Cedric Thieulot
Department of Earth Sciences
Mantle dynamics and Geophysics group
Vening Meineszgebouw A
Princetonlaan 8a
3584 CB Utrecht

Wolfgang Bangerth
Department of Mathematics
Department of Geosciences
Campus Delivery 1874
Colorado State University
Fort Collins, CO 80524
bangerth@colostate.edu

Dear colleagues:
Attached please find a revised version of our manuscript entitled “On the choice of finite element for applications in geodynamics”.

In revising the paper, we have taken the suggestions by the reviewer into careful account – we would like to express our thanks to the reviewers for the careful reading and the useful suggestions. We hope that we have addressed the concerns raised in the reviews to everyone’s satisfaction. A detailed discussion of our changes can be found on the next pages. We also attach a version of the paper that has all changes made during the revision highlighted, to make the comparison easier.

Sincerely,
Cedric Thieulot
Wolfgang Bangerth

Reply to reviewers

The reviewers made many good suggestions about the content of the paper, for which we are genuinely grateful. (Peer review works!) In the following, let us provide answers to all comments raised by the reviewers, omitting typos and grammar issues pointed out in review that we have fixed in the manuscript and that do not require discussion. Questions are shown in blue, with our answers in black.
This re-submission includes a version of the manuscript in which all changes from the previous version are highlighted in red/blue, for easier comparison.

Reviewer #1

This manuscript presents a detailed and thorough technical study on the choice of finite element types in geodynamical simulations. In this study, the authors investigate how the choice of the element type affects the accuracy and the convergence behaviour of iterative solvers typically used in such simulations. This study is limited to 2D quadrilateral elements that are commonly used in geodynamical codes, namely the Q1P0, Q1Q1, Q2Q1 and the Q2P-1 element. Based on three analytical benchmarks and two representative model setups, the authors demonstrate that the Q2Q1 and Q2P-1 are most likely suited best for geodynamical simulations. These conclusions are not entirely new, as other studies have also mentioned issues with these element types before, but it is the first study that quantitatively assesses them. As such, this study certainly deserves publication in Solid Earth.

The manuscript is well structured. The authors first provide an overview about applications of the finite element method in geodynamical simulations, then provide a theoretical overview on error estimates, where they outline theoretically expected results. Afterwards, they proceed to test the different element types using 5 different test cases, three of them being analytical benchmarks. For the analytical benchmarks, the convergence rates are determined and in some cases the effect of the element choice on the number of iterations for the iterative Stokes solver is determined. Finally, two scenarios without an analytical solution are presented.

The manuscript reads very well. The figures are appropriate. In some figures, annotations and colorbars are too small to read. I only have some other minor comments. In my opinion, the manuscript should therefore be published after minor revisions.

Comments:

In the first two paragraphs of the introduction, the code SLIM3D (Popov et al., 2008) should also be mentioned. In this paper, the authors also discuss issues with locking.

We have added this reference.

In the paragraph on the usage of the Taylor-Hood element in geodynamic codes is in my opinion too focused on ASPECT. There are other geodynamic codes that have routinely employed this element that should also be mentioned here. Otherwise the paragraph leaves the impression as if ASPECT is the only code in the geodynamic community to use a Taylor-Hood element. Other codes that have employed the Taylor-Hood element are (note that this list may not be complete):

- MILAMIN (Dabrowski et al. ,2008), which only uses triangular elements
- MILAMIN_VEP (e.g. Thielmann & Kaus, 2012), where MILAMIN was extended to use also quadrilateral elements
- an early version of LaMEM (e.g. Lechmann et al.,2011), which has now moved to an FD formulation.
- pTatin3D (May et al., 2014)

Popov, A., and S. Sobolev (2008), SLIM3D: A tool for three-dimensional thermomechanical modeling of lithospheric deformation with elasto-visco-plastic rheology, Physics of the Earth and Planetary Interiors, 171, 55–75.

