_k\}_{k=1}^{N_\mu}$, where $N_v = \dim V_h$, $N_q = \dim Q_h$ and $N_\mu = \dim \Sigma_h$. For the functions $(\boldsymbol{u}_h, p_h, \lambda_h) \in V_h \times Q_h \times \Sigma_h$, we write $\mathbf{u}$, $\mathbf{p}$ and $\boldsymbol{\lambda}$ for the vectors containing the respective degrees of freedom (DoFs) in $\mathbb{R}^{N_v}$, $\mathbb{R}^{N_q}$ and $\mathbb{R}^{N_\mu}$. In order to write an algebraic counterpart of (3.2c), we need to introduce the discrete normal trace operator

$$\boldsymbol{\gamma_n} : \mathbb{R}^{N_v} \to \mathbb{R}^{N_\mu}$$

that returns the average normal components of a vector $\boldsymbol{v}_h \in V_h$ along the edges on Γ_b. That is, for each $i \in \{1, 2, ..., N_\mu\}$

$$(\boldsymbol{\gamma_n}\mathbf{v})_i = \frac{1}{|e_i|}\int_{e_i} \boldsymbol{v}_h \cdot \boldsymbol{n}\,\mathrm{d}s,$$

where $e_i \in \mathcal{E}(\Gamma_b)$ is the unique edge along Γ_b associated to the degree of freedom in Σ_h with index i. Then, the algebraic counterpart of (3.2) can be written in terms of matrices and vectors as

$$\mathbf{A}_\varepsilon(\mathbf{u}) + \mathbf{G}_\varepsilon(\mathbf{u}) - \mathbf{B}\mathbf{p} - \mathbf{D}\boldsymbol{\lambda} = \mathbf{f}, \tag{4.7a}$$

$$\mathbf{B}^\top \mathbf{u} = 0, \tag{4.7b}$$

$$\boldsymbol{\lambda} + \mathbf{C}(\boldsymbol{\lambda}, \mathbf{u}) = 0. \tag{4.7c}$$

Here, we have introduced the matrices $\mathbf{B} \in \mathbb{R}^{N_v \times N_q}$ and $\mathbf{D} \in \mathbb{R}^{N_v \times N_\mu}$, the vector $\mathbf{f} \in \mathbb{R}^{N_v}$ and the nonlinear operators $\mathbf{A}_\varepsilon : \mathbb{R}^{N_v} \to \mathbb{R}^{N_v}$, $\mathbf{G}_\varepsilon : \mathbb{R}^{N_v} \to \mathbb{R}^{N_v}$ and $\mathbf{C} : \mathbb{R}^{N_\mu} \times \mathbb{R}^{N_v} \to \mathbb{R}^{N_\mu}$. The matrices are given by the elements $\mathbf{B}_{ij} = \langle Bq_j, \boldsymbol{v}_i \rangle_V$ and $\mathbf{D}_{ij} = \langle \mu_j, \gamma_n \boldsymbol{v}_i \rangle_\Sigma$ and the vector by $\mathbf{f}_i = \langle F, \boldsymbol{v}_i \rangle_V$. The nonlinear operators are defined as

$$[\mathbf{A}_\varepsilon(\mathbf{u})]_i = \int_\Omega \alpha \left(\varepsilon + |\mathbf{D}\boldsymbol{u}|\right)^{r-2} (\mathbf{D}\boldsymbol{u} : \mathbf{D}\boldsymbol{v}_i) \,\mathrm{d}x, \tag{4.8}$$

$$[\mathbf{G}_\varepsilon(\mathbf{u})]_i = \int_\Omega \tau \left(\varepsilon + |\mathbf{T}\boldsymbol{u}|\right)^{r-2} (\mathbf{T}\boldsymbol{u} \cdot \mathbf{T}\boldsymbol{v}_i) \,\mathrm{d}x, \tag{4.9}$$

$$\mathbf{C}(\boldsymbol{\lambda}, \mathbf{u}) = \max\{0, -\boldsymbol{\lambda} + c(\boldsymbol{\gamma_n}\mathbf{u})\}, \tag{4.10}$$

for an arbitrary $c > 0$ and a regularisation term $\varepsilon > 0$. This regularisation term is commonly used to avoid numerical complications caused when $|\mathbf{D}\boldsymbol{u}|$ or $|\mathbf{T}\boldsymbol{u}|$ is equal to or very close to zero [32]. In (4.10), the max operation is understood to be carried out componentwise on each of the elements in the vector $-\boldsymbol{\lambda} + c(\boldsymbol{\gamma_n}\mathbf{u}) \in \mathbb{R}^{N_\mu}$. The use of the operator $\mathbf{C}$ in (4.10) is a common way of expressing contact conditions. A particular advantage is that the nonlinear system (4.7) can be solved with a semi-smooth Newton method that enjoys superlinear convergence in a neighbourhood of the solution [28]. Equation (4.7c) is equivalent to (3.2c) whenever Σ_h is defined as in (4.1c). By solving (4.7c) we enforce

$$\boldsymbol{\gamma_n}\mathbf{u} \le 0, \quad \boldsymbol{\lambda} \le 0 \quad \text{and} \quad (\boldsymbol{\gamma_n}\mathbf{u}) \cdot \boldsymbol{\lambda} = 0, \tag{4.11}$$

exactly, which is the algebraic equivalent of the discrete contact conditions (3.3).

4.3. Numerical results. We present numerical results computed for a Stokes variational inequality with a manufactured solution on the domain $\Omega = (0,1)^2$. Computations of subglacial cavitation with the algorithm presented in this document can be found in [10]. Subglacial cavitation is a time-dependent problem in which a free surface is evolved with an advection equation. We do this in [10] by coupling the algorithm presented in this study with a solver for the advection equation.

