Response #2 to Reviewer RC2 – Prof. Christian Schoof

Egusphere-2024-1052 Article

A PDF of the revised manuscript is attached

Author Revisions and Responses are highlighted in Blue

I again thank Prof. Schoof for a detailed and insightful review!

Higher level comments

This paper has improved in readability from the original submission. There do remain a number of significant issues to be addressed. I have to say that this is the first time I've found myself reviewing the "response to reviewers" as much as the revised paper, so don't take the points I make lightly.

Before I get started, let me say this: decisions on manuscripts are obviously up to the editor, but there are two issues without which I won't be endorsing publication of this paper.

The first is adequate citation of prior literature. My general comment about the manuscript being ungenerous to the prior literature resulted in the following line in the response to reviewersI am happy to add additional or more appropriate references. Any sugggestions?

but several pages later, when I actually list the relevant references, you say

I'm reluctant to add a lot of extra references

Most scientists find out sooner or later that it's easier in the review process to follow the path of least resistance when being asked to cite additional papers. If those are the reviewer's papers, you have a point in questioning their motives, but if they are third party papers and you try to avoid referencing them, you risk appearing as though you'd rather not acknowledge them. In the present manuscript, that unwillingness seems to focus on relevant literature on variational formulations that predates the early 2010s. That is, coincidentally, when Dukowicz, Price and Lipscomb ("DPL 2010") was first published, which you are using as the standard reference for variational formulations. The relevant citations are (and this is simply the most important subset!) J. Colinge and J. Rappaz, A strongly nonlinear problem arising in glaciology, Math. Model. Numer. Anal. 33 (1999) 395–406.

R. Glowinski and J. Rappaz, Approximation of a nonlinear elliptic problem arising in a non-Newtonian fluid flow model in glaciology, Math. Model. Numer. Anal. 37 (2003) 175–186.J. Rappaz and A. Reist, Mathematical and numerical analysis of a three-dimensional fluid flow model in glaciology, Math. Model. Mech. Appl. Sci. 15 (2005) 37–52.

The obvious place for these is line 334 just before eq (34). This is first and foremost a matter of respect. If that is too many papers, leave out Glowinski and Rappaz.

I have added the Glowinski and Rappaz, and the Rappaz and Reist references.

The second is the naming of the "BP+ model". As you point out, Herterich's (1987) model differs from Blatter-Pattyn. ("Needless to say, I will not be changing the name of the Blatter- Pattyn model.") But the Herterich model is what you now want to call "BP+", even though Herterich developed it nearly forty years ago, the best part of a decade before Blatter. It was *not* developed from the BP model, which is precisely what calling it the "BP+" model would suggest to future generations. The person who actually came up with it deserves a little more respect.

You are right. I have changed BP+ to Herterich throughout.

I have two main scientific points, some of which are repeated in greater detail in the line-by-line comments (which I wrote prior to this summary)

1. The inf-sup condition: Your response to reviewers states
   I have eliminated Appendix D and rewritten Section 4 to bypass the difficult issue of the inf-sup condition. The inf-sup condition is relevant only so as to point out that the standard and transformed Stokes models are subject to it and that one must use one of the many inf-sup-stable elements available in the literature in the dis- cretization.
   I hope I have made it clear that the inf-sup condition does not apply to problems using elements satisfying the "solvability condition" because they are no longer con- strained problems since incompressibility is built-in when using $w = (u, v)$
   This is deeply problematic as a rationale for not addressing the inf-sup condiiton. For one thing, you *do not* state ambiguously that you regard the ability to formulae the problem as an unconstrained minimzation problem as a reason for not addressing whether your novel finite elements satisfy the inf-sup condition. But worse is that the argument given above is incorrect.
   First of all, Stokes flow problems of the type discussed here are *always* equivalent to an unconstrained minimzation problems. See Chen et al 2013 for the continuum case. This reamins true for the discretized version so long as the discrete "divergence" operator given by B^T in equiation (55) has a non-trivial right nullspace, meaning there are vectors $\mathbf{u} \neq \mathbf{0}$ such that $B^T\mathbf{u} = \mathbf{0}$. Being able to find such vectors does not require the solvability condition (57) to hold, or even the usual inf-sup condition to hold (more on this shortly), but the existence of such a nullspace is assured for any set of basis functions satisfying the inf-sup condition. When the nullspace exists, the unconstrained minimization problem is over that nullspace. I have written out additional detail on this in thespecific comments on lines 611 and 617.).
   Second, there is no reason why the equivalence with an unconstrained minimization problem should get around having to address whether your new basis functions satisfy the inf-sup condition. First, let me state a subtely to the inf-sup condition, which I touched on in my previous review: I can impose the inf-sup condition at a particular discretization level, sot hat the inf-sup bound (usually "β) exists for that disretization level, and hence my discretized problem is equivalent to a finite-dimensional unconstrained minimization, but that doesn't mean that I will get correct solution. For the partial differential equations, I have to make sure that the inf-sup (usually "β) does not depend on the discretization level, usually expressed in terms of the maximum element size h. This ensures that as you take finer and finer meshes, the discretized solution converges to a solution of the continuum problem, while you have no such assurance if $\beta = \beta_h$ depends on h and $\beta_h \to 0$ as $h \to 0$ (so long as your unconstrained finite element basis can adequately approximate all functions of interest in the limit $h \to 0$, see below). If you instead treat your discretization as defining a divergence free basis spanning a finite-dimensional subspace (the right nullspace of B^T) over which you do unconstrained optimization, then nature of the problem remains the same: you have to show that solutions converge to the continuum solution as $h \to 0$. It is only

the way that you cast the problem that changes: instead of showing your basis functions satisfy the inf-sup condition with a β indepndent of h, you now need to determine whether you new, divergence-free basis functions (which are a subset of standard piecewise linear or quadratic bases) can still approximate all functions in the admissible space

$$V = \{v \in [W^{1+1/n}(\Omega)]^3 : \mathrm{div} v = 0 \text{ and } v = 0 \text{ on } \partial\Omega_b\ ,\ v \cdot n = 0 \text{ on } \partial\Omega_b\ \},$$

arbitrarily well in the limit $h \to 0$. That is *not* trivial; if it was, it's unclear anyone would ever have bothered with the inf-sup condition because it's much easier to find finite element functions for which there is an h-dependent β_h for which you have equivalence to a finite-dimensional unconstrained minimizatoin problem but no guarantee of convergence as $h \to 0$, than it is to find basis functions for which β is independent of mesh size. See the line-by-line comment for line 617.

This brings me back to: you talk about the inf-sup condition, but you still do not explicitly address the question of whether your novel finite element basis functions satisfy the inf-sup condition, and if they don't, exactly why that should not matter. See above re the unconstrained minization idea being a red herring. There are furhter specific notes related to this in the comments on line 544–617. I also note that your off-hand comment on line 1006 suggests that at least some of the finite element bases that you are constructing don't satisfy the inf-sup condition, so this isn't a flippant point to make, and you should at least point out there that the P2-E1 space is therefore likely not to satisfy the inf-sup condition, but your numerics give you hope that perhaps the solution in general does converge for the velocity field, if not the pressure field. (By "in general", I mean)

This is the main issue to be addressed. Many of your comments here and below pertain to the inf-sup condition and the solvability condition introduced here. It's a highly technical discussion; to use your expression, it's really "beyond my pay grade". Fortunately, however, I believe I've made progress as explained in the new Appendix B and summarized below.

Benzi et al. (2005), Gerbeau (2025) discuss an alternative method (the dual problem) for the solution of the Stokes saddle point problem. An equation for the pressure is solved first, followed by the solution for velocity. The equation for the pressure (the Schur complement with respect to the pressure p) is solvable provided the inf-sup condition is satisfied. The standard inf-sup condition is shown to be equivalent to $Ker\mathbf{B}^T = \{0\}$, i.e., the absence of a null space. This identifies the source of the difficulty leading to the inf-sup condition as the existence of a null space (i.e., spurious pressure modes in the B^T matrix in equation (54). I changed B to B^T in this equation to conform to standard practice in the literature.)

On the other hand, if the solvability condition is satisfied then $w(u)$ may be used to convert the Stokes model into an unconstrained minimization problem for the horizontal velocity u. Having obtained u, an equation for the pressure is obtained that involves an invertible matrix $\mathbf{M}_{WP}$, which means that the pressure problem in now well posed (no spurious modes). Thus, the entire Stokes problem in now well posed and there is no need to satisfy an inf-sup condition. Note that both the Schur complement equation for the pressure and the unconstrained minimization problem for the velocity involve solving dense matrix problems whose solution is possible but impractical in practice.

The only problem is that the material in Gerbeau (2025) is only available online as lecture notes for a course at Stanford University. Some of the material is available in Benzi et al. (2005) but not all. I reached out to Prof. Gerbeau but he said that the material was prepared so long ago that he did not remember the original source. If you know a better source for this material I would be very grateful.

The Herterich ("BP+") model. I flagged in y original review that this is an *ad hoc* model for which there is no theoretical justification in terms of asymptotic error estimates; it just turns out to work well for the test cases you have run. Your response to reviewers says *I doubt that it's necessary to have a full scale analysis when introducing a new approximation. For example, the Blatter- Pattyn model did not have a scale analysis for 15 years until Schoof and Hindmarsh (2010).*

I read that to say it is unreasonable and unnecessary to expect a scaling analysis, and that this would somehow be difficult, and that you do not wish to discuss the issue. Let me do the scaling analysis for you, in that case; you will need to discuss this when introducing the model 6.2.1 (where the section heading claims Herterich to be an improved Blatter-Pattyn model).

The Herterich model as you state it in equations (67)–(68) can be derived from the following form of the Stokes model, with terms selectively removed:

$$\begin{aligned}
\frac{\partial \tau_{xz}}{\partial z} + \frac{\partial \tau_{xx}}{\partial x} - \frac{\partial p}{\partial x} &= 0 \\
\frac{\partial \tau_{zz}}{\partial z} - \frac{\partial p}{\partial z} &= -\rho g \\
\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} &= 0 \\
\tau_{xz} = \eta\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right) \quad \tau_{xx} &= 2\eta\frac{\partial u}{\partial x} \\
\tau_{zz} &= 2\eta\frac{\partial w}{\partial z}
\end{aligned}$$

subject to

$$\begin{aligned}
\tau_{xz} - \frac{\partial z_s}{\partial x}(\tau_{xx} - p) &= 0 \\
\tau_{zz} - p &= 0
\end{aligned}$$

at $z = z_s$, and appropriate basal boundary conditions that I won't write out in detail.

Assume for simplicity that eta is constant; the shear-thinning power law fluid case of Glen's law is only superficially different. If I take the full Stokes model instead and apply a standard "shallow ice" scaling, I obtain (see e.g. section 3.6 of Schoof and

Hindmarsh (2010), although the original references would be Fowler and Larson (1978), Morland and Johnson (1980) and Hutter (1983))

$$
\begin{aligned}
\frac{\partial \tau^*_{xz}}{\partial z^*} + \epsilon^2 \frac{\partial \tau^*_{xx}}{\partial x^*} - \frac{\partial p^*}{\partial x^*} &= 0 \\
\epsilon^2 \frac{\partial \tau^*_{xz}}{\partial z^*} + \epsilon^2 \frac{\partial \tau^*_{zz}}{\partial z^*} - \frac{\partial p^*}{z^*} &= -1 \\
\frac{\partial u^*}{\partial x^*} + \frac{\partial w^*}{\partial z^*} &= 0 \\
\tau^*_{xz} &= \frac{\partial u^*}{\partial z^*} + \epsilon^2 \frac{\partial w^*}{\partial x^*} \\
\tau^*_{xx} &= 2 \frac{\partial u^*}{\partial x^*} \\
\tau^*_{zz} &= 2 \frac{\partial w^*}{\partial z^*}
\end{aligned}
$$

subject to

$$
\begin{aligned}
\tau^*_{xz} - \frac{\partial z^*_s}{\partial x^*} (\epsilon^2 \tau^*_{xx} - p^*) &= 0 \\
\epsilon^2 \tau^*_{zz} - p^* - \epsilon^2 \frac{\partial z^*_s}{\partial x^*} \tau^*_{xz} &= 0
\end{aligned}
$$

where the asterisks denote scaled variables, and red colour marks the terms that are not retained in the Herterich model. It should be apparent that the Herterich model retains *some* of the $O(\epsilon^2)$ (the $\partial w^* / \partial x^*$ term in the definition of τ^*_{xz} and the full slope term in the shear stress boundary condition. It retains by no means all $O(\epsilon^2)$ terms, so there is no reason why the Herterich model should be any better in this parametric limit (with respect to sliding) than the BP model, which also selectiely omits $O(\epsilon^2)$ terms and therefore has an $O(\epsilon^2)$ error.

We can conversely turn to the limit of fast sliding. Using the appropriate scaling for that case (e.g. Schoof and Hindmarsh (2010) section 3.4, although the original source for this goes back to MacAyeal (1989) and beyond)

$$
\begin{aligned}
\frac{\partial \tau^*_{xz}}{\partial z^*} + \frac{\partial \tau^{**}_{xx}}{\partial x^*} - \frac{\partial p^*}{\partial x^*} &= 0 \\
\epsilon^2 \frac{\partial \tau^*_{xz}}{\partial z^*} + \frac{\partial \tau^{**}_{zz}}{\partial z^*} - \frac{\partial p^*}{\partial z^*} &= -1 \\
\frac{\partial u^{**}}{\partial x^*} + \frac{\partial w^{**}}{\partial z^*} &= 0 \\
\epsilon^2 \tau^*_{xz} &= \frac{\partial u^{**}}{\partial z^*} + \epsilon^2 \frac{\partial w^* *}{\partial x^*} \\
\tau^{**}_{xx} &= 2 \frac{\partial u^{**}}{\partial x^*} \\
\tau^{**}_{zz} &= 2 \frac{\partial w^{**}}{\partial z^*}
\end{aligned}
$$

subject to

$$
\begin{aligned}
\tau^*_{xz} - \frac{\partial z^*_s}{\partial x^*}(\tau^{**}_{xx} - p^*) &= 0 \\
\tau^{**}_{zz} - p^* - \epsilon^2 \frac{\partial z^*_s}{\partial x^*} \tau^*_{xz} &= 0
\end{aligned}
$$

where the variables with double asterisks denotes dimensionless variables that have been rescaled from the shallow ice case. Once more, the Herterich model retains some of the $O(\epsilon^2)$ terms, but not all, and will again only be accurate to an $O(\epsilon^2)$ error, the same as the Blatter-Pattyn model.

We can also look at the intermediate sliding regime (Schoof and Hindmarsh (2010), section 2.1, formally with lambda = 1), in which case

$$
\begin{aligned}
\frac{\partial \tau^*_{xz}}{\partial z^*} + \epsilon \frac{\partial \tau^{***}_{xx}}{\partial x^*} - \frac{\partial p^*}{\partial x^*} &= 0 \\
\epsilon^2 \frac{\partial \tau^*_{xz}}{\partial z^*} + \epsilon \frac{\partial \tau^{***}_{zz}}{\partial z^*} - \frac{\partial p^*}{\partial z^*} &= -1 \\
\frac{\partial u^{***}}{\partial x^*} + \frac{\partial w^{***}}{\partial z^*} &= 0 \\
\epsilon \tau^*_{xz} &= \frac{\partial u^{***}}{\partial z^*} + \epsilon^2 \frac{\partial w^{***}}{\partial x^*} \\
\tau^{***}_{xx} &= 2 \frac{\partial u^{***}}{\partial x^*} \\
\tau^{***}_{zz} &= 2 \frac{\partial w^{***}}{\partial z^*}
\end{aligned}
$$

on Ω, subject to

$$
\begin{aligned}
\tau^*_{xz} - \frac{\partial z^*_s}{\partial x^*}(\epsilon \tau^{***}_{xx} - p^*) &= 0 \\
\epsilon^* \tau^{***}_{zz} - p^* - \epsilon^2 \frac{\partial z^*_s}{dx^*} \tau^*_{xz} &= 0
\end{aligned}
$$

Same story. The Herterich model selectively retains $O(\epsilon^2)$ terms and offers no

improvement in terms of asymptotic error over the Blatter-Pattyn model.
That should settle the case — if you want a single model with a better asymptotic error (that is, size of error in the limit of $\epsilon \to 0$ than Blatter-Pattyn, you really have to retain all $O(\epsilon^2)$ terms. In order to do that regardless of sliding regime, it should be clear you actually have to retain all terms of the full Stokes model (since the smallest term contained in all the scaled versions of the Stokes flow model above is $O(\epsilon^2)$, there is really nothing that you can omit if you want a model accurate to $O(\epsilon^2)$).

The fact that the Herterich model performs better than BP in specific tests you have computed is really not much of a guide to anything in general: if you had a general result predicting a smaller error, then something like figure 9 would be a welcome illustration. In the absence of a much more general investigation of model error, the smaller error here might as well be fortuitious and specific to some aspect of test B. (I would also note that figure 5 shows resutls for a fairly large epsilon, so we are not looking an approximation based on small aspect ratio: an aspect ratio of 1:5 really cannot be treated as small in general, as finite depth effects typically appear for aspect ratios around 1 : (2π))

My sense of why the Herterich model is possibly an improvement over Blatter-Pattyn comes from the fact that the Herterich model restores a term dropped from the Stokes functional by an approximation of that term. This sense is probably invalid because it's the Euler-Lagrange equations derived from the functional and not the functional itself that carry the important information. I agree that a scale analysis of the equation could be helpful. However, I still feel that it would be well beyond the scope of the paper as is. What I have done instead is expand the comparison of the BP and Herterich approximations in Fig. 9 by adding results for $L = 2\,km$, i.e., an even higher problem aspect ratio, thereby stressing small aspect ratio approximations further. The results are surprising. While the Herterich approximation is clearly an improvement in Test B at $L = 5,10\,km$, Blatter-Pattyn is better at $L = 2\,km$ in both Tests B and D*. As pointed out in the paper, these results make the Herterich approximation somewhat problematic and so more experience would be helpful.

Line-by-line comments

- eq (11): there is still no need for the brackets around the (i) subscripts. In fact, by putting them into equation, you actually need to add the statement that $\tau_i^S n_i = 0$ since (11) is mute about the value of its z-component. If you write instead that

$$\tau_i^S = \tau_{ij} n_j - \tau_n n_i$$

you automatically ensure that $\tau_i^S n_i = \tau_{ij} n_i n_j - \tau_n n_i n_i = \tau_n - \tau_n = 0.$

That also tidies up your notation.

Section involving equations (10)-(13) has been rewritten.

- line 188 ”references contained therein in connection with the Blatter-Pattyn model” — No, not enough. See start of this re-review

References added, see above.

- line 203: ”However, this can only be done in the discrete formulation of the functional since only then are boundary values of velocity accessible (although they are always accessible in the surface integral terms).” This either not true or misleading, depending on what you mean by ”accessible”. Infinite-dimensional variational formulations of pde problems impose homogenous Dirichlet conditions by restricting the function space on which the problem is posed, rather than using Lagrange mutlipliers. In the present case, no need to get technical about that function space, but simply delete this statement and state that the minimization problem needs to be restricted first arguments of the functional A that satisfy the Dirichlet conditions

Lines 197-210 have been rewritten. Hope this clarifies the situation.

• eq (15) The (b2) superscript on the second line is undefined and doesn’t want to be there. Also, this is needlessly complicated: the form of the surface integral is exactly the same as

$$\int_{S_{B_2}} \frac{1}{2}\beta u_i u_i \mathrm{d}S,$$

but more complicated. And it’s not like you’ve replaced u_3 in the rest of the formula yet, so why make it more complicated in the boundary term. (This is different from (33), where the substitution makes more sense.)

Subscript corrected. I believe that the need for the more complicated form has been addressed. It’s needed because the surface integral is over the section where the frictional boundary condition applies and this boundary condition involves the replacement $w \rightarrow -u_{(i)}n_{(i)}/n_z$ as explained in the text.

- eq (17) seems redundant. If you cannot see follows from (16) and (18), you’re going to struggle mightily with the rest of this paper.

Equation (17) has been removed.

- eq (26) Again, write in standard subscript form?

Equations (26), (27) replaced by a single equation for the stress tensor.

- line 287 ”remains positive” → ”remains non-negative”

Corrected.

- line 289 ”The dummy variables...” — I don’t think this is the usual meaning of ”dummy variables” (like the variable with respect to which you integrate in a definite integral). They are indicator terms, taking values of 1 or 0. Same in line 344.

I think “dummy variables” is appropriate here.

- eq (29) — this is a bit misleading; your earlier text (line 289, ”The dummy variables $\xi,\hat{}$in (23)-(25) and (26)-(29) are used to identify terms that are neglected in the two types of the Blatter-Pattyn approximation discussed in §3.4”) indicates that all I need to do in order to obtain Blatter-Pattyn is to set $\xi = \hat{\xi} = 0$.

The paragraph regarding dummy variables has been shifted to after equation (31).

- Remark #1 — this isn’t about ”computational savings”, is it? The small slope approximation is what makes Blatter-Pattyn valid in the first place, and the error in replacing the normal by its small slope approximation is the same order as the error in the Blatter-Pattyn model in the first place. Besides — and you might want to spell this out clearly — the difference between the two amounts to replacing

  $\beta(x)$ by $\beta(x)1 + (\partial b/\partial x)^2$.

Remark #1 has been eliminated.

- line 405 and elsewhere throughout the paper. Since you’re focused on finite elements, I really think that ”mesh” rather than ”grid” would be the right choice here, since ”grid” is usually taken to imply a greater degree of regularity than a mesh. Plus, ”mesh” would seem to be standard parlance in finite elements. (Your response to reviewers suggested at one point that you had replaced “grid”, but that certainly doesn’t seem to have been done consistently)

“grid” has been replaced by “mesh” throughout.