Dabrowski, M., M. Krotkiewski, and D. Schmid (2008), MILAMIN: MATLAB-based finite element method solver for large problems, Geochemistry Geophysics Geosystems, 9, Q04030, doi:10.1029/2007GC001719.

Thielmann, M., and B. J. P. Kaus (2012), Shear heating induced lithospheric-scale localization: Does it result in subduction? Earth Plan. Sc. Lett., 359-360, 1–13, doi:10.1016/j.epsl.2012.10.002.

Lechmann, S. M., D. A. May, B. J. P. Kaus, and S. M. Schmalholz (2011), Comparing thin-sheet models with 3-D multilayer models for continental collision, Geophysical Journal International, 187, 10–33, doi:10.1111/j.1365-246X.2011.05164.x.

May, D. A., J. Brown, and L. Le Pourhiet (2014), pTatin3D: High-Performance Methods for Long-Term Lithospheric Dynamics.

May et al. was already cited since it relies on the $Q_2 x P_{-1}$ element but we prefer the 2015 article which is much more extensive than the 2014 one. MILAMIN_VEP and Lechmann et al. rely on Q,xP_{-1} elements and were added. The MILAMIN code needs not be cited here as it relies on the Crouzeix-Raviart element $P_2^+ x P_{-1}$ but we do mention it at the end of the manuscript when we mention the Crouzeix-Raviart element in the discussion.

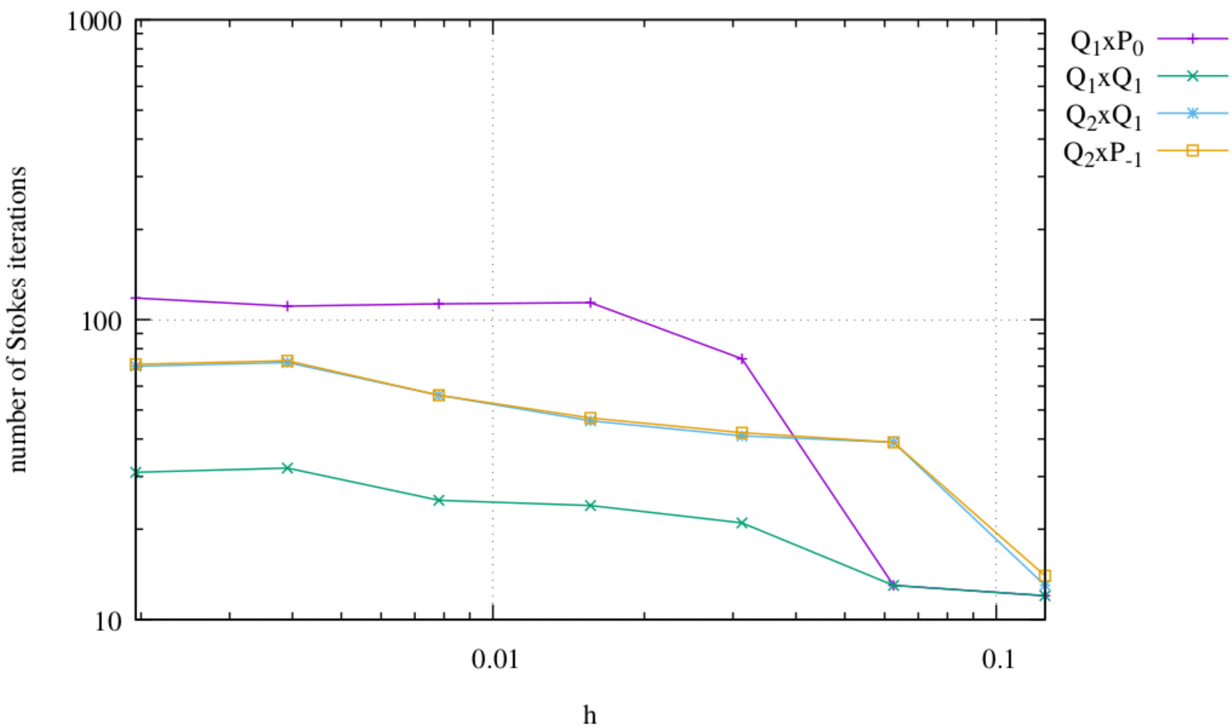
In Figure 1, the colorbars and annotations should be made larger, they are very hard to read. This is true for all figures where fields are shown.