The manufactured solution considered here is taken from [4] and is given by

$$\hat{\boldsymbol{u}}(\boldsymbol{x}) = |\boldsymbol{x}|^{\alpha-1}(x_2, -x_1)^\top, \quad \hat{p}(\boldsymbol{x}) = |\boldsymbol{x}|^\gamma, \tag{4.12}$$

where the parameters α and γ are chosen such that $\boldsymbol{u} \in \boldsymbol{W}^{2,r}(\Omega)$, $p \in W^{1,r'}(\Omega)$, $\mathbf{F}(\mathbf{D}\boldsymbol{u}) \in \boldsymbol{W}^{1,2}(\Omega)$ and $\mathbf{F}(\mathbf{T}\boldsymbol{u}) \in \boldsymbol{W}^{1,2}(\Gamma_b)$ hold. This is ensured whenever $\alpha > 1$ and $\gamma > -1 + \frac{2}{r}$, so we set $\alpha = 1.01$ and $\gamma = -1 + \frac{2}{r} + 0.01$ in order to be critically close to the regularity assumed in Theorem 4.2.

Contact boundary conditions are enforced on the lower boundary $\{y = 0\}$. Given the velocity and pressure fields defined in (4.12), we have that

$$(\hat{\boldsymbol{u}} \cdot \boldsymbol{n})(x_1) = -x_1^\alpha, \quad \hat{\lambda}(x_1) = -x_1^\gamma,$$

on $\{y = 0\}$. In order to define the contact boundary conditions in such a way that both the kinematic and dynamic conditions are active, we define the "obstacles"

$$\chi(x_1) = \begin{cases} (\hat{\boldsymbol{u}} \cdot \boldsymbol{n})(x_1) & \text{if } x_1 \le 0.5 \\ -2^{-\alpha} & \text{if } x_1 > 0.5 \end{cases}, \quad \rho(x_1) = \begin{cases} 0 & \text{if } x_1 \le 0.5 \\ \hat{\lambda}(x_1) & \text{if } x_1 > 0.5 \end{cases}.$$

Then, for this numerical test we solve the r-Stokes system (2.1) together with the boundary conditions

$$\boldsymbol{u}\cdot\boldsymbol{n} \leq \chi, \quad \lambda \leq \rho \quad \text{and} \quad (\boldsymbol{u}\cdot\boldsymbol{n} - \chi)(\lambda - \rho) = 0 \quad \text{on } \{y = 0\}, \tag{4.13a}$$

$$\boldsymbol{u}\cdot\boldsymbol{n} = \hat{\boldsymbol{u}}\cdot\boldsymbol{n} \quad \text{on } \{x = 0\}, \tag{4.13b}$$

$$\sigma_{nn} = \hat{\sigma}_{nn} \quad \text{on } \partial\Omega \setminus (\{y = 0\} \cup \{x = 0\}), \tag{4.13c}$$

$$\sigma_{nt} = \hat{\sigma}_{nt} \quad \text{on } \partial\Omega, \tag{4.13d}$$

where $\hat{\sigma} = \sigma(\hat{\boldsymbol{u}}, \hat{p})$. Boundary conditions are set for the normal velocity along $\{x = 0\}$ in (4.13b) to mimic the boundary conditions enforced at Γ_d and make $\dim R_V = 1$. In this case, R_V is a one-dimensional vector space containing vertical motions. Therefore,

$$R_V \cap K = \{(0, \theta) : \theta \leq \min \chi\},$$

and, since $F = A_\varepsilon \hat{\boldsymbol{u}} + G_\varepsilon \hat{\boldsymbol{u}} - B\hat{p} - \gamma_n' \hat{\lambda}$, we have that

$$\langle F, \boldsymbol{r}\rangle_V = \theta \int_0^1 x^\gamma \,\mathrm{d}x < 0$$

for all $\boldsymbol{r} = (0, \theta)$ with $\theta < 0$ (note that $\min \chi < 0$). This proves that the compatibility condition (2.20) holds and the system is well-posed. Although the functional setting of this numerical test differs slightly from the setting studied in this paper, the numerical test contains the fundamental elements of the setting analysed.

We compute solutions to the r-Stokes system on $\Omega = (0, 1)^2$ with boundary conditions (4.13) on a sequence of uniformly refined meshes using the finite element spaces in (4.1). The regularisation parameter in (4.8) is set to $\varepsilon = 10^{-4}$. In Glen's law (2.2) we fix $\mathcal{A} = 0.5$ and for the friction boundary condition we set $\tau = 1$. We consider the values $n = 1$, 2, 3 and 4, which correspond with $r = 2$, 1.5, 1.33 and 1.25. The resulting orders of convergence for the velocity are shown in Tables 1 and 2, and for the pressure and Lagrange multiplier in Table 3. For the Lagrange multiplier error, we use the discrete norm

$$\|\mu_h\|_{\Sigma', h} = h^{1/r'} \|\mu_h\|_{L^{r'}(\Gamma_b)},$$

which should yield the same order of convergence as the one that would be obtained with the Σ' norm by a standard inverse inequality.

Table 1 indicates that the orders of convergence for the velocity in the seminorm $\|\mathbf{D}(\cdot)\|_{L^r(\Omega)}$ and in the V-norm coincide. This demonstrates that the presence of rigid modes in the velocity space does not affect the accuracy of the velocity computation in the V-norm. The computed orders of convergence for the velocity in the V-norm coincide with those estimated in (4.2a). In Table 2 we see that the orders of convergence for the velocity in the $L^r(\Omega)$-norm appear to increase by one when compared to the orders computed with the V-norm. On the other hand, the orders of convergence obtained for the pressure appear to be independent of r. A closely related problem (without contact boundary conditions) is solved in the work of Belenki et al. [4]. In the work of Belenki et al., the problem is formulated as an r-Stokes problem with Dirichlet boundary conditions and the MINI element is used for the velocity and pressure. Interestingly, their numerical results deliver the predicted orders of convergence for the pressure error. Hence, the apparent suboptimality of (4.2b) for the pressure could be due to the finite elements used here or to the presence of contact boundary conditions and a Lagrange multiplier. Regarding the Lagrange multiplier, the estimated orders of convergence are exceeded, but a dependence on r is observed.

Table 1

Calculated orders of convergence for the velocity computed with a manufactured solution together with estimated orders according to Theorem 4.2.