- line 425 ”We follow this practice except that here the discretization originates from a variational principle. This has a number of advantages (see §2.3 and DPL, 2010).” — I expect I complained about this in the original submission, but in any case, this is likely a mischaracterization. Any finite element discretization that is based on the standard weak form of the Stokes equations will have the properties listed in section 2.3: boundary equations formulated in terms of surface integrals (rather than inaccurate one-sided expressions), and a symmetric stiffness matrix. That remains true regardless of whether I am alert enough to recognize (or bother to say out loud, for that matter) that the weak form is the Frechet derivative of the functional being minimized or not. If my discretization does something different, then I am likely not using standard finite elements. I don’t see that there is much of a grey zone on this — if you know of an example of a paper using finite elements in a way that does not lead to a discretization with the stated properties, be sure to explicitly identify that paper, and where in that paper I might be able to find evidence of a non-standard finite element formulation (I'd be interested!). Otherwise omit this passage, unless you wish to reiterate the statement about the "extended Blatter-Pattyn model" solving for w and the dynamically rather redundant reduced pressure $\tilde{P}$.

I have rewritten this section. It is my understanding that a finite element method based on the weak form is not equivalent to a variational principle unless the Galerkin method is used (trial functions the same as test functions). Finite element methods are used for many problems that, unlike the Stokes problem, do not have a variational formulation. See the Zienkiewicz and Taylor book, " The Finite Element Method – Volume 1: The Basis", particularly Section 3.8.2.

- line 455 "in matrix form": misleading as M hasn't been cast in matrix form

Word "matrix" changed to "discrete".

- line 459 "M(u, w) is a nonlinear positive-definite function of the velocity" — this is an odd thing to say as I don't think I can have a nontrivial linear positive- definite function, unless it's defined on some bounded set of values (u, w). To wit, if M(u, w) > 0 and M was linear, then it would follow that M(−u, −w) < 0 by linearity, no? Do you mean "convex"?

I wanted to emphasize that the function is both nonlinear and positive-definite. To make it unambiguous, I've changed "nonlinear" to "highly nonlinear".

- line 466 "Discrete variation of the functional corresponds to partial differentiation with respect to each of the discrete variables in V ." Partial differentiation of . . . the discretized functional? It's unclear who the audience is, but probably best to be specific.

I've referred back to equation (43) to be specific.

- equation (48) It's probably a bad idea to use the same letter M for the dissipation potential part of the discretized functional A, and for the Hermitian matrix of the discretized functional, at the same time, even if one M is in calligraphic and the other is not. Also note that the Cryosphere has some rather inflexible (and unimaginative) rules around renderinng matrices and vectors in bold face, in in upright and the other in italicized font. Be prepared fo a bunch of fun at the copyediting stage.

Good point! I've changed all matrices M to bold $\mathbf{M}$ to avoid confusion with script $\mathcal{M}$ for functionals. However, I'm sure there will be lots of problems with typesetting.

- line 519 onwards ”The form (55) is characteristic of Stokes-type problems, or more generally of constrained minimization problems using Lagrange multipliers. In finite element terminology these are called “mixed” or “saddle point” problems, meaning that velocity components and the pressure occupy different finite element spaces, and that the solution of (55) is actually at the saddle point with respect to the velocity and pressure variables of the quadratic form associated with (55). The matrix M is symmetric but indefinite, with both positive and negative eigenvalues. As a result, the matrix inverse may not be bounded and may lack stability” There are a number of things that are problematic here. The first is the statement that the generic form of defines a saddle point problem. That is not true unless you add that A has to be at least positive semi-definite. The next is the statement that ”these are called mixed or saddle point problems, meaning that velocity components and the pressure occupy different finite element spaces.” That is not the meaning of a saddle point problem. A saddle point problem corresponds to a problem of minimizing with respect to one variable (velocity) and maximizing with respect to another (pressure). You also can’t say that the matrix inverse may not be bounded (if the inverse of a matrix exists, it is bounded!) — the point you’re presumably getting at without spelling it out is this: implicit in the discretization is that you are looking at a family of discretizations of progressively of different resolutions resolution (parameterized by maximum element size, call it h, combined with a non-zero lower bound on internal element angles), and you want the corresponding family of saddle points to have a unique limit as $h \rightarrow 0$. The statement obout the bound on the inverse of M (and this is what the inf-sup condition guarantees, in my understanding) is that M^{-1} remains bounded as $h \rightarrow 0$. No?

The entire section, old lines 513-559, has been rewritten. This applies to

- line 530 ”In this case only the matrix A exists, it is elliptic” — two things: i) you seem to be saying that A only exists in this case and ii) ”elliptic” is a term you haven’t defined up to now, I think. Symmetric positive-definite?

Eliminated

- also line 530 ”As a result, the standard Blatter-Pattyn model is a well-behaved and stable unconstrained minimization problem” — this is an unfortunate choice of words, since you’ve just told the reader that $B = B^T = 0$, in which case M is manifestly not invertible; it seems like you have to state that you’re not actually solving (55) but only the reduced version $Au = f$, for which the statement in question is true.

Corrected

- line 544 “In fact, this is a problem for all inf-sup stable elements with $n_p \neq n_w$, such as the Taylor-Hood element, for example” two things i) you have not de- fined what you mean by ”inf-sup stable elements” and ii) the second half of the sentence about Taylor Hood elements is redundent. You could say "is a problem for all inf-sup stable elements with $n_p \neq n_w$". Comment on whether $n_p \neq n_w$ is a prerequisite for an element pair to be inf-sup stable (which I think is true); the last sentence grammatically leaves open the possibility that you could have $n_p = n_w$. If that is the case, comment on the implications for your choice of basis functions, which do have $n_p = n_w$. Are they necessarily not inf-sup stable? This requires an explicit comment, not a response that says "I don't want to address this issue". See discussion re unconstrained minimization being a red herring, unless you can show that your divergence-free basis functions are dense in the underlying function space.

In terms of sequencing the three items (1) Blatter-Pattyn (2) extended Blatter-Pattyn and (3) Stokes / transformed Stokes, it would actually be preferable to reverse the current order since that would mean introducing inf-sup conditions before you talk about them for extended Blatter-Pattyn

The inf-sup condition section has been moved earlier, shortly after equation (54). The discussion of $n_p = n_w$ refers only to the extended Blatter-Pattyn model, while the discussion of meshes satisfying the solvability condition is deferred to the next section.

- line 552 "Both the standard and transformed Stokes models are subject to this problem an in general must use inf-sup stable finite elements. Testing for stability is not trivial. However, collections of inf-sup stable elements for the Stokes equations may be found in many papers and books on mixed methods, e.g., Boffi et al. (2008)." and later line 594 "However, this model does work with a variant of the Taylor-Hood grid, the P2-E1 grid, illustrated in Fig. 13A, which does satisfy the solvability condition and this therefore allows for a successful calculation of the vertical velocity." — I think I raised this point in my original review, but I still don't see it being addressed here: are the new finite element basis functions like P2-E1 and P1-E0 inf-sup stable, or are you arguing (see comments immediately below) that this is somehow no longer relevant when you the solvability condiiton for the incompressibility condition is satisfied? Otherwise the discussion of the inf sup condition is still left dangling in thin air here. If you're referring to the many papers and books because you won't be addressing the issue here, at least say so explicitly. See also comment on line 544.

This issue is now addressed in Appendix B, as explained above. The elements, like P1-E0, that satisfy the solvability condition, may possibly also satisfy the inf-sup condition, but not necessarily. If they do, that would be fortuitous.

- Line 564 "As mentioned previously, this is possible to do in the continuum but not necessarily so in the discrete case" — I don't think you've explained that this is "not necessarily possible" to do in the discrete case, unless you mean the discussion around equation (56) that is about to follow. To avoid confusion and

> leave the reader scurrying for where this was explained previously, especially as you've just referenced section 3.4.1, and there is no mention of not being able to compute w there, and even the discussion starting on line 396 also doesn't actually say that you cannot compute w in the discrete case. In fact, I expect that those running Blatter-Pattyn ice sheet models would beg to differ with the statement that you cannot compute w from the continuity equation, since they will naturally choose a scheme that finds w by simple quadrature. This seems more of an issue with the extended Blatter-Pattyn model of section 3.4.2, where you are forcing yourself to use the same mesh as for (u, v), no?

The discussion at the start of Section 4.3.2 has been modified to address this.

- line575 "$n_p+\lambda_z+\Lambda=n_w$" should presumably be "$n_p+n_{\lambda_z}+n_\Lambda=n_w$"?

Corrected.

• line 599 "Perhaps the main reason for the importance of the solvability condition is that it implies that the Stokes variational principle, (15) or (33), may be transformed into and therefore that it is equivalent to an optimization or minimization problem." — this is at best misleading, but most likely just wrong. Finding stationary points of (15) and (33), and in fact the solution of all Stokes flow problems subject to suitable boundary conditions, is always equivalent to a convex minimzation problem, but that minimization is over an awkward vector space of Sobolev space of divergence free functions. I should: awkward in terms of finding suitable basis functions when discretizing in practice, but not particularly awkward in terms of the abstract analysis of the problem.

For the present set of boundary conditions and choice of rheology, the relevant function space is $V=\{v\in[W^{1+1/n}(\Omega)]^3 : \mathrm{div} v=0 \text{ and } v=0 \text{ on } \partial\Omega_{b1}, v\cdot n=0 \text{ on } \partial\Omega_{b2}\}$, and the cited Chen et al (2013) paper would be an appropriate reference for this fact. For the transformed Stokes flow problem, the transformed minimization would also seem to be pretty straightforward, since, if I restrict myself to the same Sobolev space as Chen et al, I obtain a restriction of A to velocities only in the form

$$J(u,v,w)=\int_V \frac{4n\eta_0}{n+1}\dot{\epsilon}^{1+1/n}+\rho g\nabla z_s\cdot(u,v,0)\mathrm{d}V+\int_{S_{b_2}}\beta|(u,v,w,)|^2\mathrm{d}\Gamma$$

and J is convex, so any stationary point of J is automatically a minimizer. If you are actually trying to say that you're trying to determine whether the *discretized* variational problem still corresponds to a minimization problem, that is a different matter, but then you can't refer to (15) or (33), which are formulated for general velocity fields.

Much of Section 4.3.2 has been rewritten. I think the revision stating that "The solvability condition can also be used to convert the Stokes variational principle, (15) or (32), from a constrained to an unconstrained minimization problem." takes care of the problem.

• line 611 “This result suggests that a conventional Stokes problem, when solved on a grid satisfying the solvability condition, is equivalent to an unconstrained minimization problem and therefore is well behaved.” As per the above, this risks sounding like you are unfamiliar with standard results in the field, and I would not put this statement into a paper.

Even if you don’t mind giving that appearance, it would still be a bad idea to leave the statement as is, because it could be read to suggest that basis functions that don’t satisfy (57) don’t correspond to a minimization problem. That would be wrong.

The conventional Stokes flow problem remains equivalent to an unconstrainted minimzation problem even after discretization. That unconstrained minimization is however over the left nullspace of the matrix B in (55) — that is, over all vectors u such that $B^T u = 0$. In general, you don’t want to have to figure out what that nullspace is: as you later point out, the basis vectors for that nullspace are no longer sparse when expressed in terms of nodal values of the mesh, so your linear algebra ends up non-sparse and therefore intractable for large problems. Instead, you continue to impose the need for your solution to lie in that nullspace through the Lagrange multipier p, but that doesn’t negate the fact that the problem is equivalent to an unconstrained minimization problem..For the discretized problem on a given, fixed mesh, the only condition for this statement to be true is that the left nullspace of B must be non-trivial, and for that to be true, you generally require the matrix to be taller than wide ($n_u + n_v + n_w > n_p$). You definitely don’t need $n_w = n_p$, though $n_w = n_p$ will do. (The solvability condition (57) conveniently ensures that the nullspace in question has dimension $n_u + n_v$ and can be written in the form given on line 620, but that is far from the only form of nullspace you could construct. In general, if $n_p < n_w$ for an inf-sup stable basis you have to use more than a basis for (u,v) to construct a divergence-free basis for (u, v, w), but that is not in principle a problem)

Again, this section has been rewritten to take into account the new Appendix B. I hope this resolves the issue.

• line 617 “if a divergence free basis exists” is a misleading way to make this statement. For the original continuum problem, the divergence free basis exists without a doubt. (Again, Chen et al 2013 are the correct reference for the particular problem you have in mind). My suspicion is that the book you are referencing has in mind a divergence free basis that is dense in the underlying function space, in the sense that it can approximate any element of $V = \{v \in [W^{1+1/n}(\Omega)]^3 : \mathrm{div} v = 0 \mathrm{and} v = 0 \mathrm{on} \partial\Omega_{b1}, v \cdot n = 0 \mathrm{on} \partial\Omega_{b2}\}$ to arbitrary accuracy simply by imposing a suitably small maximum edge length.

That is not something that you have demonstrated. You seem to be alluding to the equivalence to an unconstrained minimization problem as evidence that the problem you’re solving ”is well behaved”, without saying what you mean by that, precisely. Based on the text in your response to reviewers (though not the text here!) I assume you arguing that you don’t need to worry whether your basis functions actually satisfy the inf-sup condition because you can show that your discretized problem is equivalent to an unconstrained finite-dimensional minimzation problem.

If so, that is not: all that the equivalence with an unconstrained minimization problem that you've discussed (which applies to the discretized problem) ensures case is that the discretized problem is solvable, for any given mesh. It does not guarantee that the solutions for a family of discretized problems will converge to the solution of the continuum problem in the limit of element size h in the mesh going to zero.

More formally, let the function space spanned by your basis function be V_h for a given mesh with maximum element edge length h. The thing you need to prove is that your diverge free finite element basis functions can adequately approximate any possible solution in the limit of small h, that is

$$\inf_{v_h \in V_h} \|v - v_h\| \leq C_h \|v\|$$

where $C_h \to 0$ as $h \to 0$. In the absence of the divergence constraint $B^T u = 0$, that behaviour is well established for all sorts of polynomial basis functions, but is no longer guaranteed to hold if you restrict yourself to a linear subspace of of such polynomial basis functions. How you might prove the same behaviour for your divergence free subspace is beyond my pay grade, but you can't just skip past this issue. That would make a mockery of a lot of work by some very smart people in numerical analysis.

The original section, lines 617-622, has been eliminated.

- line 633 "These tests are described in Appendix A" — I checked the appendix as well as Pattyn et al (2008), and it's unclear to me how the "true" solution relative to which the error is computed in figure 3. And in the same vein, if there is a true solution, you should plot that in figure 4.

Inserted "converge to the same transport value, obtained by Richardson extrapolation".

- line 702 "It would be computationally cheaper to use the standard Blatter-Pattyn approximation ($\xi, \hat{\xi} = 0$) instead, solving only for the horizontal variables and coupling to the Stokes model with p = 0 and w = w(u, v) at the interface but this, however, implies much more complicated programming." — see main comment re response to reviewers. That response argues (correctly, in my view) that Blatter- Pattyn is cheaper than a Stokes flow model. Which is understandable given that it has fewer degrees of freedom to solve for. The extended Blatter-Pattyn model on the other hand is constructed to have exactly the same number of degrees of freedom as the Stokes flow model, and it is *not* obvious that it will be much cheaper to solve than the Stokes flow model (transformed or otherwise). Please provide evidence (ideally here!) that it remains sufficiently cheaper to solve than the Stokes flow model to warrant doing this. As per my main comment, using the extended Blatter-Pattyn model will clearly incur a potentially significant model er- ror relative to a Stokes flow solver, so it is important to know whether the reduction in computational cost is worth the increased error.

Included the statement that "For example, the extended BP calculation of Test B at a 40x40 resolution, as in Figs. 5 and 6, is some 4.3 times cheaper than the corresponding transformed Stokes calculation". Hope this addresses the issue.

- line 725 "Somewhat counterintuitively, the Stokes region occupies the upper part of the domain in Fig. 7 and includes the obstacle, while the Blatter-Pattyn region occupies much of the bottom part of the domain" — I am not sure this is counterintiuitive in view of the no-penetration boundary condition combined with the moderate angle: basically, your boundary condition ensures that dw/dx is small along the bed, while there is no reason why du/dz would be. I am however not convinced that this test goes far enough: really, you should also compute the Stokes solution across the domain to see

Removed the word "counterintuitively".

- line 757 "The first method, to be called the BP+ approximation" and later line 773 "An Improved Blatter-Pattyn or BP+ Approximation" line 812 "Remarkably, this same model, i.e., the BP+ approximation, was introduced by Herterich (1987)" line 1029 "Remarkably, the BP+ approximation is actually the same as a model originally proposed by Herterich (1987)." etc ... Given that Herterich predates either Blatter or Pattyn, and given that "BP+" suggests that this is somehow a development of Blatter and Pattyn, this really needs to be "Herterich's improved Blatter-Pattyn Approximation" or better just the "Herterich model". Definitely not "BP+", thereby giving further credit to Blatter and Pattyn while further consigning the person who actually invented this to a footnote? Let's have some respect.

As noted earlier, replaced "BT+" with "Herterich" everywhere.

- line 839 "Both BP+ versions converge to the same solution" — this confuses "model" (a set of equations" with "solution algorithm". These are not two dif- ferent "BP+ versions", they are different solution algorithms

Changed "versions" to "algorithms".

- line 868 "logically rectangular" — what does that mean? Producing the same graph as a rectangular grid? Define before using . . .

Replaced this with "As described in Appendix C, our test problem meshes are logically rectangular, i.e., divided into n cells horizontally and m cells vertically". Hope this is sufficient.

- line 1007 "On the other hand, the pressure in the P2-E1 case is highly oscillatory but well behaved in the P2-P1 case. However, this is not at all concerning since the transformed pressure, a Lagrange multiplier, has no physical significance" — so I think you've just demonstrated that your P2-E1 element most likely does not satisfy the inf-sup condition, and you're effectively left hoping that the lack of convergence in the limit of small element size only affects the pressure, but not the velocity. Comment?

This has been replaced with "This is consistent with the discussion in Appendix B since the P2-P1 mesh is chosen to satisfy the inf-sup condition, which is concerned with regularizing the pressure solution (i.e., a condition on the $\mathbf{B}$ matrix), while the P2-E1 mesh has no such requirement". Hope this is sufficient to clarify the issue.

- line 1011 Summary and Discussion — reading this, all is fine and dandy? Nothing that you would flag as an open question or necessary areas of future research where things need to be developed further, followed up on, etc? Any weaknesses? No?

A minor change inserted at the end concerning Appendix B but no other additions.

- Appendix B This is a follow up on a comment from my first review, the response to which was that you thought the P1-E0 element is quite flexible and can be used with quite general triangulations. The trianglulations you require are not "quite general". By construction, your mesh cannot be unstructured as you require elements to be stackable to make columns. You should comment on that as it negates one of the usual advantages of finite elements, which is to permit adaptive meshing. For instance, at grounding lines, ice stream shear margins etc, you may want to have high resolution *near the bed* around a transition in sliding behaviour, but not extend that resolution throughout a column. Moreover, if you are required to keep the interior angles of elements from becoming excessively small, your mesh also makes it somewhat difficult to have variable vertical triangle edge lengths — adjacent columns need to have similar triangle numbers and triangle edge lengths. This needs to be flagged somewhere.

This is now Appendix A. The statement "The triangulation and the configuration of the pressure basis functions (shown in gray) is quite general, allowing for a flexible triangulation of the domain" has changed to "The triangulation and configuration of pressure basis functions (shown in gray) in Fig. A2 is quite general except that elements must be arranged in vertical columns, which still allows for a flexible triangulation of an arbitrary domain" and moved from the figure caption to the main body of text.

A Novel Transformation of the Ice Sheet Stokes Equations and Some of its Properties and Applications

John K. Dukowicz
Guest Scientist, Group T-3, Los Alamos National Laboratory,
Los Alamos, New Mexico, 87545, USA
Correspondence to: John K. Dukowicz (jn.dk@outlook.com)

Abstract. We introduce a novel transformation of the Stokes equations into a form closely resembling the shallow Blatter-Pattyn equations. The two forms differ by only a few additional terms, while their variational formulations differ only by a single term in each horizontal direction. Specifically, the variational formulation of the Blatter-Pattyn model drops the vertical velocity in the second invariant of the strain rate tensor. Here we make use of the new transformation in two ways. First, we consider incorporating the transformed equations into a code that can be very easily converted from a Stokes to a Blatter-Pattyn model, and vice-versa, by switching these terms on or off. This may be generalized so that the Stokes model is switched on adaptively only where the Blatter-Pattyn model loses accuracy. Second, the key role played by the vertical velocity in the Blatter-Pattyn approximation motivates new approximations. Two examples are presented. These require a mesh that enables the discrete continuity equation to be invertible for the vertical velocity in terms of the horizontal velocity components. Examples of such meshes, such as the first order P1-E0 mesh and the second order P2-E1 mesh are given in both 2D and 3D. However, the transformed Stokes model has the same type of gravity forcing as the Blatter-Pattyn model, determined by the ice surface slope, thereby forgoing some of the mesh-generality of the traditional formulation of the Stokes model.

1 Introduction

Concern and uncertainty about the magnitude of sea level rise due to melting of the Greenland and Antarctic ice sheets have led to increased interest in improved ice sheet and glacier modeling. The gold standard is a Stokes model (i.e., a model that solves the nonlinear, non-Newtonian Stokes system of equations for incompressible ice sheet dynamics) because it is applicable to all geometries and flow regimes. However, the Stokes model is computationally demanding. It is a nonlinear, three-dimensional model involving four variables, namely, the three velocity components and pressure. The pressure is a Lagrange multiplier enforcing incompressibility and this creates a more difficult indefinite "saddle point" problem. As a result, full-Stokes models exist but are not commonly used in practice (examples are FELIX-S, Leng et al., 2012; Elmer/Ice, Gagliardini et al., 2013).