We have redone every figure where fields are shown and the colorbars have been made substantially larger.

Section 5.2:

In section 5.1, both the convergence rates and the number of FGMRES iterations are shown for the Donea & Huerta benchmark. Is there a specific reason as to why the number of FGMRES iterations is not shown in this section? Otherwise I would suggest to also show a diagram similar to Figure 3 (right column).

We had these numbers in our log files but had not thought about visualizing them. The number of iterations for SolCx is shown in the following figure:



This is interesting that it shows a behavior different from the one in the figure for the Donea & Huerta benchmark: The condition number doesn't appear to blow up as h->0 for the Q_1xQ_0 discretization (that is, the problem is not ill-posed in the limit), though it does show that for the SolCx case, the matrix is poorly conditioned. One could speculate why this is so: It may be because for the D&H case, the matrix has the checkerboard pattern as a nullspace, whereas that pattern can not exist/is suppressed for the SolCx case due to the spatially variable coefficient.

Regardless, the Q_1xQ_0 case is still the worst. As a consequence, we think that it does not provide fundamentally different insight than what we had previously shown for the D&H case and we therefore chose not to include it. We think that that may be a debatable choice, and would be happy to include the figure if the reviewer feels strongly about the issue.

Line 326: The statement that in geodynamic applications, discontinuities do not align with mesh interfaces is in my opinion only partly true. In the case of lithospheric-scale models, it may well be possible to align mesh interfaces with discontinuities. I would therefore suggest to write "large-scale geodynamic applications" instead of "geodynamic applications".

We changed the text as suggested.

Line 336: I think the authors meant to write "pure shear" instead of "simple shear"

The reviewer is of course correct and we changed this.

Figure 7: I suggest to add a legend to the figure.

We played with ideas to incorporate what is in the caption into a legend but found no good way to do so without either putting a lot of text into the legend, or to duplicate what is already in the caption. Since we did not wish to clutter the figure which already contains text on the top, left side and bottom, we decided to not add a legend.

Line 471: Although it is not the focus of this paper, the convergence issues related to viscoplastic deformation should not be entirely discarded here. I therefore suggest to slightly extend this discussion and to also add the Spiegelman et al (2016) paper on this issue.

Spiegelman, M., D. A. May, and C. R. Wilson (2016), On the solvability of incompressible Stokes with viscoplastic rheologies in geodynamics, Geochem. Geophys. Geosyst, 17(6), 2213–2238, doi:10.1002/2015GC006228.

We have added the following sentence: “Of course, we are not the first to observe that convergence is hard to come by for these sorts of problems (see for example Spiegelman et al. (2016)) and recent approaches to regularize visco(-elasto)-plastic deformation by Duretz et al. (2020) and Jacquey et al. (2021) have been found to improve the convergence behaviour of the nonlinear solvers.”

In addition, there have been several recent approaches to regularize visco(-elasto)-plastic deformation, which improve the convergence behaviour of the nonlinear solvers. I think this should also be mentioned here:

Duretz, T., R. De Borst, P. Yamato, and L. le Pourhiet (2020), Toward Robust and Predictive Geodynamic Modeling: The Way Forward in Frictional Plasticity, Geophysical Research Letters, 47(5), 20, doi:10.1029/2019GL086027.

Jacquey, A. B., H. Rattez, and M. Veveakis (2021), Strain localization regularization and patterns formation in rate-dependent plastic materials with multiphysics coupling, Journal of the Mechanics and Physics of Solids, doi:10.1016/j.jmps.2021.104422.