	$\lVert \mathbf{D}(\boldsymbol{u}) - \mathbf{D}(\boldsymbol{u}_h) \rVert_{L^r(\Omega}$				$\lVert \boldsymbol{u} - \boldsymbol{u}_h \rVert_V$			
$h \backslash r$	2.00	1.50	1.33	1.25	2.00	1.50	1.33	1.25
3.54×10^{-1}	-	-	-	-	-	-	-	-
1.77×10^{-1}	0.96	1.05	1.08	1.11	0.97	1.10	1.14	1.18
8.84×10^{-2}	0.97	1.03	1.05	1.07	0.98	1.06	1.09	1.12
4.42×10^{-2}	0.97	1.02	1.04	1.06	0.98	1.04	1.06	1.08
2.21×10^{-2}	0.97	1.02	1.03	1.04	0.98	1.03	1.04	1.06
1.10×10^{-2}	0.97	1.01	1.02	1.03	0.98	1.02	1.03	1.04
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 2

Calculated orders of convergence for the velocity in the $L^r(\Omega)$-norm computed with a manufactured solution.

	$\lVert \boldsymbol{u} - \boldsymbol{u}_h \rVert_{L^r(\Omega}$			
$h \backslash r$	2.00	1.50	1.33	1.25
3.54×10^{-1}	-	-	-	-
1.77×10^{-1}	1.97	2.07	1.99	1.85
8.84×10^{-2}	1.95	1.95	1.87	1.72
4.42×10^{-2}	1.96	1.96	1.86	1.71
2.21×10^{-2}	1.96	1.97	1.87	1.72
1.10×10^{-2}	1.96	1.98	1.87	1.74

5. Conclusions. In this study, we present a Stokes variational inequality that arises when modelling a symmetrical marine ice sheet. We prove the well-posedness of this system whenever the subspace of rigid modes in the velocity space is of dimension at most one under the condition that a compatibility condition holds in Theorem 2.3. We consider a family of finite element discretisations for this problem in Section 3 and prove an analogous well-posedness result for the discrete system in Theorem 3.1. Using the techniques from [4, 29], we then establish error estimates for a particular finite element discretisation in Theorem 4.2. These error estimates, which are verified with a numerical test using a manufactured solution, indicate that the presence of rigid modes and the nonlinearity of the friction boundary condition do not affect the order of convergence of the scheme.

The results from this study give a theoretical justification for using a finite element discretisation belonging to the family considered in Section 3 in glaciological applications. Moreover, this analysis can be extended to different contact problems in glaciology such as the subglacial cavitation problem, where the subspace of rigid modes present in the velocity space is of dimension one whenever Dirichlet boundary conditions are enforced for the tangential component of the velocity on the top boundary, as done in [20]. This extension would require taking into consideration the presence of non-homogeneous boundary conditions for the velocity and the use of a periodic domain. However, if instead of Dirichlet boundary conditions, we enforce Neumann boundary conditions on the top boundary, as described in [10], the space

TABLE 3
Calculated orders of convergence for the pressure and the Lagrange multiplier computed with a manufactured solution together with estimated orders according to Theorem 4.2.

	$\|p-p_h\|_Q$				$\|\lambda-\lambda_h\|_{\Sigma',h}$			
$h\backslash r$	2.00	1.50	1.33	1.25	2.00	1.50	1.33	1.25
3.54×10^{-1}	-	-	-	-	-	-	-	-
1.77×10^{-1}	0.88	0.93	0.96	0.98	1.00	1.00	1.00	0.98
8.84×10^{-2}	0.90	0.94	0.97	0.98	1.00	1.00	0.98	0.93
4.42×10^{-2}	0.91	0.95	0.97	0.95	1.01	1.00	0.96	0.87
2.21×10^{-2}	0.92	0.95	0.97	0.90	1.01	1.00	0.93	0.80
1.10×10^{-2}	0.93	0.96	0.96	0.84	1.01	1.00	0.88	0.73
$2/r'$	1.00	0.67	0.5	0.4	1.00	0.67	0.5	0.4

of rigid modes is then of dimension two. This situation would require a more complicated extension that should be considered in future work. Subglacial cavitation is considered in [10] and the Stokes variational inequality that arises is solved using the finite element discretisation from Section 4.

A major assumption of this paper is that the domain is two-dimensional. An extension of the analysis presented here to three dimensions would require a careful consideration of the rigid modes present in the velocity space, since the space of rigid modes in three dimensions is larger than in two dimensions. However, in most problems of interest, three dimensional marine ice sheets are considered to be enclosed within two lateral walls, see for example [16]. In this case, if the lateral walls and the bedrock are flat, the space of rigid modes in V is once again reduced to vertical movements and is therefore one-dimensional. As a consequence, much of the analysis from this paper would still be valid in three dimensions. However, the extension operator presented in Appendix B.3, used to prove Lemma 4.1, relies heavily on the fact that the domain is two-dimensional. Therefore, the choice of finite elements used to solve the variational inequality would have to be chosen and studied carefully.

Appendix A. Equivalence of formulations.

In this appendix we demonstrate the equivalence between the strong formulation (2.1) of the contact problem with boundary conditions (2.4)-(2.6), the variational inequality (2.11), the minimisation of $\mathcal{J}$, defined in (2.17), and the mixed formulation (2.12). This analysis is similar to the one presented in [8], with the difference that in this case we consider contact boundary conditions.

LEMMA A.1. *If $(\boldsymbol{u},p)\in\boldsymbol{\mathcal{C}}^2(\Omega)\times\mathcal{C}^1(\Omega)$, then the strong formulation (2.1) with boundary conditions (2.4)-(2.6) holds if and only if the variational inequality (2.11) is satisfied.*