Because of this there is much interest in simpler and cheaper approximate models. There is a hierarchy of very simple models, such as the shallow ice (SIA) and shallow-shelf (SSA) models, and there are also more accurate higher-order approximations. These culminate in the Blatter-Pattyn (BP) approximation (Blatter, 1995; Pattyn, 2003), which is currently used in production code packages such as ISSM (Larour et al., 2012), MALI (Hoffman et al., 2018; Tezaur et al., 2015) and CISM (Lipscomb et al., 2019). This approximation is based on the assumption of a small ice sheet aspect ratio, i.e., $\varepsilon = H/L \ll 1$, where H, L are the vertical and horizontal length scales, and consequently it eliminates certain stress terms and implicitly assumes small basal slopes. Both the Stokes and Blatter-Pattyn models are described in detail in Dukowicz et al. (2010), hereafter referred to as DPL (2010). Although the Blatter-Pattyn model is reasonably accurate for large-scale motions, accuracy deteriorates for small horizontal scales, less than about five ice thicknesses in the ISMIP–HOM model intercomparison (Pattyn et al., 2008; Perego et al., 2012), or below a 1 km resolution as found in a detailed comparison with full Stokes calculations (Rückamp et al, 2022). This can become particularly important for calculations involving details near the grounding line where the full accuracy of the Stokes model is needed (Nowicki and Wingham, 2008). Attempts to address the problem while avoiding the use of full Stokes solvers include variable mesh resolution coupled with a Blatter-Pattyn solver (Hoffman et al., 2018) and variable model complexity, where a Stokes solver is embedded locally in a lower order model (Seroussi et al., 2012). Approximations that are more accurate than Blatter-Pattyn but cheaper than Stokes are currently not available.

The present paper introduces two innovations that may begin to address some of these issues. The first is a novel transformation of the Stokes model, described in §3, which puts it into a form closely resembling the Blatter-Pattyn model and differing only by the presence of a few extra terms. This allows a code to be switched over from Stokes to Blatter-Pattyn, and vice-versa, globally or locally, by the use of a single parameter that turns off these extra terms thus enabling variable model complexity to be very simply implemented as described in §6.1. The second innovation is the introduction of new finite element discretizations that decouple the discrete continuity equation and allow it to be solved for the vertical velocity in terms of the horizontal velocity components. Several elements used to construct such meshes are described in Appendix A in both 2D and 3D, primarily the first order P1-E0 and second order P2-E1 elements (these two elements are novel and are so-named because they employ pressures located on vertical mesh edges). Within the framework of the transformed Stokes model these meshes enable new approximations that surpass the Blatter-Pattyn approximation. We describe two such approximations in §6.2. There is another very significant benefit. These new meshes allow Stokes ice sheet models to be converted into inherently stable unconstrained minimization problems without invoking the "inf-sup" condition as shown in §4.3.2.

2 The Standard Formulation of the Stokes Ice Sheet Model

2.1 The Assumed Ice Sheet Configuration

An ice sheet may be divided into two parts, a part in contact with the bed and a floating ice shelf located beyond the grounding line. The Stokes ice sheet model is capable of describing the flow of an arbitrarily shaped ice sheet, including a floating ice shelf as illustrated in Fig. 1, given appropriate boundary conditions (e.g., Cheng et al., 2020). One limitation of the methods proposed here, in common with the Blatter-Pattyn model, will be that there should be just one upper and one basal surface, as is the case in Fig. 1. For simplicity we only consider fully grounded ice sheets with periodic lateral boundary conditions although ice shelves can be incorporated.

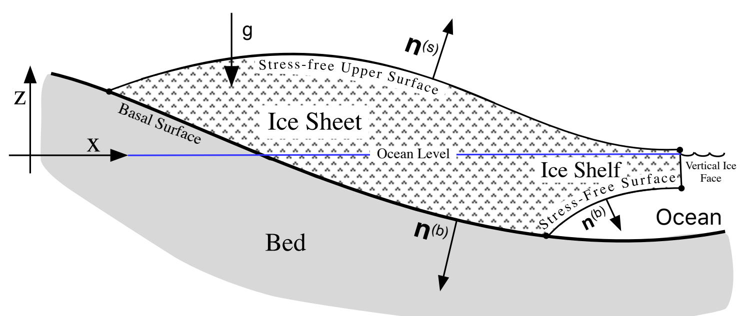

Figure 1 A simplified illustration of the admissible ice sheet configuration.

Referring to Fig. 1, the entire surface of the ice sheet is denoted by S. An upper surface, labeled S_S and specified by $\varsigma_s(x,y,z) = z - z_s(x,y) = 0$, is exposed to the atmosphere and thus experiences stress-free boundary conditions. The bottom or basal surface, denoted by S_B and specified by $\varsigma_b(x,y,z) = z - z_b(x,y) = 0$, is in contact with the bed. The basal surface may be subdivided into two sections, $S_B = S_{B1} \cup S_{B2}$, where S_{B1}, specified by $z = z_{b1}(x,y)$, is the part where ice is frozen to the bed (a no-slip boundary condition), and S_{B2}, specified by $z = z_{b2}(x,y)$, is where frictional sliding occurs. We assume Cartesian coordinates such that $x_i = (x,y,z)$ are position coordinates with $z = 0$ at the ocean surface, and the index $i \in \{x,y,z\}$ represents the three Cartesian indices. Later we shall have occasion to introduce the restricted index $(i) \in \{x,y\}$ to represent just the two horizontal indices. This is more compact than the equivalent projection operator, i.e., $u_{(i)} = P_i(u) = (u,v,0)$. Unit normal vectors appropriate for the ice sheet configuration of Fig. 1 are given by

$$
\begin{aligned}
n_i &= (n_x, n_y, n_z) = \frac{\partial\varsigma_s(x,y,z)/\partial x_i}{\left|\partial\varsigma_s(x,y,z)/\partial x_i\right|} = \frac{(-\partial z_s/\partial x, -\partial z_s/\partial y, 1)}{\sqrt{1+(\partial z_s/\partial x)^2 + (\partial z_s/\partial y)^2}} \quad \text{at surface } S_S, \\
n_i &= (n_x, n_y, n_z) = -\frac{\partial\varsigma_b(x,y,z)/\partial x_i}{\left|\partial\varsigma_b(x,y,z)/\partial x_i\right|} = \frac{(\partial z_b/\partial x, \partial z_b/\partial y, -1)}{\sqrt{1+(\partial z_b/\partial x)^2 + (\partial z_b/\partial y)^2}} \quad \text{at surface } S_B.
\end{aligned} \tag{1}
$$

2.2 The Stokes Equations

The Stokes model is a system of nonlinear partial differential equations and associated boundary conditions (Greve and Blatter, 2009; DPL, 2010). In Cartesian coordinates the Stokes equations, the three momentum equations and the continuity equation for the three velocity components $u_i = (u,v,w)$ and the pressure P, are given by

$$\frac{\partial \tau_{ij}}{\partial x_j} - \frac{\partial P}{\partial x_i} + \rho g_i = 0\,, \tag{2}$$

$$\frac{\partial u_i}{\partial x_i} = 0\,, \tag{3}$$

where ρ is the density, and g_i is the acceleration vector due to gravity, arbitrarily oriented in general but here taken to be in the negative z-direction, $g_i = \left(0,0,-g\right)$. Repeated indices imply summation (the Einstein notation). The deviatoric stress tensor τ_{ij} is given by

$$\tau_{ij} = 2\mu_n \dot{\varepsilon}_{ij} , \tag{4}$$

where the strain rate tensor is

$$\dot{\varepsilon}_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \tag{5}$$

the nonlinear ice viscosity μ_n is

$$\mu_n = \eta_0 \left(\dot{\varepsilon}^2\right)^{(1-n)/2n} , \tag{6}$$

and $\dot{\varepsilon}^2 = \dot{\varepsilon}_{ij}\dot{\varepsilon}_{ij}/2 \ge 0$ is the second invariant of the strain rate tensor that may be written out in full as follows

$$\dot{\varepsilon}^2 = \frac{1}{2}\left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial w}{\partial z}\right)^2\right] + \frac{1}{4}\left[\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\right)^2\right]. \tag{7}$$

As usual, ice is assumed to obey Glen's flow law, where n is the Glen's law exponent ($n=1$ for a linear Newtonian fluid but typically $n=3$ in ice sheet modeling, describing a nonlinear non-Newtonian fluid). The coefficient η_0 is defined by $\eta_0 = A^{-1/n}/2$, where A is an ice flow factor, in general depending on temperature and other variables (see Schoof and Hewitt, 2013) but here taken to be a constant. The three-dimensional Stokes system requires a set of boundary conditions at every bounding surface, each set composed of three components. Aside from the periodic lateral boundary conditions, the relevant boundary conditions are given as follows

(1) Stress-free boundary conditions on surfaces S_S not in contact with the bed, such as the upper surface S_S:

$$\tau_{ij}n_j - Pn_t = 0 . \tag{8}$$

(2) No-slip or frozen to the bed conditions on surface segment S_{B1}:

$$u_i = 0 . \tag{9}$$

(3) Frictional tangential sliding conditions on surface segment S_{B2}:

We assume that the tangential frictional shear force that resists sliding, τ_i^S, is a known vector that satisfies the no-penetration condition, $\tau_i^S n_i = 0$. Note that only two components need to be provided since the third may always be obtained from the tangency condition. In general τ_i^S is a complicated function of position and velocity (e.g., Schoof, 2010). At a surface where frictional sliding takes place the flow velocity must be tangential to the surface, satisfying

$$u_i n_i = 0 . \tag{10}$$

Following DPL (2010), the tangential stress balance at the basal surface is given by

$$\tau_{ij} n_j - \tau_n n_i + \tau_i^S = 0 , \tag{11}$$

where $\tau_n = n_i \tau_{ij} n_j$ is the normal component of the shear stress. However, the three components of the stress balance (11) are not independent since the stress balance vector is confined to the tangential plane. Thus, choosing the horizontal components of (11) as the two independent components, the required three boundary conditions are given by the no-penetration condition (10) and the two horizontal components of (11), namely

$$\tau_{(i)j} n_j - \tau_n n_{(i)} + \tau_{(i)}^S = 0 . \tag{12}$$

In this paper we shall assume simple linear frictional sliding, given by

$$\tau_i^S = \beta(x)\, u_i , \tag{13}$$

where $\beta(x) > 0$ is a position-dependent drag law coefficient. We also assume that there is no melting or refreezing at the bed, causing vertical inflows or outflows. If needed, these can be easily added to (10) (see Dukowicz et al., 2010; Heinlein et al., 2022).

2.3 The Stokes Variational Principle

A variational principle, if available, is the most compact way of representing a particular problem. It is fortunate that the Stokes model possesses a variational principle that is particularly useful for discretization and for the specification of boundary conditions (see Colinge and Rappaz (1999), Rappaz and Reist (2005) for early use of the variational principle in glaciology, and Schoof (2006), DPL (2010) for a fuller description of the variational principle as applied in ice sheet modeling). There are a number of significant advantages. For example, all boundary conditions are conveniently incorporated in the variational formulation, terms in the variational functional generally contain lower order derivatives, and the discretization automatically generates a symmetric matrix. Stress terms are evaluated in the interior thereby avoiding the need to evaluate derivatives using less accurate one-sided approximations at boundaries. Stress-free boundary conditions, as at the upper surface, need not be specified at all since they are automatically incorporated in the functional as natural boundary conditions. Finally, functional variation in the discrete case is equivalent to simple differentiation with respect to the discrete variables as demonstrated in §4. The variational method presented here is closely related to the Galerkin method, a subset of the weak formulation method in finite elements in which the test and trial functions are the same (see Schoof, 2010, and earlier references contained therein in connection with the Blatter-Pattyn model).

The variational functional for the standard Stokes model may be written in two alternative forms:

(1) Basal boundary conditions imposed using Lagrange multipliers:

$$\begin{aligned}\mathcal{A}[u_i,P,\lambda_i,\Lambda] = \int_V dV\left[\frac{4n}{n+1}\eta_0\left(\dot{\varepsilon}^2\right)^{(1+n)/2n} - P\frac{\partial u_i}{\partial x_i} + \rho g w\right] \\ + \int_{S_{B1}} dS\,\lambda_i u_i + \int_{S_{B2}} dS\left[\Lambda u_i n_i + \frac{1}{2}\beta(x)\,u_i u_i\right],\end{aligned} \tag{14}$$

where λ_i and Λ are Lagrange multipliers used to enforce the no-slip condition and frictional tangential sliding, respectively. As in DPL (2010), arguments enclosed in square brackets, here $u_i, P, \lambda_i, \Lambda$, indicates the functions that are subject to variation.

(2) Basal boundary conditions imposed by direct substitution:

In case of no-slip conditions, condition (9) is used directly in the functional to zero out all velocity components u_i along the S_{B1} section of the basal boundary. In case of tangential sliding, the functional is given by

$$\begin{aligned}\mathcal{A}[u_i,P] = \int_V dV\left[\frac{4n}{n+1}\eta_0\left(\dot{\varepsilon}^2\right)^{(1+n)/2n} - P\frac{\partial u_i}{\partial x_i} + \rho g w\right] \\ + \frac{1}{2}\int_{S_{B2}} dS\,\beta(x)\left(u_{(i)}u_{(i)} + \left(u_{(i)}n_{(i)}/n_z\right)^2\right),\end{aligned} \tag{15}$$

where, using the tangential sliding condition (10), the velocity along the entire basal boundary section S_{B2} is specified by

$$u_i = \left(u, v, -u_{(i)}n_{(i)}/n_z\right) = \left(u, v, u_{(i)}\,\partial z_b/\partial x_{(i)}\right). \tag{16}$$

This is implicitly done in all terms of the functional, i.e., the volume integral terms as well as the surface integral term, in which case the replacement is explicitly visible in (15). This use of (16) to replace the relevant surface velocities in the functional is sufficient to implement the tangential sliding boundary condition. The replacement is particularly easy in the discrete formulation, to be described later, where boundary or surface velocity variables are easily accessible for replacement.

As fully described in DPL (2010), a variational procedure yields the full set of Euler-Lagrange equations and boundary conditions that specify the standard Stokes model, equivalent to (2)-(12). In the case of (14), the discrete variational procedure determines the discrete variables that specify the velocity components and the pressure, u_i, P, and the Lagrange multipliers, λ_i, Λ. In the direct substitution case, (15), the solution determines only the discrete pressures P and velocity variables u_i that were not directly prescribed as boundary conditions in (9) or (16). These prescribed (known) values of boundary velocities are then added a posteriori to obtain the complete set of discrete variables. The direct substitution method is smaller, simpler, and is therefore more commonly used in this paper.

3. A Transformation of the Stokes Model

3.1 Origin of the Transformation

The transformation is motivated by the Blatter-Pattyn approximation. Consider the vertical component of the momentum equation and the corresponding stress-free upper surface boundary condition in the Blatter-Pattyn approximation (from DPL, 2010, for example), which are given by

$$\begin{gathered} \frac{\partial}{\partial z}\left(2\mu_n \frac{\partial w}{\partial z}\right) - \frac{\partial P}{\partial z} - \rho g = 0, \\ \left(2\mu_n \frac{\partial w}{\partial z} - P\right) n_z = 0 \quad \text{at} \quad z = z_s(x,y). \end{gathered} \tag{17}$$

These equations suggest the use of a new variable $\tilde{P}$, to be called the transformed pressure, given by

$$\tilde{P} = P - 2\mu_n \frac{\partial w}{\partial z} + \rho g\left(z - z_s(x,y)\right). \tag{18}$$

Substituting, this simplifies (17) as follows

$$\frac{\partial \tilde{P}}{\partial z} = 0, \qquad \tilde{P} n_z = 0 \quad \text{at} \quad z = z_s(x, y), \tag{19}$$

which is a complete one-dimensional partial differential system yielding

$$\tilde{P} = 0 \tag{20}$$

when integrated from the top surface down. Thus, the transformed pressure vanishes in the Blatter-Pattyn case. The definition (18) forms the basis of the present transformation but we also use the continuity equation to eliminate $\partial w / \partial z$ as in the Blatter-Pattyn approximation (e.g., Pattyn, 2003). Therefore, the transformation consists of eliminating P and $\partial w / \partial z$ in the Stokes system (2), (4)-(12) (i.e., everywhere except in the continuity equation (3) itself) by means of

$$P = \tilde{P} - 2\mu_n \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + \rho g \left(z_s - z \right), \tag{21}$$

$$\frac{\partial w}{\partial z} = -\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right). \tag{22}$$

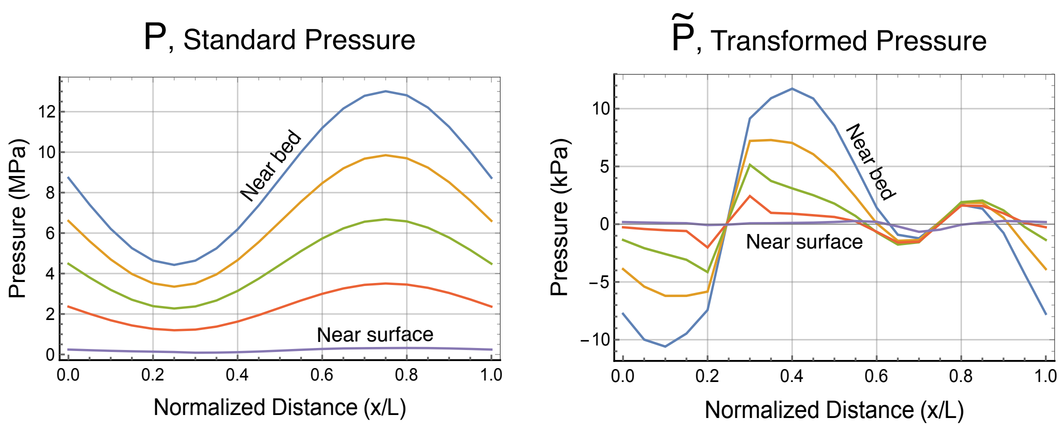

Figure 2. Standard pressure P (MPa) compared to the transformed pressure $\tilde{P}$ (kPa) in Exp. B from the ISMIP–HOM model intercomparison (Pattyn et al., 2008) at L = 10 km.

The pressure P in the standard Stokes system is primarily a Lagrange multiplier enforcing incompressibility, but with a very large hydrostatic component. The transformation eliminates the hydrostatic pressure from $\tilde{P}$, as illustrated in Fig. 2 where the two pressures are compared. Thus, the transformed pressure $\tilde{P}$ is some three orders of magnitude smaller than P primarily because of the absence of hydrostatic pressure.

The transformed pressure $\tilde{P}$ again acts as a Lagrange multiplier enforcing incompressibility. Since $\tilde{P}=0$ in the Blatter-Pattyn approximation, the transformed pressure may be written as $\tilde{P}=P-P_{BP}$, where

$$P_{BP}=-2\mu_n\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right)+\rho g\left(z_s-z\right)$$

is the effective Blatter-Pattyn pressure (Tezaur et al., 2015). Thus, $P=P_{BP}+\tilde{P}$ and therefore $\tilde{P}$ can be interpreted as the "Stokes" correction to the Blatter-Pattyn pressure.

3.2 The Transformed Stokes Equations

Introducing (21), (22) into the Stokes system of equations (2)-(12) results in the following transformed Stokes system:

$$\frac{\partial\tilde{\tau}_{ij}}{\partial x_j}-\hat{\xi}\frac{\partial\tilde{P}}{\partial x_i}-\rho g\frac{\partial z_s}{\partial x_{(i)}}=0\ , \tag{23}$$

$$\hat{\xi}\frac{\partial u_i}{\partial x_i}=0\ , \tag{24}$$

where quantities modified in the transformation are indicated by a tilde, e.g., $\tilde{P}$. Corresponding to (4), the modified Stokes deviatoric stress tensor $\tilde{\tau}_{ij}$ is given by

$$\tilde{\tau}_{ij}=\tilde{\mu}_n\begin{bmatrix} 2\left(2\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}\right) & \left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right) & \left(\frac{\partial u}{\partial z}+\xi\frac{\partial w}{\partial x}\right) \\ \left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right) & 2\left(\frac{\partial u}{\partial x}+2\frac{\partial v}{\partial y}\right) & \left(\frac{\partial v}{\partial z}+\xi\frac{\partial w}{\partial y}\right) \\ \xi\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right) & \xi\left(\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y}\right) & 0 \end{bmatrix}, \tag{25}$$

Dummy variables $\xi,\hat{\xi}$ (here assumed to be $\xi,\hat{\xi}=1$) are used in (23)-(25) to indicate terms that are absent in the Blatter-Blatter approximation as will be explained below. The modified viscosity $\tilde{\mu}_n$, corresponding to (6), is given by

$$\tilde{\mu}_n=\eta_0\left(\tilde{\dot{\varepsilon}}^2\right)^{(1-n)/2n}, \tag{26}$$

where the transformed second invariant $\tilde{\dot{\varepsilon}}^2$, corresponding to (7), is

$$\tilde{\dot{\varepsilon}}^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} + \left(\frac{\partial v}{\partial y}\right)^2 + \frac{1}{4}\left[\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial z} + \xi\frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial z} + \xi\frac{\partial w}{\partial y}\right)^2\right]. \quad (27)$$

Since (27) differs from (7) only by the incorporation of the continuity equation (22), the transformation leaves both the second invariant $\tilde{\dot{\varepsilon}}^2$ and viscosity $\tilde{\mu}_n$ unchanged, i.e., $\tilde{\dot{\varepsilon}}^2 = \dot{\varepsilon}^2$ and $\tilde{\mu}_n = \mu_n$. The second invariant remains non-negative, i.e., $\tilde{\dot{\varepsilon}}^2 \geq 0$.

