A NOVEL TRANSFORMATION OF THE ICE SHEET STOKES EQUATIONS AND SOME OF ITS PROPERTIES AND APPLICATIONS

by J. Dukowicz

Re-review by Christian Schoof, University of British Columbia

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 reviewers
I 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.

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.

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 discretization.
   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 constrained 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 solutions.

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

$$V = \{v \in [W^{1+1/n}(\Omega)]^3 : \mathrm{div} v = 0 \,\mathrm{and}\, v = 0 \,\mathrm{on}\, \partial\Omega_{b_1},\, v \cdot n = 0 \,\mathrm{on}\, \partial\Omega_{b_2}\},$$

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)

2. The Herterich ("BP+") model. I flagged in my 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:

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

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 $1 : (2\pi)$)

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.

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

- 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

- 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.)

- 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.

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

- line 287 "remains positive" $\rightarrow$ "remains non-negative"

- 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.

- eq (29) — this is a bit misleading; your earlier text (line 289, "The dummy variables ξ, ˆ 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 $\xi = \hat{\xi} = 0$.

- 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 $\beta(x)\sqrt{1 + (\partial b/\partial x)^2}$.

- 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)

- 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 $\tilde{P}$.

- line 455 ”in matrix form”: misleading as $\mathcal{M}$ hasn’t been cast in matrix form

- line 459 ”$\mathcal{M}(u, w)$ 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 $\mathcal{M}(u, w) > 0$ and $\mathcal{M}$ was linear, then it would follow that $\mathcal{M}(-u, -w) < 0$ by linearity, no? Do you mean ”convex”?

- 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.

- equation (48) It’s probably a bad idea to use the same letter M for the dissipation potential part of the discretized functional $\mathcal{A}$, 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.

- line 519 onewards "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 $h \to 0$. 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* $h \to 0$. No?
- 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?
- 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 $A\mathbf{u} = \mathbf{f}$, for which the statement in question is true.
- 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 defined 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

- 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?

- 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.

- line 575 "$n_p + \lambda_z + \Lambda = n_w$" should presumably be "$n_p + n_{\lambda_z} + n_\Lambda = n_w$"?

- 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 $V = \{v \in [W^{1+1/n}(\Omega)]^3 : \mathrm{div} v = 0 \,\mathrm{and}\, v = 0 \,\mathrm{on}\, \partial\Omega_{b_1},\, v{\cdot}n = 0 \,\mathrm{on}\, \partial\Omega_{b_2}\}$, 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.

- 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 $\mathbf{u}$ such that $B^T\mathbf{u} = 0$. 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)

- 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 $V = \{v \in [W^{1+1/n}(\Omega)]^3 : \mathrm{div} v = 0 \,\mathrm{and}\, v = 0 \,\mathrm{on}\, \partial\Omega_{b_1},\, v \cdot n = 0 \,\mathrm{on}\, \partial\Omega_{b_2}\}$ 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 $B^T \mathbf{u} = 0$, 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.

- 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.

- line 702 "It would be computationally cheaper to use the standard Blatter-Pattyn approximation (ξ, $\hat{\xi} = 0$) 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 error relative to a Stokes flow solver, so it is important to know whether the reduction in computational cost is worth the increased error.

- 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

- 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.

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

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

- 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?

- 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?

- 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.