Proof. Let $(\boldsymbol{u},p)\in\boldsymbol{\mathcal{C}}^2(\Omega)\times\mathcal{C}^1(\Omega)$ satisfy (2.1) and (2.4)-(2.6). It is clear that if (2.1b) holds, then $\langle Bq,\boldsymbol{u}\rangle_V=0$ for all $q\in Q$. Let $\boldsymbol{v}\in K$ and multiply (2.1a) by $\boldsymbol{v}-\boldsymbol{u}$ and integrate over Ω. The equality

$$-\int_\Omega \left[\nabla\cdot\left(\alpha|\mathbf{D}\boldsymbol{u}|^{r-2}\mathbf{D}\boldsymbol{u}\right)-\nabla p\right]\cdot(\boldsymbol{v}-\boldsymbol{u})\,\mathrm{d}x = \langle A\boldsymbol{u}-Bp,\boldsymbol{v}-\boldsymbol{u}\rangle_V-\int_{\partial\Omega}\sigma(\boldsymbol{v}-\boldsymbol{u})\cdot\boldsymbol{n}\,\mathrm{d}s \tag{A.1}$$

follows from the divergence theorem. We also have that

$$\int_{\partial\Omega} \sigma(\boldsymbol{v}-\boldsymbol{u})\cdot\boldsymbol{n}\,\mathrm{d}s = \int_{\partial\Omega} \left(\sigma_{nn}(\boldsymbol{v}-\boldsymbol{u})\cdot\boldsymbol{n} + \boldsymbol{\sigma}_{nt}\cdot(\boldsymbol{v}-\boldsymbol{u})\right)\,\mathrm{d}s.$$

As a result of the contact conditions (2.5a), we have

$$\int_{\Gamma_b} \sigma_{nn}(\boldsymbol{v}-\boldsymbol{u})\cdot\boldsymbol{n}\,\mathrm{d}s \geq -\int_{\Gamma_b} p_w\,(\boldsymbol{v}-\boldsymbol{u})\cdot\boldsymbol{n}\,\mathrm{d}s,$$

from which the variational inequality (2.11) follows.

The converse statement is deduced by means of the integration by parts formula (A.1) and the use of the fundamental lemma of calculus of variations with adequate test functions. The examples in [24, 27] contain similar derivations. ∎

For the next step, we need to use the inf-sup condition (2.15) between the velocity and the pressure spaces.

Lemma A.2. *Given a solution $(\boldsymbol{u},p)\in K\times Q$ of the variational inequality (2.11), the velocity field is then divergence free, i.e. $\boldsymbol{u}\in\mathring{K}$, and is a minimiser of the functional $\mathcal{J}:\mathring{K}\to\mathbb{R}$ defined in (2.17). Conversely, if $\boldsymbol{u}\in\mathring{K}$ minimises $\mathcal{J}:\mathring{K}\to\mathbb{R}$, then there is a unique $p\in Q$ such that $(\boldsymbol{u},p)\in K\times Q$ solves (2.11).*

Proof. For the first part of the Lemma, for a test function $\boldsymbol{v}\in\mathring{K}$, the variational inequality (2.11) can be written as

$$\langle A\boldsymbol{u}+G\boldsymbol{u}-F,\boldsymbol{v}-\boldsymbol{u}\rangle_V \geq 0 \quad \forall\boldsymbol{v}\in\mathring{K}. \tag{A.2}$$

By the convexity of $\mathcal{J}$ and the fact that $\langle D\mathcal{J}(\boldsymbol{u}),\boldsymbol{v}\rangle_V = \langle A\boldsymbol{u}+G\boldsymbol{u}-F,\boldsymbol{v}\rangle_V$, it follows from (A.2) that $\mathcal{J}(\boldsymbol{u})\leq\mathcal{J}(\boldsymbol{v})$ for all $\boldsymbol{v}\in\mathring{K}$.

Conversely, if we assume $\boldsymbol{u}\in\mathring{K}$ to minimise $\mathcal{J}$ over $\mathring{K}$, then $\boldsymbol{u}$ solves (A.2). Now, using [3, Lemma 3.3], we can decompose $\boldsymbol{v}\in K$ into the sum $\boldsymbol{v}=\boldsymbol{v}_0+\boldsymbol{w}$ of a divergence-free velocity field $\boldsymbol{v}_0\in\mathring{K}$ and the field $\boldsymbol{w}\in V_b$. Then, the variational inequality (2.11) will hold if there is a $p\in Q$ such that

$$\langle A\boldsymbol{u}-F,\boldsymbol{w}\rangle_V = \langle Bp,\boldsymbol{w}\rangle_V \quad \forall\boldsymbol{w}\in V_b. \tag{A.3}$$

By (2.15), there is a $p\in Q$ for which (A.3) holds and it is unique. ∎

Finally, we show that the variational inequality (2.11) is equivalent to the mixed problem (2.12). This proof relies on the fact that the range of the operator $\gamma_n : V\to\Sigma$ is closed.

Lemma A.3. *If $(\boldsymbol{u},p)\in K\times Q$ solves the variational inequality (2.11), then there is a unique $\lambda\in\Lambda$ such that $(\boldsymbol{u},p,\lambda)$ is a solution of the mixed problem (2.12). Conversely, if $(\boldsymbol{u},p,\lambda)\in V\times Q\times\Lambda$ solves (2.12), then $(\boldsymbol{u},p)$ is a solution of (2.11).*

Proof. Equation (2.12a) can be rewritten as

$$\gamma_n'\lambda = A\boldsymbol{u}+G\boldsymbol{u}-Bp-F \quad \text{in } V'. \tag{A.4}$$

Since $\gamma_n : V\to\Sigma$ has a closed range, we have that $\operatorname{Ran}\gamma_n' = (\operatorname{Ker}\gamma_n)^\circ$, where

$$(\operatorname{Ker}\gamma_n)^\circ = \left\{\mu\in\Sigma' : \langle\mu,\phi\rangle_\Sigma = 0 \quad \forall\phi\in\operatorname{Ker}\gamma_n\right\}.$$

Therefore, if $(\boldsymbol{u},p)\in K\times Q$ is a solution to (2.11), then there is unique $\lambda\in\Sigma'$ if $A\boldsymbol{u}+G\boldsymbol{u}-Bp-F\in(\operatorname{Ker}\gamma_n)^\circ$. For a $\boldsymbol{w}\in\operatorname{Ker}\gamma_n$, we clearly have that $\boldsymbol{u}+\boldsymbol{w}\in K$. Using the variational inequality (2.11), we can write

$$\langle A\boldsymbol{u}+G\boldsymbol{u}-Bp-F,\boldsymbol{w}\rangle_V = 0,$$

which means that $A\boldsymbol{u}+G\boldsymbol{u}-Bp-F \in (\operatorname{Ker}\gamma_n)^\circ$. Next, we must show that $\lambda \in \Lambda$ and that (2.12c) holds. By setting $\boldsymbol{v}=0$ and $\boldsymbol{v}=2\boldsymbol{u}$ in (2.11) we see that $\langle \lambda, \gamma_n \boldsymbol{u}\rangle_\Sigma = 0$. Since $\gamma_n \boldsymbol{u} \le 0$ in Σ, it follows that (2.12c) must hold. Finally, $\lambda \in \Lambda$ follows from (2.11), (A.4) and the fact that $\boldsymbol{v}+\boldsymbol{u} \in K$ for any $\boldsymbol{v} \in K$.