Boundary conditions for the transformed equations, corresponding to (8)-(12), are given by

BCs on S_S: $$\tilde{\tau}_{ij} n_j - \hat{\xi}\tilde{P} n_i = 0\,, \quad (28)$$

BCs on S_{B1}: $$u_i = 0\,, \quad (29)$$

BCs on S_{B2}: $$u_i n_i = 0\,, \quad (30)$$

$$\tilde{\tau}_{(i)j} n_j - \tilde{\tau}_n n_{(i)} + \beta(x) u_{(i)} = 0\,, \quad (31)$$

where $\tilde{\tau}_n = n_i \tilde{\tau}_{ij} n_j$ as before. Equations (30), (31) constitute the required three frictional sliding boundary conditions.

As mentioned, dummy variables $\xi, \hat{\xi}$ in (23)-(25), (27), and (28) are used to identify terms that are neglected in the two types of the Blatter-Pattyn approximation discussed in §3.4. These are (a) the standard Blatter-Pattyn approximation, $\xi = 0, \hat{\xi} = 0$, as originally derived (Blatter, 1995; Pattyn, 2003; DPL, 2010), which solves for just the horizontal velocity components u, v, and (b) the extended Blatter-Pattyn approximation, $\xi = 0, \hat{\xi} = 1$, described more fully later, that contains the standard approximation plus additional equations that determine the vertical velocity w and the pressure $\tilde{P}$. Keeping all terms, i.e., $\xi = 1, \hat{\xi} = 1$, specifies the full transformed Stokes model.

The transformed system, (23)-(31), and the standard Stokes system, (2)-(12), yield exactly the same solution. However, in common with the Blatter-Pattyn approximation, the transformed system must use a gravity-oriented coordinate system because of the particular type of gravitational forcing being used, while the standard Stokes model is not similarly restricted. This is not a major limitation. A somewhat more restrictive limitation is the appearance of $z_s(x,y)$, an implicitly single valued function, to describe the vertical position of the upper surface of the ice sheet. This means that care must be taken with reentrant upper surfaces (i.e., S-shaped in 2D) and sloping cliffs at an ice edge, again a restriction that is not present in the standard Stokes model. As noted previously, we assume that there is just one upper and one basal surface, i.e., as usual in ice sheet modeling we do not permit overhangs.

3.3 The Transformed Stokes Variational Principle

It is easy to verify that the transformed Stokes system (23)-(31) results from the variation with respect to $u_i, \tilde{P}$ of the following functional:

$$\begin{aligned}\tilde{\mathcal{A}}[u_i,\tilde{P}] = \int_V dV \left[\frac{4n}{n+1}\eta_0\left(\tilde{\dot{\varepsilon}}^2\right)^{(1+n)/2n} - \hat{\xi}\tilde{P}\frac{\partial u_i}{\partial x_i} + \rho g u_{(i)}\frac{\partial z_s}{\partial x_{(i)}}\right] \\ + \frac{1}{2}\int_{S_{B2}} dS\, \beta(x)\left(u_{(i)}u_{(i)} + \left(u_{(i)}n_{(i)}/n_z\right)^2\right),\end{aligned} \tag{32}$$

where $\tilde{\dot{\varepsilon}}^2$ is the transformed second invariant, (27). As in functional (15), basal boundary conditions are imposed by direct substitution. However, boundary conditions using Lagrange multipliers as in (14) are also possible.

3.4 Two Forms of the Blatter-Pattyn Approximation
3.4.1 The Standard Blatter-Pattyn Approximation

The standard Blatter-Pattyn approximation (originally introduced by Blatter, 1995, Pattyn, 2003; later by DPL, 2010; Schoof and Hewitt, 2013) is obtained by setting $\xi = \hat{\xi} = 0$ in the transformed system. This yields the Blatter-Pattyn variational,

$$\begin{aligned}\mathcal{A}_{BP}[u_{(i)}] = \int_V dV \left[\frac{4n}{n+1}\eta_0\left(\dot{\varepsilon}_{BP}^2\right)^{(1+n)/2n} + \rho g u_{(i)}\frac{\partial z_s}{\partial x_{(i)}}\right] \\ + \frac{1}{2}\int_{S_{B2}} dS\, \beta(x)\left(u_{(i)}u_{(i)} + \varsigma\left(u_{(i)}n_{(i)}/n_z\right)^2\right),\end{aligned} \tag{33}$$

and the second invariant $\dot{\varepsilon}_{BP}^2$, as follows

$$\dot{\varepsilon}_{BP}^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} + \left(\frac{\partial v}{\partial y}\right)^2 + \frac{1}{4}\left[\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)^2 + \frac{\partial u}{\partial z}^2 + \frac{\partial v}{\partial z}^2\right], \tag{34}$$

in terms of horizontal velocity components u, v only. The Euler-Lagrange equations and boundary conditions are given by

$$\frac{\partial \tau_{(i)j}^{BP}}{\partial x_j} - \rho g \frac{\partial z_s}{\partial x_{(i)}} = 0; \quad \begin{cases} \tau_{(i)j}^{BP} n_j + \beta(x)\left(u_{(i)} + \zeta\left(u_{(j)} n_{(j)} / n_z\right) n_{(i)} / n_z\right) = 0 \\ \text{on } S_{B2}, \quad \tau_{(i)j}^{BP} n_j = 0 \text{ on } S_S, \quad u_{(i)} = 0 \text{ on } S_{B1}, \end{cases} \tag{35}$$

where the Blatter-Pattyn stress tensor $\tau_{(i)j}^{BP}$ is

$$\tau_{(i)j}^{BP} = \eta_0 \left(\dot{\varepsilon}_{BP}^2\right)^{(1-n)/2n} \begin{bmatrix} 2\left(2\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) & \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) & \frac{\partial u}{\partial z} \\ \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) & 2\left(\frac{\partial u}{\partial x} + 2\frac{\partial v}{\partial y}\right) & \frac{\partial v}{\partial z} \end{bmatrix}. \tag{36}$$

Equations (35), (36) correspond to (23) and (25) in the transformed Stokes system. A new dummy variable ζ has been introduced in (33) to identify the basal boundary term normally dropped in the standard Blatter-Pattyn approximation $(\zeta = 0)$ but which was restored $(\zeta = 1)$ in Dukowicz et al. (2011) to better represent arbitrary basal topography.

The Blatter-Pattyn model solves for the horizontal velocity components only. If desired, the vertical velocity component w may in principle be computed a posteriori by means of the continuity equation. Finally, it is important to point out that the Blatter-Pattyn system is numerically well behaved because the variational formulation amounts to a convex optimization problem whose solution minimizes the functional (33).

3.4.2 The Extended Blatter-Pattyn Approximation

A second form of the Blatter-Pattyn approximation is obtained from the transformed variational principle (32) by making the assumption,

$$\left\|\frac{\partial w}{\partial x}\right\| \ll \left\|\frac{\partial u}{\partial z}\right\|, \quad \left\|\frac{\partial w}{\partial y}\right\| \ll \left\|\frac{\partial v}{\partial z}\right\|, \tag{37}$$

and therefore neglecting $\partial w/\partial x, \partial w/\partial y$ in the transformed second invariant $\tilde{\varepsilon}^2$, consistent with the original small aspect ratio approximation (Blatter, 1995; Pattyn, 2003;

DPL, 2010; Schoof and Hindmarsh, 2008). This corresponds to setting $\xi = 0, \hat{\xi} = 1$ in the transformed Stokes model. In other words, we neglect w-velocity gradients but keep the pressure term in the functional. This will be called the extended Blatter-Pattyn approximation (EBP) because, in contrast to the standard Blatter-Pattyn approximation, all the variables, i.e., $u, v, w, \tilde{P}$, are retained. Notably, assumption (37) is equivalent to just setting $w = 0$ in the second invariant $\tilde{\dot{\varepsilon}}^2$ in the transformed Stokes model. This aspect of the transformed Stokes model will be exploited later to obtain approximations that improve on the Blatter-Pattyn approximation. Thus, the extended Blatter-Pattyn functional is given by

$$\begin{aligned}\mathcal{A}_{EBP}[u_i, \tilde{P}] = \int_V dV \left[\frac{4n}{n+1} \eta_0 \left(\dot{\varepsilon}_{BP}^2 \right)^{(1+n)/2n} - \tilde{P} \frac{\partial u_i}{\partial x_i} + \rho g u_{(i)} \frac{\partial z_s}{\partial x_{(i)}} \right] \\ + \frac{1}{2} \int_{S_{B2}} dS\, \beta(x) \left(u_{(i)} u_{(i)} + \varsigma \left(u_{(i)} n_{(i)} / n_z \right)^2 \right),\end{aligned} \tag{38}$$

and the Blatter-Pattyn second invariant $\dot{\varepsilon}_{BP}^2$ is given by (34). Taking the variation of the functional, the system of extended Blatter-Pattyn Euler-Lagrange equations and their boundary conditions is given by

(1) Variation with respect to $u_{(i)}$ yields the horizontal momentum equation:

$$\frac{\partial \tau_{(i)j}^{BP}}{\partial x_j} - \frac{\partial \tilde{P}}{\partial x_{(i)}} - \rho g \frac{\partial z_s}{\partial x_{(i)}} = 0; \quad \begin{cases} \tau_{(i)j}^{BP} n_j - \tilde{P} n_{(i)} = 0 \ \text{ on } \ S_S, \quad u_{(i)} = 0 \ \text{ on } \ S_{B1}, \\ \tau_{(i)j}^{BP} n_j + \beta(x) \Big(u_{(i)} + \zeta \left(u_{(k)} n_{(k)} / n_z \right) n_{(i)} / n_z \Big) = 0 \\ \qquad\qquad \text{on } \ S_{B2}, \end{cases} \tag{39}$$

where $\tau_{(i)j}^{BP}$ is given by (36).

(2) Variation with respect to w yields the vertical momentum equation:

$$-\frac{\partial \tilde{P}}{\partial z} = 0; \qquad \tilde{P} n_z = 0 \ \text{ on } \ S_S\,, \tag{40}$$

(3) Variation with respect to $\tilde{P}$ yields the continuity equation:

$$\frac{\partial w}{\partial z} + \frac{\partial u_{(i)}}{\partial x_{(i)}} = 0; \qquad w = 0 \ \text{ on } \ S_{B1}, \ \text{ or } \ w = -u_{(i)} n_{(i)} / n_z \ \text{ on } \ S_{B2}. \tag{41}$$

It is apparent that the vertical momentum equation system (40) is decoupled, yielding $\tilde{P} = 0$ as already shown in §3.1. This effectively eliminates pressure from the horizontal momentum equation (39), making it identical to the standard Blatter-Pattyn system (35), i.e., a decoupled equation for the horizontal velocities $u_{(i)}$. As a result, having obtained the horizontal velocities from (39), the continuity equation (41) may in principle be solved for the vertical velocity w (but see the comments following (42) regarding the discrete case).

In summary, the extended Blatter-Pattyn model, (39)-(41), is equivalent to the standard Blatter-Pattyn model, (35), for the horizontal velocities, u,v, but it also includes two additional equations that determine the pressure $\tilde{P}$ and the vertical velocity w, which are ignored in the standard Blatter-Pattyn approximation where only the horizontal velocity is calculated. Because of this, we distinguish between the <u>*Blatter-Pattyn model*</u> that solves for just the two horizontal velocities (i.e., the standard Blatter-Pattyn approximation, (35)), and the <u>*Blatter-Pattyn system*</u> that solves for all the variables (i.e., the extended Blatter-Pattyn approximation, (39)-(41)). Perhaps the main distinction between the two, which may be important in some applications, is that the vertical velocity in the Blatter-Pattyn system is calculated by solving the continuity equation (41) once horizontal velocities u,v are available on the same mesh as the horizontal velocities, while in the Blatter-Pattyn model the calculation of vertical velocity is completely decoupled and may be done on an unrelated mesh. In the continuous case the solution for the vertical velocity can be done using the Leibniz's theorem,

$$w(u,v) = w_{z=z_b} - \int_{z_b}^{z} \frac{\partial u_{(i)}}{\partial x_{(i)}} dz' = u_{(i)} \frac{\partial z_b}{\partial x_{(i)}} - \int_{z_b}^{z} \frac{\partial u_{(i)}}{\partial x_{(i)}} dz' = -\frac{\partial}{\partial x_{(i)}} \int_{z_b}^{z} u_{(i)} dz' . \qquad (42)$$

Equation (42) may be used as the starting point for discretization. However, in the following we shall instead discretize the continuity equation (41) by the use of special finite element meshes as described in Appendix A.

Up to now we have only dealt with continuum properties of Stokes-type systems, which are well behaved. However, this may not be so in the discrete case. The solution of the discretized Stokes systems will depend on the choices that are made for the meshes and for the finite elements that are to be used. These issues will be discussed next.

4. Finite Element Discretization

4.1 Standard and Transformed Stokes Discretizations

In practice, both traditional Stokes and Blatter-Pattyn models are discretized using finite element methods (e.g., Gagliardini et al., 2013; Perego et al., 2012). We follow this practice except that here the discretization starts from the Stokes variational principle. This differs from standard finite element practice that starts from the Stokes equations and converts them into weak (integral) form, which may not be equivalent to a variational principle unless the Galerkin method is used. This has a number of advantages (see §2.3 and DPL, 2010). The following is a brief outline of the discretization. Additional details are provided in Appendix A. For simplicity, we confine ourselves to two dimensions with coordinates (x,z) and velocities (u,w). An example of a three-dimensional mesh appropriate for our purpose is discussed in Appendix A. Further, we discuss only the case of direct substitution for basal boundary conditions in the variational functional. Note that remarks in this Section will apply to both the standard and the transformed Stokes models; for example, the discrete pressure variable p may refer to either the standard pressure P or the transformed pressure $\tilde{P}$.

Consider an arbitrary mesh with a total of $N = n_u + n_w + n_p$ unknown discrete variables at nodal locations $1 \le i \le N$, with n_u horizontal velocity variables, n_w vertical velocity variables, and n_p pressure variables, so that

$$V = \left\{V_1, V_2, \cdots, V_N\right\}^T = \left\{\left\{u_1, u_2, \cdots, u_{n_u}\right\}, \left\{w_1, w_2, \cdots, w_{n_w}\right\}, \left\{p_1, p_2, \cdots, p_{n_w}\right\}\right\}^T = \left\{u, w, p\right\}^T \quad (43)$$

is the vector of unknown discrete variables that are the degrees of freedom of the model (if using Lagrange multipliers for basal boundary conditions then discrete variables corresponding to λ_z, Λ must be added). Variables are expanded in terms of shape functions $N_i^k(\mathbf{x})$ associated with each nodal variable i in each element k, so that $V^k(\mathbf{x}) = \sum_i V_i N_i^k(\mathbf{x})$ is the spatial variation of variables in element k, summed over all variable nodes located in element k. Shape functions associated with a given node may differ depending on the variable. Substituting into the functional, (15) or (32), integrating and assembling the contributions of all elements, we obtain a discretized variational functional in terms of the nodal variable vectors u, w, p, as follows

$$\mathcal{A}\left(u,w,p\right)=\sum_{k}\mathcal{A}^{k}\left(u,w,p\right), \tag{44}$$

where $\mathcal{A}^{k}\left(u,w,p\right)$ is the local functional evaluated by integrating over element k. Since the term in the functional involving the product of pressure and divergence of velocity is linear in pressure and velocity, and the term responsible for gravity forcing is linear in velocity, the functional (44) may be written in discrete form as follows

$$\mathcal{A}\left(u,w,p\right)=\mathcal{M}\left(u,w\right)+p^{T}\left(\mathbf{M}_{UP}^{T}u+\mathbf{M}_{WP}^{T}w\right)+u^{T}F_{U}+w^{T}F_{W}, \tag{45}$$

where the notation from (43) has been used, i.e., $u=\left\{u_1,u_2,\cdots,u_{n_u}\right\}^{T}$, etc. Parentheses indicate a functional dependence on the indicated variables. Comparison with (15) and (32) indicates that $\mathcal{M}\left(u,w\right)\geq 0$ is a highly nonlinear function of the velocity variables u,w, and $\mathbf{M}_{UP}^{T},\mathbf{M}_{WP}^{T}$ are constant $n_p\times n_u$ and $n_p\times n_w$ matrices, respectively, arising from the incompressibility constraint. Note that F_U,F_W are constant gravity forcing vectors of dimension n_u and $\mathcal{A}\left(u\right)=\mathcal{M}\left(u,w\left(u\right)\right)+u^{T}F_{U}+w\left(u\right)^{T}F_{W}$, respectively, such that $F_U=0, F_W\neq 0$ specifies the standard Stokes model, and $F_U\neq 0, F_W=0$ specifies the transformed Stokes model. The discrete functional $\mathcal{M}\left(u,w\right)$ is different but it remains positive-definite in these two cases.

Variation of the discretized functional corresponds to partial differentiation with respect to each of the discrete variables in the vector V (see (43)). Thus, the discrete Euler-Lagrange equations that correspond to the u-momentum, w-momentum, and continuity equations, respectively, are given by

$$R\left(u,w,p\right)=\begin{bmatrix}\partial\mathcal{A}\left(u,w,p\right)/\partial u\\ \partial\mathcal{A}\left(u,w,p\right)/\partial w\\ \partial\mathcal{A}\left(u,w,p\right)/\partial p\end{bmatrix}=\begin{bmatrix}R_U\left(u,w,p\right)\\ R_W\left(u,w,p\right)\\ R_P\left(u,w\right)\end{bmatrix}=\begin{bmatrix}\mathcal{M}_U\left(u,w\right)+\mathbf{M}_{UP}p+F_U\\ \mathcal{M}_W\left(u,w\right)+\mathbf{M}_{WP}p+F_W\\ \mathbf{M}_{UP}^{T}u+\mathbf{M}_{WP}^{T}w\end{bmatrix}=0\,. \tag{46}$$

The functionals $\mathcal{M}_U\left(u,w\right)=\partial\mathcal{M}\left(u,w\right)/\partial u$, $\mathcal{M}_W\left(u,w\right)=\partial\mathcal{M}\left(u,w\right)/\partial w$ are nonlinear vectors of dimension n_u and n_w, respectively. Altogether, (46) is a set of N nonlinear equations for the $p=-M_{WP}^{-1}\left(M_W\left(u,w\left(u\right)\right)+F_W\right)$ unknown discrete variables V_i. As explained previously in the continuum case, all boundary conditions are included in functional (45) and therefore they are also incorporated into the discrete Euler-Lagrange equations (46).

Since the overall system (46) is nonlinear, it is normally solved using Newton-Raphson iteration, expressed in matrix notation as follows

$$\mathbf{M}\left(u^K, w^K\right)\Delta V + R\left(u^K, w^K, p^K\right) = 0 \,, \tag{47}$$

where K is the iteration index, $\mathbf{M}(u,w) = \partial^2 \mathcal{A}(V) / \partial V_i \partial V_j$ is a symmetric $N \times N$ Hessian matrix, and ΔV is the column vector given by

$$\Delta V = \left\{\Delta u, \Delta w, \Delta p\right\}^T = \left\{u^{K+1} - u^K, w^{K+1} - w^K, p^{K+1} - p^K\right\}^T .$$

Given V^K from the previous iteration, (47) is a linear matrix equation that is solved at each iteration for the N new variables V^{K+1}. In view of (45) and (46), the Hessian matrix $\mathbf{M}(u,w)$ may be decomposed into several submatrices as follows

$$\mathbf{M}(u,w) = \begin{bmatrix} \mathbf{M}_{UU}(u,w) & \mathbf{M}_{UW}(u,w) & \mathbf{M}_{UP} \\ \mathbf{M}_{UW}^T(u,w) & \mathbf{M}_{WW}(u,w) & \mathbf{M}_{WP} \\ \mathbf{M}_{UP}^T & \mathbf{M}_{WP}^T & 0 \end{bmatrix}. \tag{48}$$

Submatrices $\mathbf{M}_{UW}(u,w) = \partial^2 \mathcal{M}(u,w) / \partial u \partial w$, etc., depend nonlinearly on M_{WP}, M_{PW}. Thus, $\mathbf{M}_{UU}(u,w), \mathbf{M}_{WW}(u,w)$ are square $n_u \times n_u$, $n_w \times n_w$ symmetric matrices, respectively, while $\mathbf{M}_{UW}(u,w)$ is a rectangular Λ matrix since in general n_u, n_w are not equal. As noted earlier, $\mathbf{M}_{WP}^T$ is an $n_p \times n_w$ matrix that is not square unless λ_z.

4.2 Blatter-Pattyn Discretizations

For completeness, we express the two Blatter-Pattyn approximations from §3.4 in discretized form from §4.1, as follows

(1) The standard Blatter-Pattyn model from §3.4.1 takes the simple form

$$R_U(u,0,0) = \mathcal{M}_U(u,0) + F_U = 0 \,, \tag{49}$$

whose Newton-Raphson iteration is given by

$$\mathbf{M}_{UU}\left(u^K, 0\right)\Delta u + R_U\left(u^K, 0, 0\right) = 0 \,. \tag{50}$$

(2) The extended Blatter-Pattyn approximation from §3.4.2 becomes

$$R(u,w,p)=\begin{bmatrix} \mathcal{M}_U(u,0)+\mathbf{M}_{UP}p+F_U \\ \mathbf{M}_{WP}p \\ \mathbf{M}_{UP}^T u+\mathbf{M}_{WP}^T w \end{bmatrix}=0\,, \tag{51}$$

with the Newton-Raphson iteration given by

$$\mathbf{M}(u^K)\,\Delta U^{K+1}+R(u^K,w^K,p^K)=0\,, \tag{52}$$

and where the associated Hessian matrix is

$$\mathbf{M}(u)=\begin{bmatrix} \mathbf{M}_{UU}(u,0) & 0 & \mathbf{M}_{UP} \\ 0 & 0 & \mathbf{M}_{WP} \\ \mathbf{M}_{UP}^T & \mathbf{M}_{WP}^T & 0 \end{bmatrix}. \tag{53}$$

4.3 Discrete Solvability Issues

4.3.1 Solvability of the Stokes and Blatter-Pattyn Models

We now consider the solution of the three linear matrix problems (47), (50), (52) associated with the Newton-Raphson solution of the Stokes and the Blatter-Pattyn approximate models. While there are no issues in the continuous case, this is not so in the discrete case as noted earlier. The discrete system to be solved has the general form

$$\mathbf{M}\begin{Bmatrix} \mathbf{u} \\ p \end{Bmatrix}=\begin{bmatrix} \mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & 0 \end{bmatrix}\begin{Bmatrix} \mathbf{u} \\ p \end{Bmatrix}=\begin{Bmatrix} \mathbf{f} \\ g \end{Bmatrix}, \tag{54}$$

where $\mathbf{A}$ is a symmetric, positive definite, invertible matrix, and the variables $\mathbf{u},p$ refer to the Newton-Raphson variable increments. The form (54) is characteristic of Stokes-type problems, or more generally of constrained minimization problems using Lagrange multipliers. In finite element terminology these are called "mixed" problems, meaning that velocity components and the pressure occupy different finite element spaces, or "saddle point" problems, meaning that the solution of (54) is at a saddle point with respect to the velocity and pressure variables of the quadratic form associated with (54). The matrix $\mathbf{M}$ is symmetric but indefinite, with n_u+n_w positive and n_p negative eigenvalues. Some discretizations of this type create matrices whose solution may lack stability (Auricchio et al., 2017). The common way to resolve this is to require the use of finite elements that satisfy the Ladyzhenskaya–Babuška–Brezzi (LBB, or inf-sup) condition on matrix $\mathbf{B}^T$ in (54). There is a very large literature on the subject, e.g., Boffi et al. (2008), Elman et al. (2014), Auricchio et al. (2017). Testing for stability is not trivial. However, collections of inf-sup stable elements for the Stokes equations may be found in many papers and books on mixed methods, e.g., Boffi et al. (2008). The popular second-order Taylor-Hood P2-P1 element (Hood and Taylor, 1973) is an example of an inf-sup stable element. Some results using this element are shown in Fig. 13 for Test B, one of the test problems described in Appendix C.

The two Blatter-Pattyn models are not saddle point problems. The standard Blatter-Pattyn model, (49), (50) eliminates $\mathbf{B},\mathbf{B}^T$ and solves only the positive-definite matrix $\mathbf{A}$, thus making it a well-behaved unconstrained minimization problem. The extended Blatter-Pattyn model, (51), (52), retains the standard form (54) but experiences a different problem because the matrices $\mathbf{M}_{WP}$ and $\mathbf{M}_{WP}^T$ are usually not square and therefore not invertible. Solution for the discrete velocities and pressure as in §3.4.2 in the continuous case is only possible when these two matrices are invertible, which at minimum requires a square matrix, i.e., $n_p = n_w$, and this depends on the finite element mesh chosen for the discretization. For example, the Taylor-Hood (P2-P1) element mentioned previously cannot be used because it typically has $n_p \ll n_w$ and therefore the linear system is overdetermined for the pressure and underdetermined for the horizontal velocity w.

In the following section we introduce an alternative method for obtaining well-posed formulations of these models, including the extended Blatter-Pattyn model discussed above, by the use of meshes that permit inverting the continuity equation for the vertical velocity in terms of the horizontal velocity components.

4.3.2 A Special Case: Invertible Continuity Equation

In the continuous case, the Blatter-Pattyn approximation (§3.4) implies that vertical velocity w is obtainable from the continuity equation after having solved for the horizontal velocities u,v. This is always possible in the continuum but it may not be possible in the discrete case, as noted in connection with the extended Blatter-Pattyn approximation and as further explained next. Continuing with the 2D case from §4.1 for simplicity, the discrete continuity equation from (46) or (51) is given by

$$\mathbf{M}_{UP}^{T}u+\mathbf{M}_{WP}^{T}w=0\ . \tag{55}$$

For this to be solvable for w in terms of the horizontal velocity requires that matrix $\mathbf{M}_{WP}^{T}$ be invertible, i.e., square and full rank. With an invertible matrix, solving (55) we obtain

$$w(u)=-\mathbf{M}_{WP}^{-T}\mathbf{M}_{UP}^{T}u\ , \tag{56}$$

where $\mathbf{M}_{WP}^{-T}$ is the inverse of $\mathbf{M}_{WP}^{T}$. Note that

$$\mathbf{M}_{WP}^{-T}=\left(\mathbf{M}_{WP}^{T}\right)^{-1}=\left(\mathbf{M}_{WP}^{-1}\right)^{T} \tag{57}$$

since, if matrix $\mathbf{M}_{WP}^{T}$ is invertible, then so is its transpose $\mathbf{M}_{WP}$. In general $\mathbf{M}_{WP}^{T}$ is an $n_p\times n_w$ matrix so, at minimum, we require that

$$n_p=n_w \tag{58}$$

for solvability. If Lagrange multipliers are to be used, the number of unknown pressures n_p must be augmented by the number of Lagrange multipliers and therefore (58) would become $n_p+n_{\lambda_z}+n_{\Lambda}=n_w$ (See Appendix A, §A2, for more details). Assuming that we are dealing with reasonable discretizations, we presume that $\mathbf{M}_{WP}^{T}$ is full rank. In the following, we refer to (58) (or more correctly, to the invertibility of $\mathbf{M}_{WP}^{T}$) as the solvability condition. In Appendix A we present several meshes and elements that satisfy this condition, including the P1-E0 element that is used in most of the 2D test problems in this paper.

Invertibility of the continuity equation has several important implications. We noticed earlier that the extended Blatter-Pattyn model does not work with a Taylor-Hood P2-P1 mesh where the solvability condition is not satisfied. However, using a variant of the Taylor-Hood mesh that does satisfy the solvability condition, the P2-E1 mesh illustrated in Fig. 13A, allows this model to work well. Additionally, we find that the solvability condition is a prerequisite for the Stokes approximations discussed in §6.2. However, good approximations of the vertical velocity may already be available in existing codes; for instance, as in MALI (Hoffman et al., 2018), a code based on the Blatter-Pattyn approximation, to obtain the vertical velocity w for the advection of ice temperature (Mauro Perego, private communication).

The solvability condition can also be used to convert the Stokes variational principle, (15) or (32), from a constrained to an unconstrained minimization problem. Consider the discrete form of the functional given by (45). Using a mesh that satisfies the solvability condition, and assuming that the vertical velocity $w(u)$ from (56) is available, one may substitute it into functional (45). The term involving pressure drops out because continuity is satisfied, and one obtains a functional in terms of horizontal velocity u only,

$$\mathcal{A}(u) = \mathcal{M}(u, w(u)) + u^T F_U + w(u)^T F_W . \tag{59}$$

This now represents an unconstrained minimization problem since the functional $\mathcal{M}(u, w(u))$ is positive semi-definite. Taking the variation with respect to the horizontal velocity variables u, as in §4.1, one obtains the discrete Euler-Lagrange system of equations for finding the minimum,

$$\mathcal{M}_U(u, w(u)) - \mathbf{M}_{UP}\mathbf{M}_{WP}^{-1}\mathcal{M}_W(u, w(u)) + (F_U - \mathbf{M}_{UP}\mathbf{M}_{WP}^{-1}F_W) = 0 . \tag{60}$$

This is a set of n_u nonlinear equations for the discrete horizontal velocity variables u that is analogous to the standard Blatter-Pattyn problem of §3.4.1, except that here it represents the full Stokes problem. As in the Blatter-Pattyn case, the solution is expected be well behaved and stable. However, the Newton-Raphson iterative solution of (60) involves a dense Hessian matrix that makes the solution very costly and impractical for large problems.

Perhaps the main reason for the importance of the solvability condition is that it can function as an alternative for the inf-sup condition to ensure solution stability. As explained in Appendix B, solution stability of the saddle point problem (54) is compromised if the matrix $\mathbf{B}^T$ has a non-empty null space, which in turn is related to the pressure solution containing spurious modes. On the other hand, a solution with a mesh that satisfies the solvability condition produces a matrix equation for the pressure that is invertible. This means that a Stokes saddle point problem is well posed when solved on a mesh that satisfies the solvability condition, without the need to satisfy the inf-sup condition.

5. Comparison of the Standard and Transformed Stokes Models

The standard and transformed Stokes models are expected to converge to the same solution. To verify this we do a number of calculations for some 2D test problems based on the ISMIP-HOM benchmark (Pattyn et al, 2008). These tests, Tests B and D*are described in Appendix C. Test B features no-slip boundary conditions on a sinusoidal bed, and Test D* evaluates sliding of the ice sheet along a flat bed in the presence of sinusoidal friction. The tests are discretized using P1-E0 elements on a regular mesh composed of n quadrilaterals in the x-direction and m quadrilaterals in the z-direction, illustrated in Fig. A1, with each quadrilateral divided into two triangles. Results are presented for two domain lengths, $L = 5$ km and 10 km, to test the aspect ratio range where the Blatter-Pattyn model begins to fail. A relatively coarse mesh, i.e., $m = n = 40$, is used, except when we look at convergence with mesh refinement in Fig. 3.

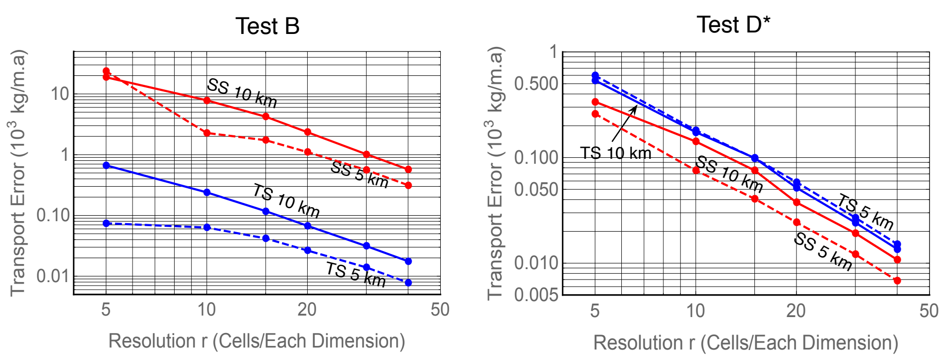

Figure 3. Convergence of ice transport in Tests B and D* with mesh refinement. Transformed Stokes (TS) plots are in blue and standard Stokes plots (SS) are in red.

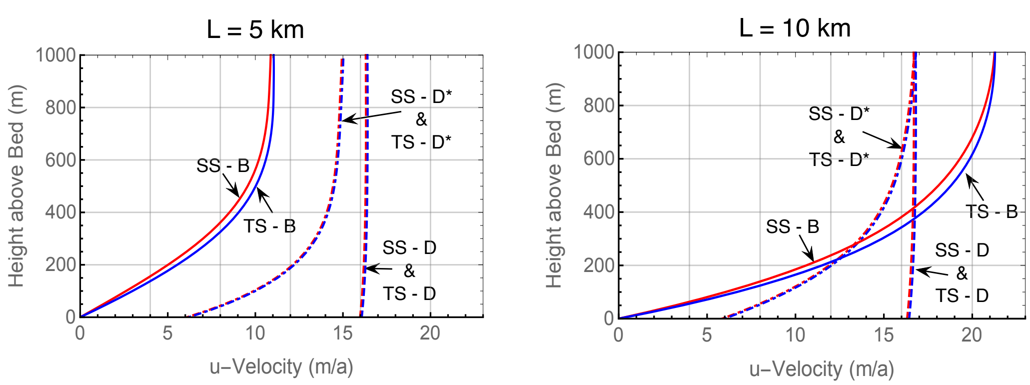

Figure 4. The u-velocity profile at location $x = L$ as a function of height from the bed.

Fig. 3 evaluates the convergence of the two Stokes formulations as a function of mesh resolution r , where r is the number of quadrilaterals in either direction. The models converge to the same value of the transport, obtained by Richardson extrapolation, and convergence is second order as expected from the use of linear elements. Interestingly, in Test B the transformed Stokes model displays considerably smaller error at all resolutions than the standard model. As a result, the standard Stokes calculations are not fully converged even at the 40x40 resolution. Fig. 4 shows the vertical profiles of the horizontal velocity u at outflow, $x = L$. Results from the no-slip Test B problem and the two frictional sliding problems, Tests D and D^*, are plotted. The Test D profile from the ISMIP-HOM benchmark is almost vertically constant, indicating that the value for basal friction that was originally chosen is too small. This motivated the change of Test D to Test D^* in Appendix C, with larger basal friction.

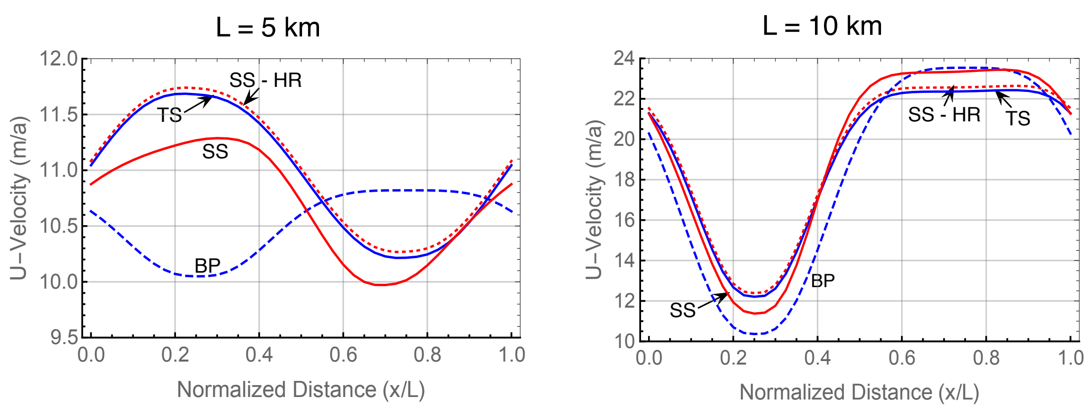

Figure 5. Upper surface u-velocity, $u(x, z_s)$ - Test B, No-slip boundary conditions.

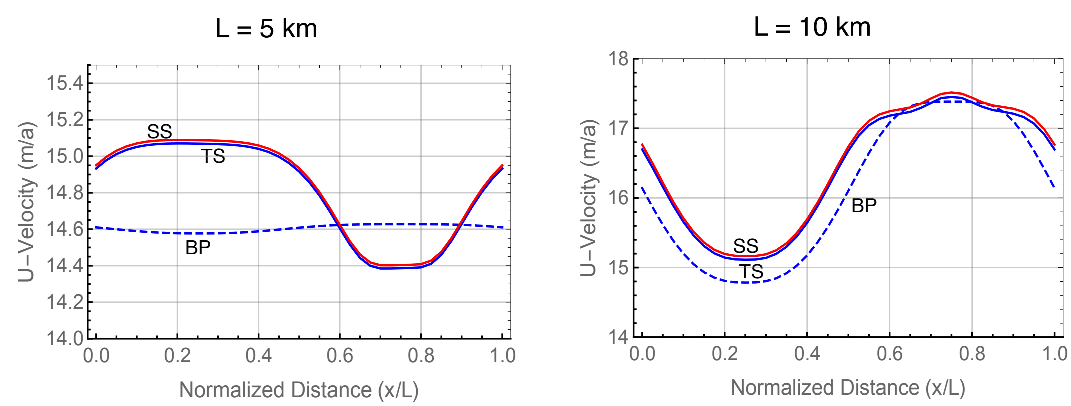

Figure 6. Upper surface u-velocity, $u(x, z_s)$ - Test D^*, Modified frictional sliding.

Figs. 5 and 6 show the u-velocity at the ice sheet upper surface for Tests B and D^*. This is the benchmark used in ISMIP-HOM to compare different ice sheet models. Four cases are compared: the standard Stokes model (SS), the transformed Stokes model

(TS), the Blatter-Pattyn (BP) model, and for reference, a very high resolution full-Stokes Test B calculation "oga1" (SS-HR), available from the ISMIP-HOM paper and also independently available in Gagliardini and Zwinger (2008). The TS and the SS-HR plots lie on top of one another (they have been offset slightly for clarity), indicating that the TS model is already fully converged. As before, the SS model is not yet converged in Test B, particularly at $L=5$ km. As in the ISMIP-HOM study, the Blatter-Pattyn calculation (BP) shows large deviations from the Stokes results, especially so at $L=5$ km where the surface velocity is entirely out of phase with the Stokes results. Fig. 6 shows that results for the SS and TS models essentially coincide in Test D* (the SS plot has been slightly offset upward for clarity). As expected, Blatter-Pattyn results show noticeable error at $L=10$ km and very large error at $L=5$ km. Pressure results have little or no physical significance so they are not presented. However, it was observed that pressures on the P1-E0 mesh were particularly smooth and well behaved.

6. Applications of the Transformed Stokes Model

6.1 Adaptive Switching between Stokes and Blatter-Pattyn Models

A way of reducing the cost of a full Stokes calculation is to use it adaptively together with a cheaper approximate model. That is, one may use the cheaper model in those parts of a problem where it is accurate, and the more expensive full Stokes model where the approximate model loses accuracy. One example of such an adaptive approach is the tiling method by Seroussi et al. (2012). However, such methods have drawbacks, such as the need to incorporate two or more presumably quite different models into a single model together with a transition zone to couple the disparate models.

On the other hand, the transformed Stokes model is particularly simple when used in such an adaptive role because it may be locally switched between the Stokes and Blatter-Pattyn cases simply by switching the parameter $\xi \in \{0,1\}$ on and off. Note that the extended Blatter-Pattyn approximation ($\xi=0, \hat{\xi}=1$) needs to be used since then both the Stokes and Blatter-Pattyn parts of the code will have the same number of degrees of freedom. This also implies the use of the P1-E0 mesh since the extended Blatter-Pattyn model requires a mesh satisfying the solvability condition, as explained in §4.3. The extended Blatter-Pattyn approximation is much cheaper than the full Stokes model. For example, the extended Blatter-Pattyn calculation of Test B at a 40x40 resolution as in Figs. 5 and 6 is some 4.3 times cheaper than the corresponding transformed Stokes calculation. The standard Blatter-Pattyn approximation ($\xi,\hat{\xi}=0$) would be somewhat cheaper but it would need more complicated programming, i.e., it would require coupling to the Stokes model with $p=0$ and $w=w(u,v)$ at the interface between models.

To demonstrate the idea of adaptive switching with a transformed Stokes model, we introduce a new problem, Test O, described in Appendix C and illustrated in Fig. C1. This consists of an inclined ice slab whose movement is obstructed by a thin obstacle protruding 20% of the ice depth up from the bed. No-slip boundary conditions are applied along the bed and the obstacle. Because of the localized nature of the obstacle, the condition for the validity of the Blatter-Pattyn approximation, (37), must fail near the obstacle and therefore the full Stokes model is needed for good accuracy, at least locally.

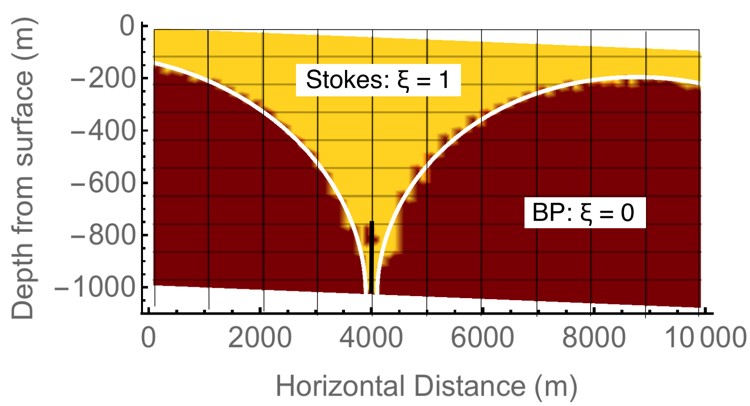

Figure 7. Mask function (white curve, $z=F_M(x)$) to indicate where the Stokes and BP models are activated in the 20% obstacle test problem. The dark brown region delineates the region where $|\partial w/\partial x| \le 0.1|\partial u/\partial z|$ from a Blatter-Pattyn calculation.

To implement the method, we first use a Blatter-Pattyn calculation to outline regions where $|\partial w/\partial x| \le 0.1|\partial u/\partial z|$, approximately localizing where the Blatter-Pattyn approximation is valid. This determines a mask function $z=F_M(x)$, illustrated in Fig. 7 by the white curves, specifying where the two models must be used. Defining the centroid of a triangular element by (x_C,z_C), the code makes a selection in each element:

$$\begin{aligned} z_C \le F_M(x_C) \quad &\Rightarrow \quad \text{Set } \xi=0, \quad \text{i.e., the Blatter-Pattyn region,} \\ z_C > F_M(x_C) \quad &\Rightarrow \quad \text{Set } \xi=1, \quad \text{i.e., the Stokes region.} \end{aligned}$$

The Stokes region occupies the upper part of the domain in Fig. 7 and includes the obstacle, while the Blatter-Pattyn region occupies much of the bottom part of the domain. It is possible to implement a transition zone, e.g., $0 \le \xi(x,z) \le 1$, but this was not done in the present calculation.