For the second part of the lemma, if $(\boldsymbol{u}, p, \lambda) \in V \times Q \times \Lambda$ solves (2.12), then $\nabla \cdot \boldsymbol{u} = 0$ a.e. in Ω, $\langle \lambda, \gamma_n \boldsymbol{v}\rangle_\Sigma \ge 0$ for all $\boldsymbol{v} \in K$ and $\langle \lambda, \gamma_n \boldsymbol{u}\rangle_\Sigma = 0$. This implies that $\langle \mu, \gamma_n \boldsymbol{u}\rangle_\Sigma \ge 0$ for all $\mu \in \Lambda$, hence $\boldsymbol{u} \in \mathring{K}$. The variational inequality then follows directly from (2.12) by testing with $(\boldsymbol{v}-\boldsymbol{u}, q)$, where $(\boldsymbol{v}, q) \in K \times Q$. ∎

Appendix B. Technical results on finite element spaces.

B.1. Approximation properties in negative order Sobolev spaces. Several technicalities arise from the need to handle the dual space of the fractional Sobolev space $W^{1-1/r,r}(\Gamma_b)$ and its finite element approximation Σ_h. Let $s \in (0,1)$ and $m \in [1,\infty)$; on Γ_b, the norm of the fractional Sobolev space $W^{s,m}(\Gamma_b)$ with $s \in (0,1)$ and $m \in [1,\infty]$ can be defined by

$$\|\phi\|^m_{W^{s,m}(\Gamma_b)} = \|\phi\|^m_{L^m(\Gamma_b)} + [\phi]^m_{W^{s,m}(\Gamma_b)}, \tag{B.1}$$

where

$$[\phi]^m_{W^{s,m}(\Gamma_b)} = \int_{\Gamma_b}\int_{\Gamma_b} \frac{|\phi(x)-\phi(y)|^m}{|x-y|^{1+sm}}\,\mathrm{d}x\mathrm{d}y,$$

see [11]. In order to prove certain approximation properties on Σ_h we need to introduce some theoretical results. We start by defining the following pair of spaces

$$L^m_0(e) = \left\{\phi \in L^m(e) : \int_e \phi\,\mathrm{d}x = 0\right\}, \quad W^{s,m}_0(e) = W^{s,m}(e) \cap L^r_0(e)$$

for an edge $e \in \mathcal{E}(\Gamma_b)$. We can use the fractional normed Poincaré inequality proved in [14, Lemma 7.1] to show that

$$\|\phi\|_{L^m(e)} \le |e|^s [\phi]_{W^{s,m}(e)} \quad \forall \phi \in W^{s,m}_0(e). \tag{B.2}$$

Inequality (B.2) can be extended to negative norms by writing the $L^m(e)$ norm for any $\phi \in L^m(\Gamma_b)$ as

$$\|\phi\|_{L^m(e)} = \sup_{\psi \in L^{m'}(e)} \frac{\int_e \phi\psi\,\mathrm{d}s}{\|\psi\|_{L^{m'}(e)}}$$

and deducing that, if $\phi \in L^m_0(e)$, we have

$$\begin{aligned}\|\phi\|_{(W^{s,m}(e))'} &= \sup_{\psi \in W^{s,m}_0(e)} \frac{\int_e \phi\psi\,\mathrm{d}s}{\|\psi\|_{W^{s,m}(e)}} \\ &\le |e|^s \sup_{\psi \in L^{m'}(e)} \frac{\int_e \phi\psi\,\mathrm{d}s}{\|\psi\|_{L^{m'}(e)}} = |e|^s \|\phi\|_{L^{m'}(e)}.\end{aligned} \tag{B.3}$$

For the finite element space Σ_h defined in (4.1c), let $\pi_\Sigma : L^m(\Gamma_b) \to \Sigma_h$ be the standard interpolation operator onto piecewise constant polynomials which takes the average of functions over each $e \in \mathcal{E}(\Gamma_b)$. Then, from inequality (B.2) we can prove error estimates in fractional norms. Moreover, (B.3) leads to

$$\|\phi - \pi_\Sigma\phi\|_{(W^{s,m}(\Gamma_b))'} \lesssim h^{2s}\|\phi\|_{W^{s,m'}(\Gamma_b)}, \tag{B.4}$$

where $h = \max\{|e| : e \in \mathcal{E}(\Gamma_b)\}$.

B.2. An interpolation operator for the velocity. Here, we compile a variety of results from different sources and prove a result regarding an interpolation operator for the velocity that preserves the discrete divergence and maps elements of K into K_h. We denote by π_V the interpolation operator introduced in [22, Section 3.1] that is defined as follows for each component of a vector-valued function: for a non-degenerate simplex $K \in \mathcal{T}_h$ with edges $\{e_i\}_{i=1}^3$ and vertices $\{\boldsymbol{a}_i\}_{i=1}^3$, we define the nodal basis functions ϕ_x with $x \in \{e_i\}_{i=1}^3 \cup \{\boldsymbol{a}_i\}_{i=1}^3$ by

$$\phi_{\boldsymbol{a}_i}(\boldsymbol{a}_j) = \delta_{ij}, \quad \int_{e_j} \phi_{\boldsymbol{a}_i}\,\mathrm{d}s = 0, \quad \int_{e_j} \phi_{e_i}\,\mathrm{d}s = \delta_{ij}, \quad \phi_{e_i}(\boldsymbol{a}_j) = 0,$$