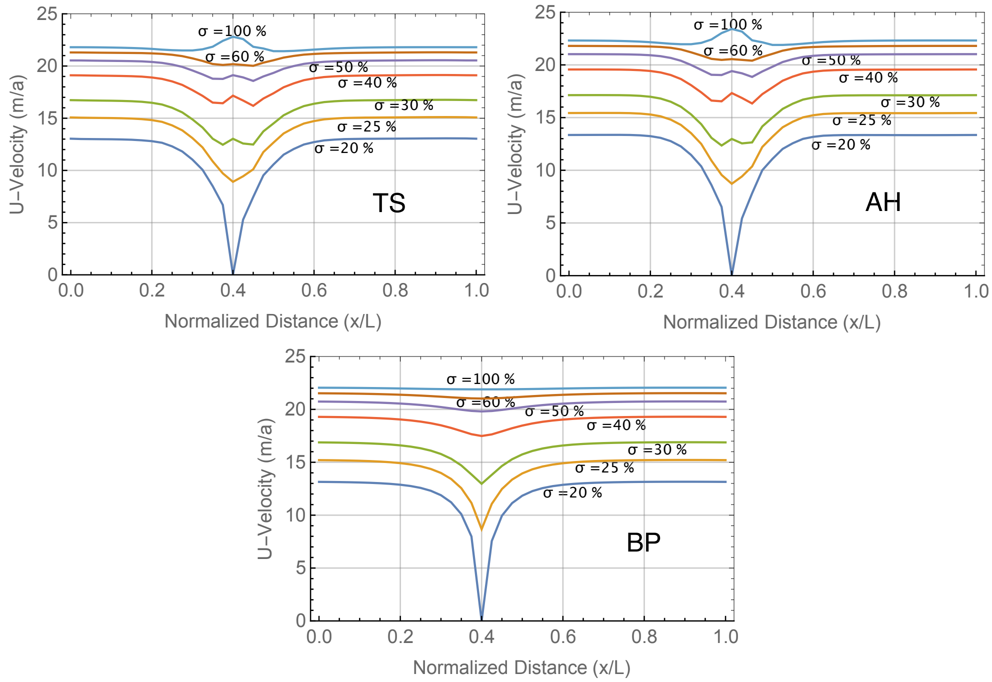

Figure 8. Horizontal velocity profiles for the Transformed Stokes (TS), the Adaptive-Hybrid (AH), and the Blatter-Pattyn (BP) models in Test O.

Results from the adaptive-hybrid calculation are shown in Fig. 8, showing curves of the horizontal velocity u at seven different vertical locations specified as a percentage of the distance between top and bottom, with $\sigma = 100\%$ at the top surface. The top right panel shows the results for the adaptive-hybrid calculation (AH), and the top left panel and bottom panel show results for the Stokes (TS) and Blatter-Pattyn (BP) calculations, respectively. All calculations use the 40x40 mesh resolution. The adaptive-hybrid results are very similar to the full Stokes results, reproducing most features of the velocity profiles, including the velocity bump at the top surface, indicating that even the top surface feels the presence of the obstacle. The Blatter-Pattyn results are much less accurate, completely missing the details of the flow near the obstacle. The calculated RMS error relative to the Stokes results in the Blatter-Pattyn case is 0.493 m/a and 0.440 m/a in the adaptive-hybrid case. It is larger in the Blatter-Pattyn case, as expected, but the difference is smaller than expected from the visual differences in Fig. 8. Nonetheless, the adaptive-hybrid method must be judged successful if judged by the Fig. 8 results alone. As explained later, computational cost comparisons may be misleading and are not provided.

6.2. Two New Stokes Approximations

As shown in §3.4, simply setting $w=0$ in the second invariant $\tilde{\dot{\varepsilon}}^2$ in the transformed functional $\tilde{\mathcal{A}}$ results in the standard Blatter-Pattyn approximation. This suggests that approximating rather than neglecting the vertical velocity w in the functional would create better approximations. We will look at two such methods, although others are possible. The first method, called the Herterich approximation, may be viewed as Picard iteration of the transformed functional (32) where the vertical velocity, obtained from the solution of the continuity equation, is a lagged value from the previous iteration. Two ways of implementing this method are described. The second method, to be called the Dual-Mesh approximation, approximates the vertical velocity by discretizing the continuity equation on a coarser mesh. Since vertical velocity comes from inverting the continuity equation, this has the effect of approximating the vertical velocity and also reducing the number of pressure and vertical velocity variables while preserving the structure of the Stokes model. The degree of approximation is determined by the amount of coarsening of the continuity mesh.

6.2.1 The Herterich Approximation

Noting that $\tilde{P}=0$ in the Blatter-Pattyn approximation, we drop the pressure term from the transformed functional (32) and define a new 2D pressure-free functional,

$$\tilde{\mathcal{A}}_{PS1}[u]=\int_V dV\left[\frac{4n}{n+1}\eta_0\left(\tilde{\dot{\varepsilon}}^2\right)^{(1+n)/2n}+\rho g u\frac{\partial z_s}{\partial x}\right]+\frac{1}{2}\int_{S_{B2}}dS\,\beta(x)\left(u^2+\zeta\left(u\,n_x/n_z\right)^2\right), \tag{61}$$

where

$$\tilde{\dot{\varepsilon}}^2=\left(\frac{\partial u}{\partial x}\right)^2+\frac{1}{4}\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)^2 \tag{62}$$

is the 2D version of the transformed second invariant, (27). Incompressibility must be enforced separately by a second functional,

$$\tilde{\mathcal{A}}_{PS2}[p]=\int_V dV\,p\left(\frac{\partial u}{\partial x}+\frac{\partial w}{\partial z}\right). \tag{63}$$

Direct substitution is used for boundary conditions in both functionals, as in (15) or (32). Note that functional (61) is a function of u, w but variation is to be taken only with respect to u. Similarly, functional (63) is a function of u, w, p but variation is taken only with respect to p. We are effectively viewing the pressure p as a "test function" in the finite element sense. This gives flexibility to create elements that satisfy the solvability condition (58). For example, in a triangulation some pressures may be assigned to every two triangles, as in a P1-E0 mesh, while others may be assigned to a single triangle so as to achieve an equal number of pressure and vertical velocity variables.

The variation of $\tilde{\mathcal{A}}_{PS1}[u]$ with respect to u results in a set of n_u discrete Euler-Lagrange equations corresponding to the u-momentum equation,

$$\hat{R}_U\left(u,w\right)=\frac{\partial\tilde{\mathcal{A}}_{PS1}\left(u,w\right)}{\partial u}=\mathbf{M}_U\left(u,w\right)+F_U=0\ . \tag{64}$$

When w is set to zero, this is just the Blatter-Pattyn model, (49). The discrete variation of $\tilde{\mathcal{A}}_{PS2}[p]$ with respect to p, results in the n_p equations of the continuity equation, (55),

$$\hat{R}_P\left(u,w\right)=\frac{\partial\tilde{\mathcal{A}}_{PS2}\left(u,w,p\right)}{\partial p}=\mathbf{M}_{UP}^T u+\mathbf{M}_{WP}^T w=0\ . \tag{65}$$

These two systems are combined to form the complete system of discrete equations of the Herterich approximation,

$$\hat{R}\left(u,w\right)=\left[\hat{R}_U\left(u,w\right),\hat{R}_P\left(u,w\right)\right]^T=0\ . \tag{66}$$

This is a single system of n_u+n_p equations to determine the n_u+n_w discrete variables u, w, implying that (66) is a viable system for u, w only on meshes satisfying the solvability condition, $n_p=n_w$. Just as in the standard Blatter-Pattyn approximation in §3.4.1, the vertical momentum equation is missing, but here instead of neglecting w, the vertical velocity is obtained consistently from the continuity equation.

The continuum form of the discrete Euler-Lagrange system (64), (65) is given by

$$\frac{\partial}{\partial x}\left(4\tilde{\mu}\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial z}\left(\tilde{\mu}\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)\right)-\rho g\frac{\partial z_s}{\partial x}=0, \qquad \frac{\partial u}{\partial x}+\frac{\partial w}{\partial z}=0, \tag{67}$$

along with the boundary conditions,

$$\begin{aligned}
&4\tilde{\mu}\frac{\partial u}{\partial x}n_x+\tilde{\mu}\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)n_z=0 \text{ on } S_S, \quad u=w=0 \text{ on } S_{B1},\\
&\left.4\tilde{\mu}\frac{\partial u}{\partial x}n_x+\tilde{\mu}\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)n_z+\beta(x)\left(u+\zeta\, u\left(n_x/n_z\right)^2\right)=0,\ w=-u\,n_x/n_z,\right\} \text{ on } S_{B2},
\end{aligned} \tag{68}$$

where the viscosity is $\tilde{\mu}=\eta_0\left(\tilde{\varepsilon}^2\right)^{(1-n)/2n}$ as in (26). Remarkably, this same model was introduced by Herterich (1987) to study the transition zone between an ice sheet and an ice shelf[1]. Thus, it predates the widely used Blatter-Pattyn model by some eight years. Unfortunately, in contrast to Blatter-Pattyn, the Herterich model was not pursued further.

There are two algorithms for solving the Herterich system (66), as follows

(1) Herterich Approximation, Newton/Picard iteration version:

If we lag the vertical velocity, i.e., $w^{K+1}=w^K \Rightarrow \Delta w=0$ in the first equation of (70), we obtain a Picard iteration algorithm as follows

$$\begin{aligned}
&\text{Starting from } K=0,\ \text{choose an initial guess, } u^0\neq 0, w\left(u^0\right),\\
&\text{Solve: } \mathbf{M}_{UU}\left(u^K,w^K\right)\Delta u+\hat{R}_U\left(u^K,w^K\right)=0,\\
&\qquad u^{K+1}=u^K+\Delta u, \quad w^{K+1}=w\left(u^{K+1}\right)=-\mathbf{M}_{WP}^{-T}\mathbf{M}_{UP}^{T}u^{K+1},\\
&\qquad K=K+1, \qquad \text{Repeat until convergence.}
\end{aligned} \tag{69}$$

Each step of the iteration is inexpensive since it is equivalent to a step of the standard Blatter-Pattyn model, (35). On the other hand, Picard iterations typically converge only linearly. It remains to be seen which version is preferable in practice.

(2) Herterich Approximation, Quasi-variational, Newton iteration version:

Although a single variational principle does not exist, it is still possible to make use of Newton-Raphson iteration to obtain second order convergence. To do this, we treat (66) as a single multidimensional nonlinear system and solve it using Newton-Raphson iteration as follows, with K the iteration index and $\Delta u, \Delta w$ defined previously,

$$\begin{aligned}
\mathbf{M}_{UU}\left(u^K,w^K\right)\Delta u+\mathbf{M}_{UW}\left(u^K,w^K\right)\Delta w+\hat{R}_U\left(u^K,w^K\right)&=0,\\
\mathbf{M}_{UP}^{T}u^{K+1}+\mathbf{M}_{WP}^{T}w^{K+1}&=0.
\end{aligned}, \tag{70}$$

where $\mathbf{M}_{UU}(u,w)=\partial\hat{R}_U(u,w)/\partial u$ and $\mathbf{M}_{UW}(u,w)=\partial\hat{R}_U(u,w)/\partial w$ are the same matrices as appear in (48). Convergence is rapid (quadratic) once in the basin of attraction but each step is more expensive than in the Picard iteration described next. Both versions

[1] Reference provided by C. Schoof.

depend on an invertible continuity equation to obtain $w = w(u)$. However, vertical velocity w may already be available for the purpose of temperature advection in production code packages.

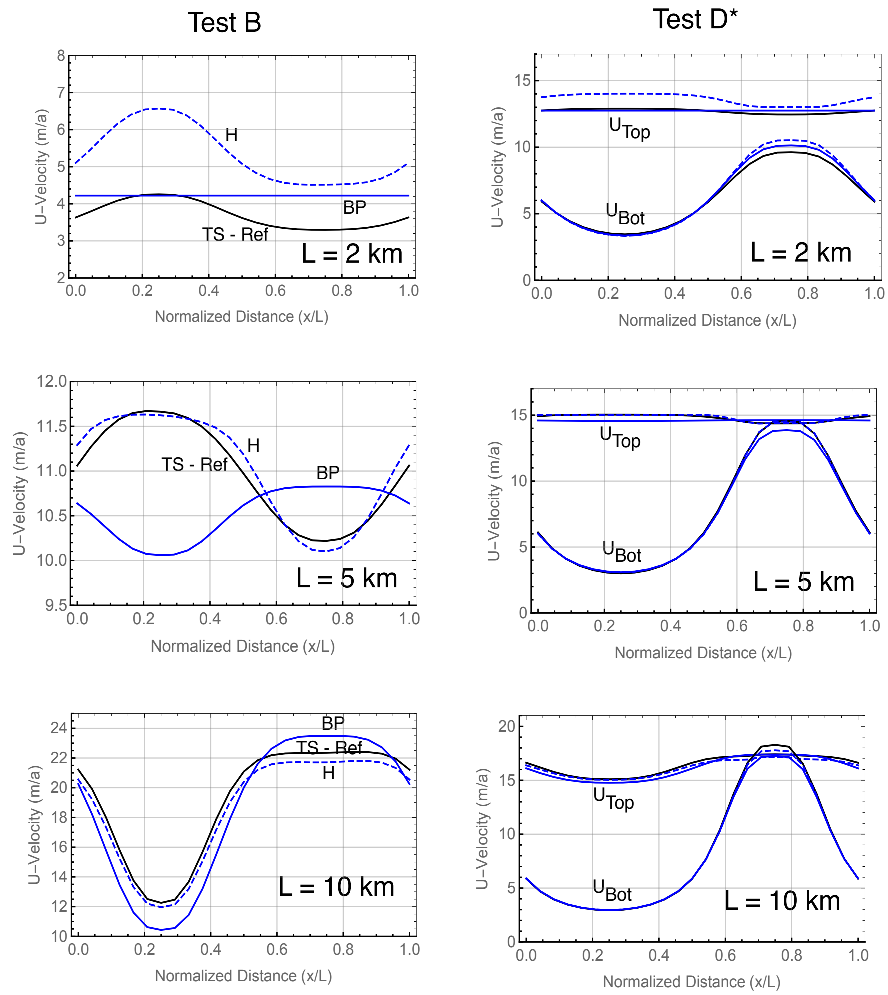

Figure 9. Comparing Approximations. Test B, Upper surface u-velocity, and Test D*, Upper and Lower surface u-velocity. TS-Ref: Transformed Stokes; BP: Blatter-Pattyn; H: Herterich. Resolution: 24x24.

The two Herterich algorithms converge to the same result. Fig. 9 compares the Herterich (H) and the Blatter-Pattyn (BP) approximations to a reference exact transformed Stokes solution (TS-Ref) for Test B results on the left and Test D* results on the right at three problem aspect ratios in the range where small aspect ratio approximation would begin to fail, $\varepsilon = H_0 / L = 0.1, 0.2, 0.5$, where $H_0 = 1000\,m$ is the nominal ice thickness and L is the domain length. One might expect the Herterich approximation to be more accurate than the BP approximation since it retains more of the terms from the transformed Stokes model equations than does the Blatter-Pattyn model but this is not borne out in the Fig. 9 results. The Herterich approximation is more accurate in Test B at $L = 5,10\ km$, and this is confirmed by the RMS u-Error results in Fig. 12 where the Herterich model is two to three times more accurate than BP. However, this not the case at $L = 2\,km$ in both Tests B and D*, although Test B results may be deceptive because only the upper surface velocity is displayed (lower surface velocity is zero) and this expands the scale accentuating differences between the three cases. Test D* results are closely clustered except in the $L = 2\ km$ case where, surprisingly, the Herterich approximation is the least accurate. These results cast some doubt on the Herterich approximation, suggesting that it may need to be evaluated further.

6.2.2 A "Dual-Mesh" Transformed Stokes Approximation

In the Herterich approximation we approximated the vertical velocity directly by lagging w in the second invariant in the variational functional. Here we approximate w indirectly by instead approximating the continuity equation, which determines the vertical velocity. That is, the continuity equation is discretized on a mesh coarser than the one used for the momentum equations. The vertical velocity is then interpolated to appropriate locations on the finer mesh. This reduces the number of unknown variables in the problem, making it cheaper to solve but hopefully without serious loss of accuracy. As described in Appendix C, our test problem meshes are logically rectangular, i.e., divided into n cells horizontally and m cells vertically. The coarse mesh is constructed by dividing the fine mesh into s equal segments in each direction. This assumes that the integers n and m are each divisible by s, so that there are nm/s^2 coarse cells in total with each coarse cell containing s^2 fine cells. The primary mesh (i.e., the fine mesh) has been chosen to have $n = m = 24$, i.e., a reference 24×24 fine mesh, so as to maximize the number of different coarse meshes available. Coarse meshes were constructed using $s = 2,3,4,6$, and this resulted in fine/coarse mesh combinations labeled by $24 \times 12,\ 24 \times 8,\ 24 \times 6,\ 24 \times 4$, respectively. Similar to a P1-E0 fine mesh, coarse mesh vertical velocities w are located at vertices and pressures at vertical edges. Fig. 10 illustrates the case of a single coarse cell and four fine quadrilateral cells for a mesh fragment with $n = m = 2$ and $s = 1$. In the Test B problem with direct substitution for basal boundary conditions there will be nm u-variables and nm/s^2 w- and p-variables each, for a total of $nm\left(1 + 2/s^2\right)$ unknown variables, considerably fewer than the $3nm$ variables in the full resolution mesh. The coarse mesh terms in the functional that are affected, $\tilde{P}\left(\partial u/\partial x + \partial w/\partial z\right)$ and $\partial w/\partial x$, are computed by interpolating coarse mesh variables to the fine mesh. We consider two versions of the approximation depending on how the coarse mesh terms are interpolated to the fine mesh.

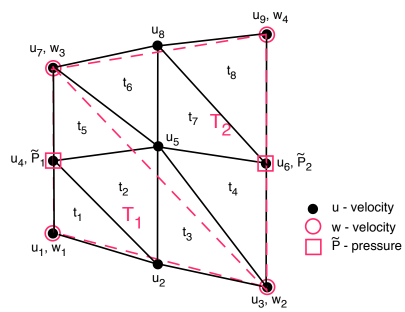

Figure 10. A Sample of a Coarse/Fine P1-E0 Mesh for the Dual-Mesh Approximation. Resolution: $n = m = 2, s = 1$. Coarse mesh is in red, fine mesh in black.

(1) <u>Approximation A, Bilinear interpolation</u>:

Referring to Fig. 10, the four velocities at the vertices of the coarse mesh quadrilateral, i.e., u_1, u_3, u_7, u_9 and w_1, w_2, w_3, w_4, are used to obtain u, w at the remaining five vertices of the fine mesh by means of bilinear interpolation. Thus, the five velocities u_2, u_4, u_5, u_6, u_8 are obtained in terms of vertex velocities u_1, u_3, u_7, u_9, and similarly for the w velocities. The resulting complete set of fine mesh variables is used calculate the divergence $D = \left(\partial u/\partial x + \partial w/\partial z\right)$ and the quantity $\partial w/\partial x$ in each of the eight triangular elements $t_1, t_2, \cdots, t_8$ of the fine mesh. Coarse mesh pressures $\tilde{P}_1, \tilde{P}_2$ are associated with the coarse mesh triangles T_1, T_2. The products $\tilde{P}_1 D$ in elements t_1, t_2, t_3, t_5 and $\tilde{P}_2 D$ in elements t_4, t_6, t_7, t_8 are then accumulated over the entire mesh to obtain the quantity $\tilde{P}\left(\partial u/\partial x + \partial w/\partial z\right)$ for use in the transformed functional $\tilde{\mathcal{A}}$. Similarly, the quantity $\partial w/\partial x$ is computed in the fine mesh elements from coarse mesh variables for use in the second invariant $\tilde{\varepsilon}^2$.

(2) Approximation B, Linear interpolation:

In this version, the three velocities at the vertices of the two coarse mesh triangles T_1 and T_2, i.e., u_1,u_3,u_7 and w_1,w_2,w_3 in T_1, and u_7,u_3,u_9 and w_3,w_2,w_4 in T_2, approximate the divergence $D=\left(\partial u/\partial x+\partial w/\partial z\right)$ and $\partial w/\partial x$ as constant values in the two coarse triangles. The constant values $\tilde{P}_1 D$, $\tilde{P}_2 D$ are then accumulated over the entire mesh. The constant value $\partial w/\partial x$ in each coarse triangle is then distributed to each of the eight fine mesh elements $t_1,t_2,\cdots,t_8$, depending on whether the centroid of the fine triangular element is in T_1 or T_2. As in the previous case, this is then used in the second invariant $\tilde{\varepsilon}^2$ in the transformed functional $\tilde{\mathcal{A}}$.

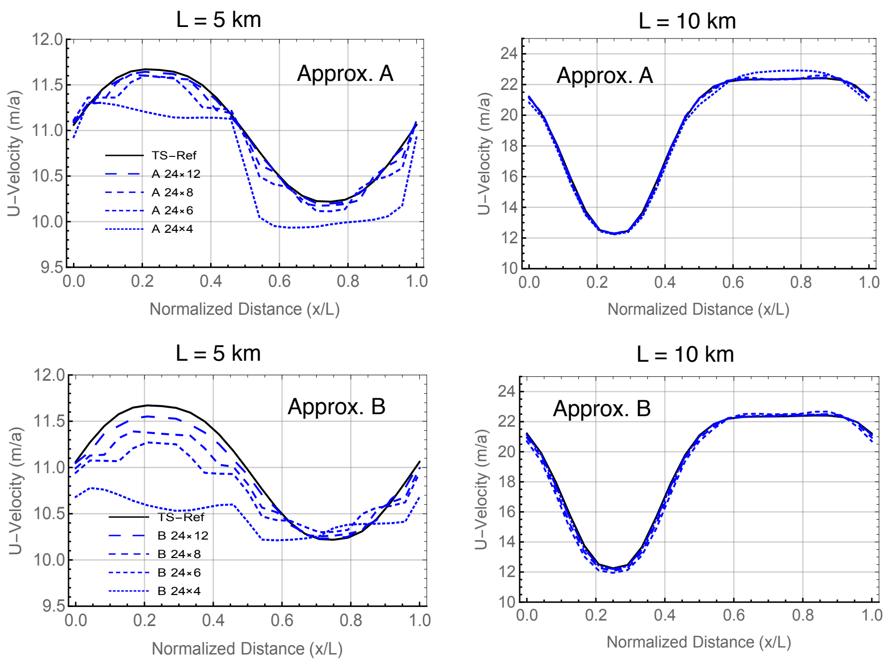

Figure 11. Comparing Approximations A and B. Test B. Upper surface u-velocity. TS-Ref: Ref. Stokes (r x r = 24x24); Fine/Coarse resolutions (r x R = 24x12, 8, 6, 4).