for all $i, j \in \{1, 2, 3\}$. For each vertex $\boldsymbol{a}_i$ we choose a an edge $e_{\boldsymbol{a}_i} \in \{e_i\}_{i=1}^3$ such that $\boldsymbol{a}_i \in \overline{e_{\boldsymbol{a}_i}}$. We then define the dual basis functions $\{\psi_{\boldsymbol{a}_i}\}_{i=1}^3$ by

$$\int_{e_{\boldsymbol{a}_i}} \psi_{\boldsymbol{a}_i}\phi_x\,\mathrm{d}s = \delta_{\boldsymbol{a}_i x}, \quad \psi_{\boldsymbol{a}_i} \in \mathbb{P}_2(e_{\boldsymbol{a}_i}),$$

where x denotes the edge $e_{\boldsymbol{a}_i}$ or its two end-points. We then define (the scalar version of) π_V by

$$(\pi_V u)(\boldsymbol{x}) = \sum_{i=1}^{3} \left(\left[\int_{e_{\boldsymbol{a}_i}} u\psi_{\boldsymbol{a}_i}\,\mathrm{d}s \right] \phi_{\boldsymbol{a}_i}(\boldsymbol{x}) + \left[\int_{e_i} u\,\mathrm{d}s \right] \phi_{e_i}(\boldsymbol{x}) \right).$$

When considering the definition of π_V in terms of a triangulation $\mathcal{T}_h$ of Ω, for vertices $\boldsymbol{a} \in \partial\Omega$, we set the associated edge $e_{\boldsymbol{a}}$ to also be contained in $\partial\Omega$. Then, we have that $\pi_V(V) \subset V_h$ and we can prove (B.9).

If the spaces V_h, Q_h and Σ_h are defined as in (4.1), a straightforward consequence of the definition of π_V is that

$$\text{(B.5)} \qquad \langle Bq_h, \boldsymbol{v}\rangle_V = \langle Bq_h, \pi_V \boldsymbol{v}\rangle_V \qquad \forall(\boldsymbol{v}, q_h) \in V \times Q_h,$$

$$\text{(B.6)} \qquad \langle \mu_h, \gamma_n \boldsymbol{v}\rangle_\Sigma = \langle \mu_h, \gamma_n \pi_V \boldsymbol{v}\rangle_\Sigma \qquad \forall(\boldsymbol{v}, \mu_h) \in V \times \Sigma_h.$$

Additionally, the interpolation operator π_V has two key approximation properties. First, the optimal approximation property

$$\text{(B.7)} \qquad \|\boldsymbol{v} - \pi_V \boldsymbol{v}\|_{\boldsymbol{W}^{s,m}(\Omega)} \lesssim h^k \|\boldsymbol{v}\|_{\boldsymbol{W}^{s+k,m}(\Omega)}$$

holds for all $m \geq 0$ and $s, k \in \mathbb{N}$ such that $0 \leq s \leq 3$ and $0 \leq k \leq 3 - s$. Finally, given the operator $\mathbf{F}$ defined in (3.6), the additional approximation property holds:

$$\text{(B.8)} \qquad \|\mathbf{F}(\mathbf{D}\boldsymbol{v}) - \mathbf{F}(\mathbf{D}\pi_V \boldsymbol{v})\|_{L^2(\Omega)} \lesssim h \|\nabla \mathbf{F}(\mathbf{D}\boldsymbol{v})\|_{L^2(\Omega)}.$$

Property (B.7) is shown to hold in [22]. On the other hand, (B.8) follows from [4, Theorem 3.4] by applying Poincaré's inequality once points (a) and (b) from Assumption 2.9 in that reference are shown to hold. These two points result from (B.5) and (B.7).

Finally, we may also prove that

$$\text{(B.9)} \qquad \|\mathbf{F}(\mathbf{T}\boldsymbol{v}) - \mathbf{F}(\mathbf{T}\pi_V \boldsymbol{v})\|_{L^2(\Gamma_b)} \lesssim h \sum_{\substack{e \in \mathcal{E}(\partial\Omega) \\ \overline{e} \cap \overline{\Gamma}_b \neq \emptyset}} \|\nabla \mathbf{F}(\mathbf{T}\boldsymbol{v})\|_{L^2(e)}$$

by imitating the proof for [4, Theorem 3.4] and applying Poincaré's inequality. Most of the steps in this proof draw from algebraic relations for the function $\mathbf{F}$ and the N-functions considered therein that continue to be valid in our context. Additionally, we need the following Orlicz-continuity result analogous to that of [4, Theorem 3.2]: for an N-function ψ with $\Delta_2(\psi) < \infty$ and an edge $e \in \mathcal{E}(\Gamma_b)$,

$$\int_e \psi(|\mathbf{T}\pi_V \boldsymbol{v}|)\,\mathrm{d}s \lesssim \sum_{\substack{e' \in \mathcal{E}(\partial\Omega) \\ \overline{e}\cap\overline{e'}\neq\emptyset}} \int_{e'} \psi(|\mathbf{T}\boldsymbol{v}|)\,\mathrm{d}s.$$

We may show the above inequality to hold by following the proof of [12, Theorem 4.5] and using the local L^1-estimate for $e \in \mathcal{E}(\Gamma_b)$:

$$\int_e |\mathbf{T}\pi_V \boldsymbol{v}|\,\mathrm{d}s \lesssim \sum_{\substack{e' \in \mathcal{E}(\partial\Omega) \\ \overline{e}\cap\overline{e'}\neq\emptyset}} \int_{e'} |\mathbf{T}\boldsymbol{v}|\,\mathrm{d}s.$$

To prove this inequality, we turn to the definition of π_V and use the bounds

$$\|\psi_{\boldsymbol{a}}\|_{L^\infty(e_{\boldsymbol{a}})} \lesssim |e|^{-1}, \quad \|\phi_{\boldsymbol{a}}\|_{L^1(e_{\boldsymbol{a}})} \lesssim |e| \quad \text{and} \quad \|\phi_e\|_{L^1(e)} \lesssim 1.$$

B.3. An extension operator. In this section we prove an auxiliary result required for the proof of Lemma 4.1. We build an extension operator $\Phi : \Sigma \to V_h$ which is uniformly bounded and satisfies

$$\langle \mu_h, \gamma_n(\Phi\phi)\rangle_\Sigma = \langle \mu_h, \phi\rangle_\Sigma \quad \forall \mu_h \in \Sigma_h. \tag{B.10}$$

Step 1. We will first find a uniformly bounded linear operator $\Pi : \Sigma \to \gamma_n(V_h)$ with the property that

$$\int_e (\phi - \Pi\phi)\,\mathrm{d}s = 0 \quad \text{for any } e \in \mathcal{E}(\Gamma_b) \text{ and } \phi \in \Sigma. \tag{B.11}$$

Let

$$Z_h = \{\phi_h \in \mathcal{C}(\Gamma_b) : \phi_h|_e \in \mathcal{P}_2(e) \quad \forall e \in \mathcal{E}(\Gamma_b)\}$$

and note that $Z_h \subset \gamma_n(V_h)$. For $\phi \in \Sigma$ and $e \in \mathcal{E}(\Gamma_b)$, we define $\Pi_2 : \Sigma \to Z_h$ by setting

$$\begin{aligned} (\Pi_2\phi)(a) &= 0 \quad \text{for the endpoints } a \text{ in } e, \\ \int_e \Pi_2\phi\,\mathrm{d}s &= \int_e \phi\,\mathrm{d}s. \end{aligned}$$

We clearly have that $\Pi_2\phi = 0$ if and only if $\sum_{e\in\mathcal{E}(\Gamma_b} \int_e |\phi|\,\mathrm{d}s = 0$, so the latter defines a norm on $\Pi_2(\Sigma)$. By exploiting this fact and the norm equivalence on finite dimensional spaces, one can see that

$$\|\Pi_2\phi\|_{W^{1-1/r,r}(e)} \lesssim |e|^{-1/r'} \|\phi\|_{L^r(e)} \quad \forall e \in \mathcal{E}(\Gamma_b)$$

for all $\phi \in \Sigma$. Now, let $\pi_Z : \Sigma \to Z_h$ be the quasi-interpolation operator defined in [14]. This operator is uniformly bounded in the $W^{1-1/r,r}(\Gamma_b)$ norm and satisfies

$$\|\phi - \pi_Z\phi\|_{L^r(e)} \lesssim |e|^{1-1/r} \|\phi\|_{W^{1-1/r,r}(e)}$$

for any edge $e \in \mathcal{E}(\Gamma_b)$ and function $\phi \in \Sigma$. As a result, the operator $\Pi = \pi_Z + \Pi_2(I - \pi_Z)$ is uniformly bounded and possesses the required property (B.11).

Step 2. For the final step, we define a uniformly bounded operator $\gamma_{n,h}^{-1} : \gamma_n(V_h) \to V_h$ for which $\gamma_{n,h}^{-1}\phi_h \cdot \boldsymbol{n} = \phi_h$ on Γ_b. This operator can be defined as the solution of the problem:

$$\begin{aligned}
\int_\Omega \nabla(\gamma_{n,h}^{-1}\phi_h) : \nabla \boldsymbol{v}_h \,\mathrm{d}x &= 0 && \forall \boldsymbol{v}_h \in V_h,\\
\gamma_{n,h}^{-1}\phi_h \cdot \boldsymbol{n} &= \phi_h && \text{on } \Gamma_b,\\
\gamma_{n,h}^{-1}\phi_h \cdot \boldsymbol{n} &= 0 && \text{on } \Gamma_d.
\end{aligned}$$

Then, the operator $\Phi = \gamma_{n,h}^{-1} \circ \Pi$ is uniformly bounded and property (B.10) holds.

REMARK B.1. *The construction of the uniformly bounded operator $\Pi : \Sigma \to \gamma_n(V_h)$ in step 1 above closely resembles that of the Fortin operator in [5, Proposition 8.4.3]. In fact, the operator $\Phi \circ \gamma_n : V \to V_h$ effectively acts as a Fortin operator in the proof of Lemma 4.1.*

REFERENCES

[1] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, Pure and Applied Mathematics, vol. 140, Academic Press, Elsevier, 2003.

[2] S. Adly and D. Goeleven, *A discretization theory for a class of semi-coercive unilateral problems*, Numer. Math. **87** (2000), 1–34.

[3] C. Amrouche and V. Girault, *Decomposition of vector spaces and application to the Stokes problem in arbitrary dimension*, Czech. Math. J. **44** (1994), no. 1, 109–140.

[4] L. Belenki, L. C. Berselli, L. Diening, and M. Růžička, *On the finite element approximation of p-Stokes systems*, SIAM J. Numer. Anal. **50** (2012), no. 2, 373–397.

[5] D. Boffi, F. Brezzi, and M. Fortin, *Mixed finite element methods and applications*, Springer Series in Computational Mathematics, Springer, 2013.

[6] S. Brenner and R. Scott, *The mathematical theory of finite element methods*, Texts in Applied Mathematics, Springer New York, 2007.

[7] O. Chadli, J. Gwinner, and N. Ocharova, *On semicoercive variational-hemivariational inequalities—existence, approximation, and regularization*, Vietnam J. Math. **46** (2018), 329–342.

[8] Q. Chen, M. Gunzburger, and M. Perego, *Well-posedness results for a nonlinear Stokes problem arising in glaciology*, SIAM J. Math. Anal. **45** (2013), no. 5, 2710–2733.

[9] P. Coorevits, P. Hild, K. Lhalouani, and T. Sassi, *Mixed finite element methods for unilateral problems: convergence analysis and numerical studies*, Math. Comput. **71** (2002), no. 237, 1–25.

[10] G. G. de Diego, P. E. Farrell, and I. J. Hewitt, *Numerical approximation of viscous contact problems applied to glacial sliding*, J. Fluid Mech. **938** (2022), A21.

[11] E. Di Nezza, G. Palatucci, and E. Valdinoci, *Hitchhiker's guide to the fractional Sobolev spaces*, B. Sci. Math. **136** (2012), no. 5, 521–573.

[12] Lars Diening and Michael Ruzicka, *Interpolation operators in Orlicz-Sobolev spaces*, Numerische Mathematik **107** (2007), no. 1, 107–129.