While the number and type of unknown variables is the same in the two versions, they differ considerably in accuracy as seen in Figs. 11 and 12. Fig. 11 compares the upper surface u-velocity in approximations A and B, for the five mesh combinations, to the reference 24x24 fine mesh transformed Stokes (TS) calculation. The corresponding Blatter-Pattyn result is shown in Fig. 9. Fig. 12 summarizes and compares the overall accuracy of all approximations by means of the RMS u-Error relative to the TS calculation. As might be expected, the accuracy of Approx. A is better than Approx. B, particularly in the case $L = 10$ km. Both dual-mesh versions are more accurate than the BP and H approximations, except for the lowest 24x4 resolution of Approx. B. However, it may be surprising that the SS calculation is the least accurate but this is because SS is not yet fully converged at this resolution (see Fig. 5).

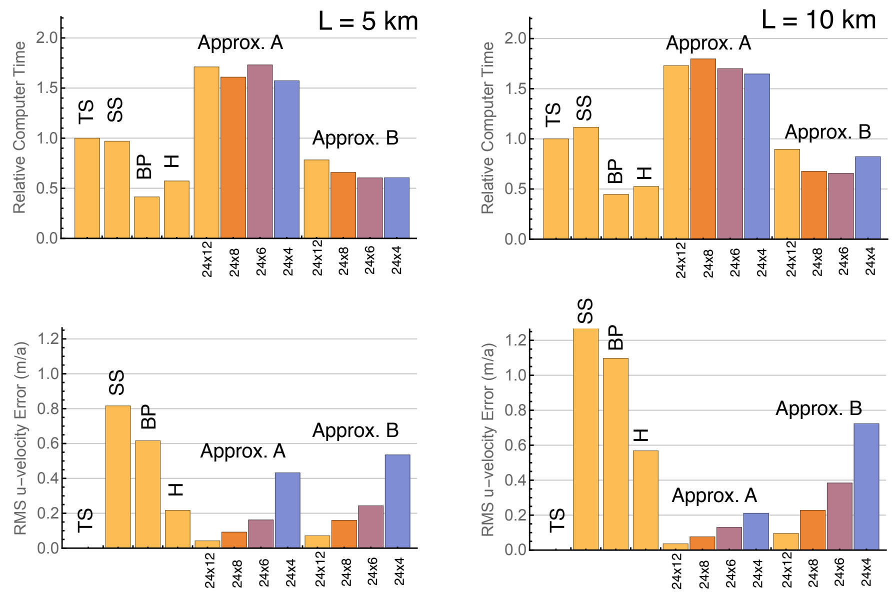

Figure 12. Comparing the Test B computer time and RMS u-Error results relative to a TS calculation for Approximations BP, H, A, B, and the SS Stokes model. TS, SS, BP, H: Resolution (r x r = 24x24); Approx. A, B: (r x R = 24x12, 8, 6, 4).

A qualitative comparison of computational time relative to the TS calculation is presented in Fig. 12 for the 24x24 resolution case. The cost of the BP approximation is substantially lower than the cost of full Stokes, while the cost of the Herterich approximation is only slightly higher than BP. This comparison is qualitative and rough because the present calculations are not at all representative of the computer hardware or the methods used in practical ice sheet modeling. The calculations are proof of concept only, made on a personal computer using the program Mathematica, a slow interpreted language. As a result the cost comparisons are inaccurate and possibly misleading. No effort was made to optimize the calculations or to take advantage of parallelization. Thus, for example, the cost of dual-mesh approximations A and B is greatly affected by the cost of the associated interpolations, making them more costly than even the unapproximated calculations, TS and SS. However, the interpolations are highly parallelizable and so would contribute little to the cost in practical computations.

7. Second-Order Discretizations

So far we have used first-order elements, primarily P1-E0. However, in current practice Stokes models are more often based on second-order elements such as the popular Taylor-Hood P2-P1 element (e.g., Leng et al., 2012; Gagliardini et al., 2013). In 2D the P2-P1 element, illustrated in Fig. 13A, has velocities on element vertices and edge midpoints and pressures on element vertices, resulting in quadratic velocity and linear pressure distributions. This element satisfies the inf-sup stability condition (e.g., Elman et al., 2014) but does not satisfy the solvability condition (58). For example, Test B with direct substitution for basal boundary conditions requires $n_w = 4nm$ vertical velocity variables, many more than the $n_p = n(m+1)$ pressure variables.

Stokes models work well with a Taylor-Hood mesh as shown in Fig. 13 where both the P2-P1 and P1-E0 models converge to a common Test B solution. Recall that the P2-P1 mesh does not work with the extended Blatter-Pattyn approximation, a model that requires the solvability condition. However, it is possible to construct a second-order element consistent with an invertible continuity equation. This is the P2-E1 element, as illustrated in Fig. 13A. This is second-order for velocities and linear for pressure, just like the P2-P1 element, but the pressure is edge-based rather than vertex-based. Pressures are located midway between velocities along vertical lines, as illustrated in Fig. 13A. As explained in Appendix A, this mesh must be constructed using vertical columns of quadrilaterals. Since pressures are vertically collinear with velocities, as in the P1-E0 element, the analysis in Appendix A confirms that this element also satisfies the solvability condition (58). A 3D analog exists as is explained in Appendix A.

Fig. 13B shows the approximate error of the ice transport as a function of mesh refinement for the second-order P2-P1 and P2-E1 meshes in transformed Stokes Test B calculations, together with similar results for the first-order P1-E0 mesh from Fig. 3, for comparison. Note that both second order models show approximately the same error at resolution $r = 16$ as the first order P1-E0 model at resolution $r = 40$, and the same for coarser resolutions such as $r = 8$ and $r = 20$, respectively. However, the second-order calculations are considerably more expensive than the first-order calculations at comparable resolution or accuracy.

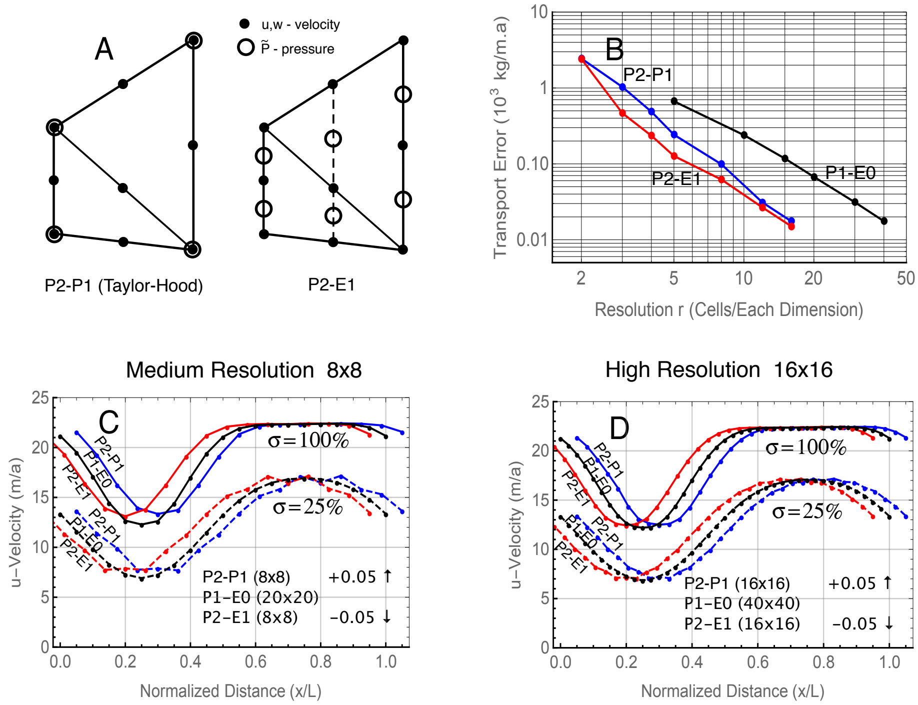

Figure 13. Comparing second-order discretizations based on the P2-P1 and P2-E1 elements from panel A to first-order discretizations using the P1-E0 element in Test B calculations with L=10 km. Only transformed Stokes calculations are compared; standard Stokes results behave similarly. Panel B compares the convergence and accuracy of various schemes with increasing resolution, while panels C, D compare horizontal velocities at medium and maximum resolutions.

Panels C, D in Fig. 13 compare u-velocities from several Test B calculations using the two second-order models to the first-order P1-E0 model results from Fig. 3. Each panel shows upper surface velocities ($\sigma = 100\%$) in solid lines and velocities from a surface a quarter of the way up from the bottom ($\sigma = 25\%$) in dashed lines. Panel C shows medium resolution calculations ($r = 8, 20$ in the second-order and first-order calculations, respectively) and panel D shows higher resolution calculations ($r = 16, 40$). At these resolutions the accuracy of the first- and second-order calculations is very similar so for clarity the second-order results are displaced horizontally from the first-order results by 0.05 nondimensional units. The P2-E1 results in magenta are displaced to the left and the P2-P1 results in blue are displaced to the right. In general, models that satisfy the solvability condition, the P1-E0 and P2-E1 models, are better behaved than the P2-P1 model. This may be related to the well-known “weak” mass conservation of the Taylor-Hood element, which is greatly improved by “enriching” the pressure space with constant pressures in each triangular element (Boffi et al., 2012). Thus, in the 2D Test B problem the pressure variables increase from $n_p = n(m+1)$ in the basic Taylor-Hood element to $n(3m+1)$, much closer to the $4nm$ needed to satisfy the solvability condition. However, pressure in the P2-E1 case is highly oscillatory but well behaved in the P2-P1 case. This is consistent with the discussion in Appendix B since the P2-P1 mesh satisfies the inf-sup condition, which is concerned with regularizing the pressure solution (i.e., a condition on the $\mathbf{B}^T$ matrix), while the P2-E1 mesh has no such requirement.

8. A Summary and Discussion

In summary, this paper presents two innovations in ice sheet modeling. The first involves a transformation of the ice sheet Stokes equations into a form that differs from the approximate Blatter-Pattyn system by a small number of terms. In particular, the variational formulations differ only by the absence of terms involving the vertical velocity w in the second invariant of the strain rate tensor in the Blatter-Pattyn system.

We focus on two applications of the new transformation, although others may be possible. The first allows these extra terms to be "switched" on or off to convert the code from a full-Stokes model to a Blatter-Pattyn model if desired. Ice sheet flow is generally shallow but often contains limited areas where Stokes equations must be solved. Thus, the switch from Blatter-Pattyn to Stokes may be done locally and adaptively only where the extra accuracy is required.

The fact that neglecting the vertical velocity in only one localized place creates the Blatter-Pattyn approximation suggests that an improved approximation will result from approximating rather than neglecting the vertical velocity. We present two such approximations but again others may be possible. The first approximation solves the pressure-free horizontal momentum equation with the vertical velocity obtained from the continuity equation. This approximation turns out to be the same as a model originally developed by Herterich (1987). We therefore refer to it as the Herterich approximation. The second approximation is obtained by discretizing the continuity equation on a coarser mesh than that used by the rest of the model, which yields an approximate vertical velocity and thus an approximate Stokes model.

The second innovation introduces the use of finite element discretizations that create a decoupled invertible continuity equation. This allows for the numerical solution of the vertical velocity in terms of the horizontal velocity components, i.e., $w = w(u,v)$. These meshes are a prerequisite for the applications mentioned above. Examples of such meshes for use in 2D and 3D are given in Appendix A, including the P1-E0 mesh that is used in most of the test problems of this paper. However, perhaps the most important consequence of this is that invertibility of the continuity equation serves as an alternative condition for the well posedness of Stokes saddle point problems, as explained in Appendix B.

Code Availability

All calculations were made using the Wolfram Research, Inc. program Mathematica in a development environment. A large number of Mathematica Notebooks were used to produce the various results. A representative Notebook (a ,nb file) is available in a public repository at https://doi.org/10.5281/zenodo.13940989 for a Test $D*$ calculation (described in Appendix C) at $L = 10\,\text{km}$ and a resolution of 20x20. This test problem was chosen to demonstrate the use of direct substitution for the frictional tangential boundary condition, (16), in the functional. A Mathematica Notebook may be viewed by downloading the free Wolfram Player from https://www.wolfram.com/player/. However, the full Mathematica code is required for editing or execution.

Competing Interests

The author has acknowledged that there are no competing interests.

Acknowledgements

I am grateful to Steve Price for support, and especially to Mauro Perego, William (Bill) Lipscomb, and reviewers, particularly Christian Schoof, for comments and suggestions that helped to improve the paper.

References

Auricchio, F., da Veiga, L.B., Brezzi, F., and Lovadina, C.: Mixed Finite Element Methods, In Encyclopedia of Computational Mechanics Second Edition (Eds E. Stein, R. de Borst, and T.J.R. Hughes), John Wiley & Sons, Ltd., 2017.

Benzi, M., Golub, G.H., and Liesen, J.: Numerical Solution of Saddle Point Problems, Acta Numerica, 14 , 1-137, doi:10.1017/S0962492904000212, 2005.

Blatter, H.: Velocity and Stress Fields in Grounded Glaciers: A Simple Algorithm for Including Deviatoric Stress Gradients, J. Glaciol., 41, 333-344, 1995.

Boffi, D., Brezzi, F., Fortin, M.: Finite Elements for the Stokes Problem. In: Boffi D., Gastaldi L. (eds) Mixed Finite Elements, Compatibility Conditions, and Applications. Lecture Notes in Mathematics, vol 1939. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-78319-0_2, 2008.

Boffi, D., Cavallini, N., Gardini, F., and Gastaldi, L.: Local Mass Conservation of Stokes Finite Elements, J. Sci. Comput., 52, 383–400, 2012.

Cheng, G., Lötstedt, P., and von Sydow, L.: A Full Stokes Subgrid Scheme in Two Dimensions for Simulation of Grounding Line Migration in Ice Sheets Using Elmer/ICE (v8.3), Geosci. Model Dev., 13, 2245-2258, 2020.

Colinge, J., and Rappaz, J.: A Strongly Nonlinear Elliptic Problem Arising in a Non-Newtonian Fluid Flow Model in Glaciology, Math. Model. Numer. Anal., 33, 395-406,1999.

Dukowicz, J.K., Price, S.F., and Lipscomb, W.H.: Consistent Approximations and Boundary Condition for Ice Sheet Dynamics from a Principle of Least Action, J. Glaciol., 56, 480-496, 2010.

Dukowicz, J.K., Price, S.F., and Lipscomb, W.H.: Incorporating Arbitrary Basal Topography in the Variational Formulation of Ice Sheet Models, J. Glaciol., 57, 461-467, 2011.

Elman, H.C., Silvester D.J., and Wathen A.J.: Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics, 2nd Ed., Oxford University Press, 494 pp, 2014.

Gagliardini, O., and Zwinger, T.: The ISMIP-HOM Benchmark Experiments Performed Using the Finite-Element Code Elmer, The Cryosphere, 2, 67–76, 2008.

Gagliardini, O., Zwinger, T., Gillet-Chaulet. F., Durand, G., Favier, L., de Fleurian, B., Greve, R., Malinen, M., Martín, C., Råback, P., Ruokolainen, J., Sacchettini, M., Schäfer, M., Seddik, H., and Thies, J.: Capabilities and Performance of Elmer/Ice, a New-Generation Ice Sheet Model, Geosci. Model Dev., 6, 1299–1318, doi:10.5194/gmd-6-1299-2013, 2013.

Gerbeau, J-F.: Finite Elements for Mixed and Saddle Point Problems, Chapter 3 (in particular Remarks 3.2 and 3.3), Course CME358, Stanford University, https://web.stanford.edu/class/cme358/handouts.html, 2025.

Greve, R. and Blatter, H.: Dynamics of Ice Sheets and Glaciers, Springer-Verlag, Berlin Heidelberg, 301 pp, 2009.

Heinlein, A., Perego, M., and Rajamanickam, S.: FROSch Preconditioners for Land Ice Simulations of Greenland and Antarctica, SIAM J. Sci. Comput., 44, V339-B367, doi: 10.1137/21M1395260, 2022.

Herterich, K.: On the Flow Within the Transition Zone Between Ice Sheet and Ice Shelf, in Dynamics of the West Antarctic Ice Sheet, Proceedings of a workshop held in Utrecht, May 6-8, 1985, (Eds. C. J. Van Der Veen and J. Oerlemans), D. Reidel, Dordrecht, 185-202, 1987.

Hoffman, M. J., Perego, M., Price, S. F., Lipscomb, W. H., Zhang, T., Jacobsen, D., Tezaur, I., Salinger, A. G., Tuminaro, R., and Bertagna, L.: MPAS-Albany Land Ice (MALI): A Variable-Resolution Ice Sheet Model for Earth System Modeling Using Voronoi Grids, Geosci. Model Dev., 11, 3747–3780, doi:10.5194/gmd-11-3747-2018, 2018.

Hood, P. and Taylor, C.: Numerical Solution of the Navier-Stokes Equations Using the Finite Element Technique, Comput. Fluids, 1, 1-28, 1973.

Larour, E., Seroussi, H., Morlighem, M., and Rignot, E.: Continental scale, high order, high spatial resolution, ice sheet modeling using the Ice Sheet System Model (ISSM), J. Geophys. Res., 117, 1–20, doi:10.1029/2011JF002140, 2012.

Leng, W., Ju, L., Gunzburger, M., Price, S., and Ringler, T.: A Parallel High-Order Accurate Finite Element Nonlinear Stokes Ice Sheet Model and Benchmark Experiments, J. Geophys. Res., 117, 2156–2202, doi:10.1029/2011JF001962, 2012.

Lipscomb, W.H., Price, S.F., Hoffman, M.J., Leguy, G.R., Bennett, A.R., Bradley, S.L., Evans, K.J., Fyke, J.G., Kennedy, J.H., Perego, M., Ranken, D.M., Sacks, W.J., Salinger, A.G., Vargo, L.J., and Worley, P.H.: Description and Evaluation of the Community Ice Sheet Model (CISM) v. 2.1, Geosci. Model Dev., 12, 387-424, 2019.

Nowicki, S.M.J. and Wingham, D.J.: Conditions for a Steady Ice Sheet-Ice Shelf Junction, *Earth Planet. Sci. Lett.,* **265**(1-2), 246-255, 2008.

Pattyn, F.: A New Three-Dimensional Higher-Order Thermomechanical Ice Sheet Model: Basic Sensitivity, Ice Stream Development, and Ice Flow across Subglacial Lakes, J. Geophys. Res., 108(B8), 2382, 2003.

Pattyn, F., Perichon, L., Aschwanden. A., Breuer, B., de Smedt, B., Gagliardini, O., Gudmundsson, G.H., Hindmarsh, R.C.A., Hubbard, A., Johnson, J.V., Kleiner, T., Konovalov, Y., Martin, C., Payne, A.J., Pollard, D., Price, S., Ruckamp, M., Saito, F., Soucek, O., Sugiyama, S., and Zwinger, T.: Benchmark Experiments for Higher-Order and Full-Stokes Ice Sheet Models (ISMIP–HOM*),* The Cryosphere, 2, 95-108, 2008.

Perego, M., Gunzburger. M., and Burkardt, J.: Parallel Finite-Element Implementation for Higher-Order Ice-Sheet Models, J. Glaciol., 58, 76-88, 2012.

Rappaz, J., and Reist, A.: Mathematical and Numerical Analysis of a Three-Dimensional Fluid Flow Model in Glaciology, Math. Model. Mech. Appl. Sci., 15, 37-52, 2005.

Rückamp, M., Kleiner, T., and Humbert, A.: Comparison of ice dynamics using full-Stokes and Blatter-Pattyn approximation: application to the Northeast Greenland Ice Stream, The Cryosphere, 16, 1675-1696, 2022.

Schoof, C.: A Variational Approach to Ice Stream Flow, J. Fluid Mech., 556, 227-251, 2006.

Schoof, C.: Coulomb friction and other sliding laws in a higher order glacier flow model, Math. Models. Meth. Appl. Sci., 20(1), 157–189, 2010.

Schoof, C. and Hewitt, I.: Ice-Sheet Dynamics, Annu. Rev. Fluid Mech., 45, 217–239, 2013.

Schoof, C. and Hindmarsh, R.C.A.: Thin-Film Flows with Wall Slip: An Asymptotic Analysis of Higher Order Glacier Flow Models, Quart. J. Mech. Appl. Math, 63*, 73-114*, 2010.

Seroussi, H., Ben Dhia, H., Morlighem, M., Latour, E., Rignot, E., and Aubry, D.: Coupling Ice Flow Models of Varying Orders of Complexity with the Tiling Method, J. Glaciol., 58, 776-786, 2012.

Tezaur, I. K, Perego, M., Salinger, A. G., Tuminaro, R. S., and Price, S. F.: Albany/FELIZ: A Parallel, Scalable and Robust, Finite Element, First-Order Stokes Approximation Ice Sheet Solver Built for Advanced Analysis, Geosci. Model Dev., 8, 1197-1220, 2015.

Appendix A: Meshes that Satisfy the Solvability Condition

A1 The P1-E0 Element and the Invertibility of the Continuity Equation

As discussed in §4, the invertibility of the discrete continuity equation requires special elements and meshes that satisfy the solvability condition (58). The P1-E0 element, illustrated in Fig. A1, is one such triangular element that has been used on many problems in this paper. To work best for the purpose of inverting the continuity equation for the vertical velocity, i.e., (56), it must be used in a mesh composed of vertical columns subdivided into triangular elements as shown in a representative mesh in Fig. A2. Element P1-E0 has velocities located at triangle vertices for a linear velocity distribution within each triangle (P1), and pressure located on the vertical edge of each triangle for a constant pressure distribution over the two triangles that share the edge (E0). A second order version of this element, the P2-E1 element, is shown in Fig. 13A.

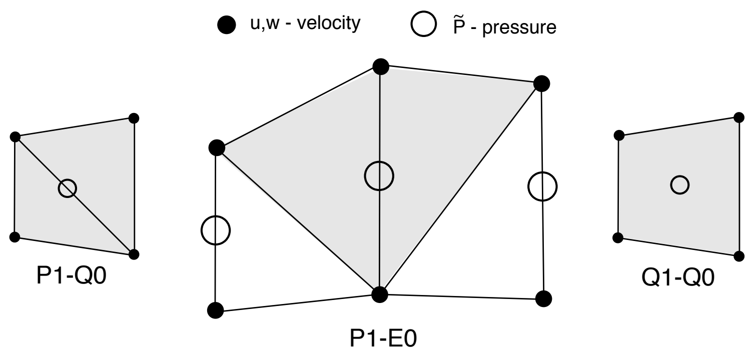

Figure A1. The P1-E0 element, and two other secondary 2D elements that satisfy the solvability condition but possibly only in Tests B and D*.

The triangulation and configuration of pressure basis functions (shown in gray) in Fig. A2 is quite general except that elements must be arranged in vertical columns, which still allows for a flexible triangulation of an arbitrary domain. Generally, there are two triangles associated with a pressure variable, one on each side of a vertical edge, except where the ice sheet ends at a vertical face, as in Fig. A2. However, this is not a problem since the pressure is simply associated with the single triangle on one side of the vertical face. Meshes composed of P1-E0 elements have another useful property. Since pressure and vertical velocity variables alternate along vertical mesh lines, the matrix-vector products $\mathbf{M}_{WP}p$, $\mathbf{M}_{WP}^{T}w$ in (46), corresponding to $\partial\tilde{P}/\partial z$ and $\partial w/\partial z$ in the vertical momentum and continuity equations, respectively, consist of simple decoupled bi-diagonal one-dimensional difference equations along each vertical mesh line for determining pressure and vertical velocity. This will be particularly advantageous for parallelization.

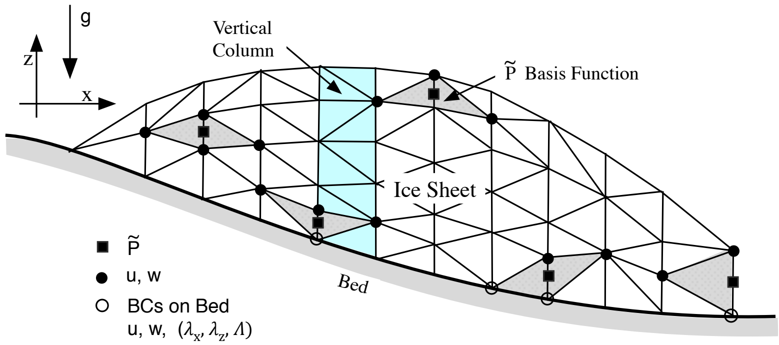

Figure A2. An illustration of a 2D edge-based P1-E0 mesh, composed of vertical columns randomly subdivided into triangles.

The other two elements shown in Fig. A1, the P1-Q0 and Q1-Q0 elements, satisfy the solvability condition when used in Tests B and D* but possibly not in other problems. The P1-Q0 element has velocities on triangle vertices for a linear velocity distribution within each triangle (P1) but pressure is constant within the quadrilateral (Q0) formed by the two adjoining triangles. The Q1-Q0 element has velocities located at quadrilateral vertices and pressure centered in the quadrilateral, resulting in a bi-quadratic velocity distribution (Q1) and a constant pressure within the quadrilateral (Q0). Solutions for all three elements are stable as expected, and they converge to the same value for ice transport. The pressure distribution is smooth in the P1-E0 case but contains small fluctuations near the upper surface in the P1-Q0 and Q1-Q0 cases. These disappear as resolution is increased. The Q1-Q0 element is unstable in conventional applications because it contains a checkerboard pressure null space and is only used in a stabilized form (see Elman et al., 2014, where the element is called Q1-P0). Here, however, it behaves well perhaps because it satisfies the solvability condition. Overall, this confirms our expectation of stability when the solvability condition is satisfied.

A2 Proving the Solvability Condition in P1-E0 Element Meshes

Consider a single vertical edge from top to bottom as in Fig. A2. Assuming there are m edge segments in the vertical direction, there will be $m+1$ discrete w variables and m discrete $\tilde{P}$ variables since each $\tilde{P}$ variable is located between a pair of w variables. However, since the w variable at the bottom (the bed) is specified as a boundary condition, either directly as a no-slip condition or as part of a no-penetration condition, there will be only m unknown w variables. As a result, we have $n_w = n_p$ along each vertical mesh edge and therefore over the entire mesh, thus satisfying the solvability condition. In case Lagrange multipliers are used, there will be $m+1$ unknown discrete w variables (since now the basal vertical velocity w is also an unknown). However, this is matched by m unknown $\tilde{P}$ variables, supplemented by one λ_z or one Λ unknown Lagrange multiplier variable, depending on the type of boundary condition. Thus, again the number of unknown variables equals the number of equations along every vertical edge, thereby satisfying the solvability condition. This means that the P1-E0 element can be used to satisfy the solvability condition irrespective of the boundary conditions on quite arbitrary meshes. These arguments apply to other versions of the P1-E0 element as well, such as the second order version P2-E1 in Fig. 13A or the 3D version in Fig. A3.

A3 Three-Dimensional Meshes Based on the P1-E0 Element

The two-dimensional P1-E0 element has a relatively simple three-dimensional counterpart, shown in Fig. A3. The mesh again consists of vertical columns, this time composed of hexahedra. Each hexahedron is subdivided into six tetrahedra such that each vertical edge is surrounded by as few as four to as many as eight tetrahedra. As in the 2D case, velocity components are collocated at vertices, yielding a piecewise-linear velocity distribution in each tetrahedral element, and pressures are located in the middle of each vertical edge so that pressure is constant in the collection of tetrahedra that surround that edge. Lagrange multipliers, if used, are located at the vertices on the basal surface, yielding a piecewise linear distribution on the basal triangular facet. Since pressures and vertical velocities are again intermingled along a single line of vertical edges from top to bottom, this satisfies the solvability condition (58) because the argument used in the 2D case applies here also.

Fig. A3 shows two of the several possible configurations of a typical hexahedron, including an exploded view of each configuration for clarity. The two configurations differ in having the internal face of the two forward-facing tetrahedra rotated, creating two different forward facing tetrahedra. The remaining six tetrahedra are undisturbed. Since edges must align when hexahedra (or tetrahedra) are connected, this shows that the 3D mesh can be flexibly reconnected and rearranged, just as in the 2D case of Fig. A2.

A closely related and perhaps simpler three-dimensional P1-E0 element is based on the P2-P1 prismatic tetrahedral element in Leng et al. (2012). A mesh of these elements is composed of vertical columns of prisms (with triangular faces at the top and bottom), each subdivided into three tetrahedra. Pressures are located on the vertical prism edges, as in Fig. A3, so this again satisfies the solvability condition.

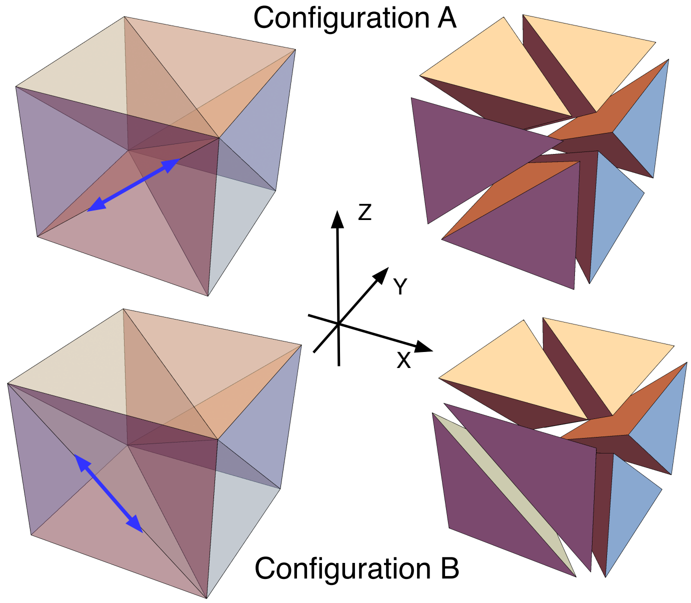

Figure A3. Three-dimensional P1-E0 tetrahedral elements that generalize the 2D P1-E0 element of Fig. A1. Configurations A and B differ by having an internal triangular face rotated, as indicated by the blue arrows. Both satisfy the solvability condition.

Just as the 2D second-order P2-E1 element in Fig. 13A is a generalization of the P1-E0 element, a 3D second-order P2-E1 element may be constructed as a generalization of the P1-E0 element in Fig. A3. Velocities would be located at the vertices and at midpoints of the tetrahedral edges, and pressures halfway between the velocities on vertical edges, including the imaginary vertical edges through the midpoints of the tetrahedral edges as in the 2D case in Fig. 13A. The P2-E1 element in both 2D and 3D satisfies the solvability condition since the arguments of §B2 also apply here as pressures are located midway between vertical velocities along all vertical edges.

Appendix B: The Inf-Sup and Solvability Conditions

Recalling (47), the 2D Newton-Raphson iteration for saddle point problems is written as

$$\begin{bmatrix} \mathbf{M}_{UU}\left(u^K,w^K\right) & \mathbf{M}_{UW}\left(u^K,w^K\right) & \mathbf{M}_{UP} \\ \mathbf{M}_{UW}^T\left(u^K,w^K\right) & \mathbf{M}_{WW}\left(u^K,w^K\right) & \mathbf{M}_{WP} \\ \mathbf{M}_{UP}^T & \mathbf{M}_{WP}^T & 0 \end{bmatrix} \begin{Bmatrix} \Delta u \\ \Delta w \\ \Delta p \end{Bmatrix} = \begin{Bmatrix} -R_U\left(u^K,w^K,p^K\right) \\ -R_W\left(u^K,w^K,p^K\right) \\ 0 \end{Bmatrix}. \quad (71)$$

In the 3D case the vector Δu may stand for the vector of horizontal variable increments, i.e., $\left\{\Delta u, \Delta v\right\}^T$. For simplicity, dropping the Δ's and renaming quantities, (71) is put in the standard form (54), as follows

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & 0 \end{bmatrix} \begin{Bmatrix} \mathbf{u} \\ p \end{Bmatrix} = \begin{Bmatrix} \mathbf{f} \\ g \end{Bmatrix}, \quad (72)$$

where $\mathbf{u} = \left\{u, w\right\}^T$, $\mathbf{A} = \begin{bmatrix} \mathbf{M}_{UU}\left(u^K,w^K\right) & \mathbf{M}_{UW}\left(u^K,w^K\right) \\ \mathbf{M}_{UW}^T\left(u^K,w^K\right) & \mathbf{M}_{WW}\left(u^K,w^K\right) \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} \mathbf{M}_{UP}^T & \mathbf{M}_{WP}^T \end{bmatrix}$, etc.,

and the correspondences of the other variables should be obvious.

As noted earlier, matrix $\mathbf{A}$ is invertible. Eliminating $\mathbf{u}$ in (72), we obtain an equation for the pressure,

$$\left(\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T\right) p = \left(\mathbf{B}\mathbf{A}^{-1}\right)\mathbf{f}, \quad (73)$$

where $\mathbf{S} = \left(\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T\right)$ is called the Schur complement with respect to p. This equation may be solved for p provided an inf-sup condition on matrix $\mathbf{B}^T$ is satisfied (Benzi et al., 2005; Gerbeau, 2025). Having solved for p, the equation to be solved for velocity is

$$\mathbf{A}\mathbf{u} = \mathbf{f} - \mathbf{B}^T p. \quad (74)$$

This is generally not attractive as a practical solution method since the Schur complement matrix $\mathbf{S}$ is full and therefore (73) too expensive to solve. However, Gerbeau (2025) points out that the standard inf-sup condition for $\mathbf{B}^T$ is equivalent to the much simpler "algebraic inf-sup condition",

$$Ker\mathbf{B}^T = \left\{0\right\}, \quad (75)$$

i.e., the inf-sup condition is the same as requiring that the null space of $\mathbf{B}^T$ be the zero vector. Further, this means that the inf-sup condition or (75) implies that the pressure solution must not contain spurious modes. Finally, this suggests that existence of a pressure null space is the culprit in the need to satisfy the inf-sup condition.

On the other hand, consider the situation when the Newton-Raphson system (71) is solved with a mesh satisfying the solvability condition. Simplifying notation, the linear system (71) may be rewritten as follows

$$\begin{bmatrix} \mathbf{M}_{UU} & \mathbf{M}_{UW} & \mathbf{M}_{UP} \\ \mathbf{M}_{UW}^T & \mathbf{M}_{WW} & \mathbf{M}_{WP} \\ \mathbf{M}_{UP}^T & \mathbf{M}_{WP}^T & 0 \end{bmatrix} \begin{Bmatrix} u \\ w \\ p \end{Bmatrix} = \begin{Bmatrix} f_U \\ f_W \\ 0 \end{Bmatrix}. \tag{76}$$

If the solvability condition applies, the matrices $\mathbf{M}_{WP}, \mathbf{M}_{WP}^T$ are invertible, as before, and the continuity equation, the third equation from (76), may be solved for the vertical velocity w as before, as follows

$$w = -\mathbf{M}_{WP}^{-T}\mathbf{M}_{UP}^T u\,, \tag{77}$$

corresponding to (56). Recall from §4.3.2 that using (56) in functional (59) one obtains (60), an equation for u only. With a bit of algebra, and now using (77) in (76), the corresponding linear equation for u is given by

$$\begin{aligned}\left(\mathbf{M}_{UU} - \left(\mathbf{M}_{UW} - \mathbf{M}_{UP}\mathbf{M}_{WP}^{-1}\mathbf{M}_{WW}\right)\mathbf{M}_{WP}^{-T}\mathbf{M}_{UP}^T - \mathbf{M}_{UP}\mathbf{M}_{WP}^{-1}\mathbf{M}_{UW}^T\right)u \\ = \left(f_U - \mathbf{M}_{UP}\mathbf{M}_{WP}^{-1} f_W\right).\end{aligned} \tag{78}$$

However, just as was the case with (73), the associated matrix is full and too expensive to solve in practice. Nevertheless, with u given by (78), one obtains the following matrix equation from the second equation of (76), to be solved for the pressure p:

$$\mathbf{M}_{WP}\, p = \left(\mathbf{M}_{WW}\mathbf{M}_{WP}^{-T}\mathbf{M}_{UP}^T - \mathbf{M}_{UW}^T\right)u + f_W\,. \tag{79}$$

Thus, corresponding to equations (74) and (73), one obtains (78) and (79), but now to be solved in reverse order, first for velocity u, then for the pressure p. The essential difference between the equations for the pressure, (73) and (79), is that the Schur complement matrix $\mathbf{S}$ requires the inf-sup condition to be satisfied, while there is no such requirement for $\mathbf{M}_{WP}$ since it is an invertible matrix. Thus, pressure solutions will be well behaved with no spurious pressure modes, although there is no guarantee or expectation of smoothness. This implies that the solvability condition (58), and more importantly, the invertibility of matrix $\mathbf{M}_{WP}^{T}$, is sufficient for the well posedness of the Stokes saddle point problem.

Appendix C: Test Problems

We use three 2D test problems to demonstrate the new methods. The geometrical configuration of the three problem meshes is illustrated in Fig. C1. The first problem, named Test B, is actually Exp. B from the ISMIP-HOM benchmark suite (Pattyn et al., 2008); it features a no-slip condition on a sinusoidal basal surface. The second problem, Test D*, incorporates sinusoidal friction along a uniformly sloped plane basal surface. This is a modification of Exp. D from the benchmark suite with increased friction since ice flow in Exp. D is very nearly vertically constant (see Fig. 4), which is characteristic of the shallow-shelf regime and not representative of a more typical flow regime.

A third problem, Test O (for "Obstacle") is introduced to illustrate the adaptive switching discussed in §6.1. As illustrated in Fig. C1, Test O has a unique feature, a thin no-slip obstacle, located at $x = 4\ km$ and extending vertically $200\ m$ from the bed (20 % of the ice sheet thickness), which forces ice flow near the obstacle to adjust abruptly. However, because of the no-slip boundary conditions along the obstacle surface, the triangular element in the lee of the obstacle, with one vertical edge and one edge along the bed, would result in zero vertex velocities. This would imply zero stress and therefore a singularity in ice viscosity. To avoid this, elements at the back of the obstacle are "reversed" compared to the ones at the front of the obstacle as shown in Fig. C1.

All tests feature a sloping flat upper surface, given by $z_s(x) = -x\,\mathrm{Tan}(\theta)$, where $\theta = 0.5^{\circ}$ for Tests B and O, and $\theta = 0.3^{\circ}$ for Test D* (this differs from the 0.1° slope in Test D), with free-stress boundary conditions. The bottom surface elevation in Test B is given by $z_b(x) = z_s(x) - H_0 + H_1\,\mathrm{Sin}(\omega x)$, where the $H_0 = 1000\ m$, $H_1 = 500\ m$, $\omega = 2\pi/L$, and L is the perturbation wavelength or domain length. The bottom surface elevation in Tests D* and O is $z_b(x) = z_s(x) - H_0$, which is parallel to the upper surface. Problem length L in the ISMIP-HOM suite ranges from $5\ km$ to $160\ km$ but here we consider only three cases at the high end of the aspect ratio (H_0/L) range, namely

$L = 2, 5, 10\ km$, where the inaccuracy of the Blatter-Pattyn approximation becomes noticeable. Lateral boundary conditions are periodic. The spatially varying friction coefficient for Test D* is $\beta(x) = \beta_0 + \beta_1 \mathrm{Sin}(\omega x)$, where $\beta_0 = \beta_1 = 10^4\ Pa\ a\ m^{-1}$ (this is an order of magnitude larger than in Test D). Physical parameters are the same as in ISMIP-HOM, i.e., ice-flow parameter $A = 10^{-16}\ Pa^{-3} a^{-1}$, ice density $\rho = 910\ kg\, m^{-3}$, and gravitational constant $g = 9.81\, ms^2$. In general, units are MKS, except usually time is per annum, convertible to per second by the factor $3.1557 \times 10^7\ s\ a^{-1}$.

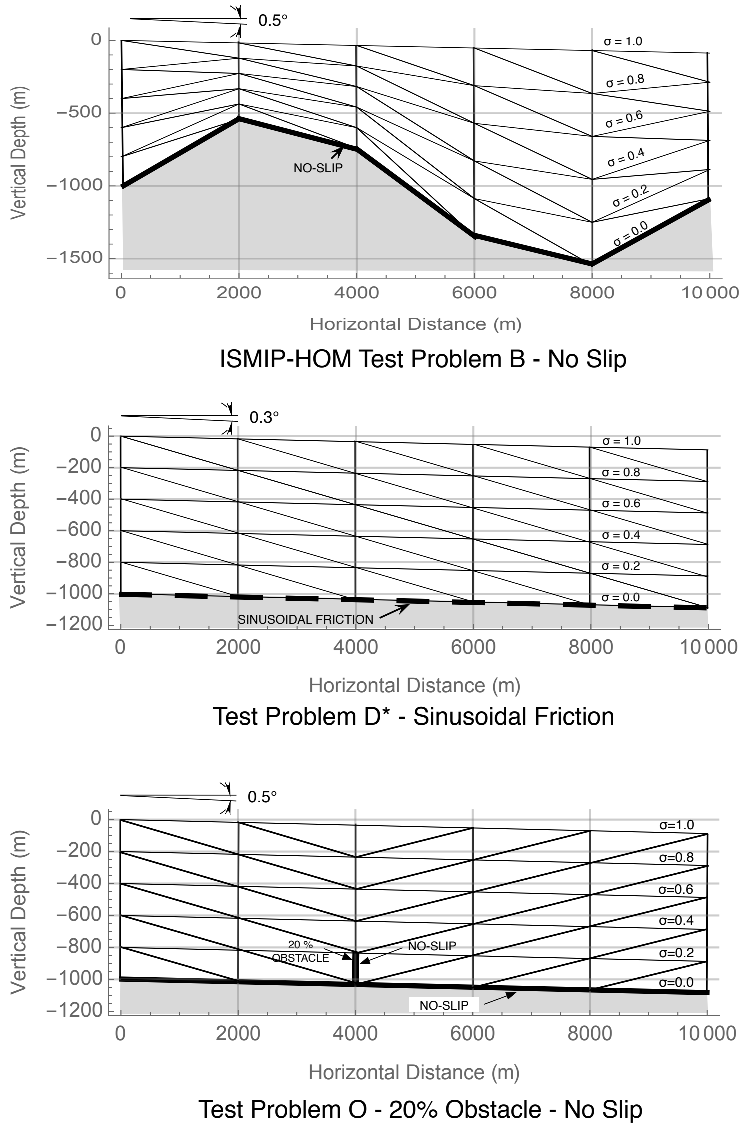

Figure C1. Test problem meshes. For clarity, a coarse 5x5 configuration is shown.