[13] G. Durand, O. Gagliardini, B. de Fleurian, T. Zwinger, and E. Le Meur, *Marine ice sheet dynamics: Hysteresis and neutral equilibrium*, J. Geophys. Res. - Earth **114** (2009), no. F3.

[14] A. Ern and J. L. Guermond, *Finite element quasi-interpolation and best approximation*, ESAIM: M2AN **51** (2017), no. 4, 1367–1385.

[15] L. C. Evans, *Partial differential equations*, Graduate studies in mathematics, American Mathematical Society, 1998.

[16] L. Favier, O. Gagliardini, G. Durand, and T. Zwinger, *A three-dimensional full Stokes model of the grounding line dynamics: effect of a pinning point beneath the ice shelf*, Cryosphere **6** (2012), no. 1, 101–112.

[17] G. Fichera, *Boundary value problems of elasticity with unilateral constraints*, pp. 391–424, Springer Berlin Heidelberg, Berlin, Heidelberg, 1973.

[18] A. C. Fowler, *A sliding law for glaciers of constant viscosity in the presence of subglacial cavitation*, P. Roy. Soc. A-Math. Phy. **407** (1986), no. 1832, 147–170.

[19] ———, *Mathematical geoscience*, Interdisciplinary Applied Mathematics, Springer London, 2011.

[20] O. Gagliardini, D. Cohen, P. Raback, and T. Zwinger, *Finite-element modeling of subglacial cavities and related friction law*, J. Geophys. Res. - Earth **112** (2007), no. F2.

[21] J. Garbe, T. Albrecht, A. Levermann, J. F. Donges, and R. Winkelmann, *The hysteresis of the Antarctic Ice Sheet*, Nature **585** (2020), 538–544.

[22] V. Girault and L. R. Scott, *A quasi-local interpolation operator preserving the discrete divergence*, Calcolo **40** (2003), no. 1, 1–19.

[23] J. W. Glen, *The flow law of ice: a discussion of the assumptions made in glacier theory, their experimental foundation and consequences*, Physics of the Movement of Ice: Symposium at Chamonix 1958, Int. Assoc. Hydrol. Sci., Wallingford, UK, 1958, pp. 171–183.

[24] R. Glowinski, J. L. Lions, and R. Trémoliéres, *Numerical analysis of variational inequalities*, North Holland, 1981.

[25] R. Greve and H. Blatter, *Dynamics of ice sheets and glaciers*, Springer Science & Business Media, 2009.

[26] J. Gwinner, *Discretization of semicoercive variational inequalities*, Aequationes Math. **42** (1991), 72–79.

[27] J. Haslinger, I. Hlaváček, and J. Nečas, *Numerical methods for unilateral problems in solid mechanics*, Finite Element Methods (Part 2), Numerical Methods for Solids (Part 2), Handbook of Numerical Analysis, vol. 4, Elsevier, 1996, pp. 313–485.

[28] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth Newton method*, SIAM J. Optimiz. **13** (2002), no. 3, 865–888.

[29] A. Hirn, *Approximation of the p-Stokes equations with equal-order finite elements*, J. Math. Fluid Mech. **15** (2013), no. 1, 65–88.

[30] I. Hlaváček, *Dual finite element analysis for semi-coercive unilateral boundary value problems*, Aplikace Matematiky **23** (1978), 52–71.

[31] G. Jouvet, *Modélisation, analyse mathématique et simulation numérique de la dynamique des glaciers*, Ph.D. thesis, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, 2010.

[32] G. Jouvet and J. Rappaz, *Analysis and finite element approximation of a nonlinear stationary Stokes problem arising in glaciology*, Adv. Numer. Anal. **2011** (2012).

[33] N. Kikuchi and J. T. Oden, *Contact problems in elasticity*, Society for Industrial and Applied Mathematics, 1988.

[34] F. Pattyn, C. Schoof, L. Perichon, R. C. A. Hindmarsh, E. Bueler, B. de Fleurian, G. Durand, O. Gagliardini, R. Gladstone, D. Goldberg, G. H. Gudmundsson, P. Huybrechts, V. Lee, F. M. Nick, A. J. Payne, D. Pollard, O. Rybak, F. Saito, and A. Vieli, *Results of the marine ice sheet model intercomparison project, MISMIP*, The Cryosphere **6** (2012), no. 3, 573–588.

[35] C. Ritz, T. Edwards, G. Durand, A. Payne, V. Peyaud, and R. Hindmarsh, *Potential sea-level rise from Antarctic ice-sheet instability constrained by observations*, Nature **528** (2015).

[36] C. Schoof, *The effect of cavitation on glacier sliding*, P. Roy. Soc. A-Math. Phy. **461** (2005), no. 2055, 609–627.

[37] ———, *Ice sheet grounding line dynamics: steady states, stability, and hysteresis*, J. Geophys. Res. - Earth **112** (2007), no. F3.

[38] ———, *Marine ice-sheet dynamics. Part 1. The case of rapid sliding*, Journal of Fluid Mechanics **573** (2007), 27–55.

[39] ———, *Coulomb friction and other sliding laws in a higher order glacier flow model*, Math. Mod. Meth. Appl. S. **20** (2010), no. 01, 157–189.

[40] W. Spann, *Error estimates for the approximation of semicoercive variational inequalities*, Numer. Math. **69** (1994), 103–116.

[41] A. G. Stubblefield, M. Spiegelman, and T. T. Creyts, *Variational formulation of marine ice-sheet and subglacial-lake grounding-line dynamics*, J. Fluid Mech. **919** (2021), A23.

[42] T. Van Bon, *Finite element analysis of primal and dual variational formulations of semicoercive elliptic problems with nonhomogeneous obstacles on the boundary*, Aplikace Matematiky **33** (1988), 1–21.

[43] J. Weertman, *On the sliding of glaciers*, J. Glaciol. **3** (1957), no. 21, 33–38.

[44] ———, *Stability of the junction of an ice sheet and an ice shelf*, J. Glaciol. **13** (1974), no. 67, 3–11.

[45] L. K. Zoet and N. R. Iverson, *Rate-weakening drag during glacier sliding*, J. Geophys. Res. - Earth **121** (2016), no. 7, 1206–1217.