Figure 15: Personally, I found the blue and green colors hard to distinguish. In addition, I am not sure whether the red and green color could be distinguished by people with a red-green weakness.

All colour scales employed in the manuscript are obtained from https://www.fabiocrameri.ch/colourmaps/ and these are considered universally readable ("The colour combinations are readable both by colour-vision deficient and colour-blind people, and even when printed in black&white"). We had also strived to use the same colors for the same elements throughout the entire paper. As a consequence, we have not changed this one figure, but we have made the subfigures larger so that the blue and green regions are better visible.

Line 528: I suggest to write "robust mantle convection and crustal dynamics simulations that employ finite elements" instead of " robust mantle convection and crustal dynamics" simulations.

So changed.

Line 540: MILAMIN should also be mentioned here (see reference above).

We added the reference.

Reviewer #2

Review of: "On the choice of finite element for applications in geodynamics" Authors: Cedric Thieulot and Wolfgang Bangerth

Submitted to: Solid Earth Discussions

Reviewer: Dave A. May

1 Summary

In this manuscript the authors discuss and compare several mixed finite-elements for solving the incompressible Stokes problem in the context of geodynamic applications. The study focuses on finite-element spaces which have been traditionally used (i.e. implicitly advocated) in the field of geodynamics ($Q_1 - P_0$) or those adopted in more recent studies ($Q_1 - Q_1$ stabilized, $Q_2 - Q_1$, $Q_2 - P_{-1}$). The intention of the study is to elucidate which element pair is "the best at accurately simulating typical geodynamic situations". This point is meaningful for both practitioners and developers of geodynamic software. The evaluation of finite-elements for "typical geodynamic situations" is assessed by examining the solution quality and solver performance for several well known analytic solutions of viscous flow and an idealised model problem.

Overall the rational and design of the evaluation conducted is sound. Furthermore, the conclusions reached are correct. The conclusions do not identify an answer to the question "what is the best element-pair to use for geodynamics simulations?”. Rather, by a process of elimination, the authors identify that the $Q_2 - Q_1$ and the $Q_2 - P_{-1}$ elements are the only suitable candidates (of those elements under examination). No clear advice is provided as to which of these two should be preferred in general, or in specific modelling contexts. It would be helpful if further discussion was provided to elaborate on when $Q_2 - Q_1$ might be preferred over $Q_2 - P_{-1}$ (and vice-versa).

There are two major benefits of the $Q_2 - P_{-1}$ element which make it distinct from $Q_2 - Q_1$ that have not been discussed. These are that only the $Q_2 - P_{-1}$ element: (i) provides local (elementwise) conservation, i.e. $\int_{\Omega e} \nabla \cdot u_h \, dV = 0$, where Ω_e is the domain of element e; (ii) allows the element face geometry to be described by a quadratic (2D) or bi-quadratic (3D) representation without degrading the a priori error estimates (or committing a finite element crime). I detail these points and why they are important in the general comments section below. There is one other technical point about the a priori error estimates related to the $Q_2 - P_{-1}$ element which should be clarified in the revision - I expand upon this in the general comments below.

2 Major comments

2.1 Conservation

Given a domain Ω partitioned into non-overlapping elements Ω_e, such that $\Omega = \cup^{N}_{e=1}\Omega_e$, the discrete solution u_h obtained with $Q_2 - P_{-1}$ element satisifies

$$\int_{\Omega_e} \nabla \cdot u_h \, dV = 0. \quad (1)$$

In contrast, the solution obtained with $Q_2 - Q_1$ only satisfies

$$\int_{\Omega} \nabla \cdot u_h \, dV = 0. \quad (2)$$

That is, the former element ($Q_2 - P_{-1}$) provides *local* conservation (element-wise, Ω_e), whilst the latter element ($Q_2 - Q_1$) only provides *global* conservation (domain-wise, Ω).

The type of the conservation provided by an element (or lack there-of in the case of $Q_1 - Q_1$) is important for the solution quality of buoyancy driven flows. You express this point in your own results when you examine the solution associated with $Q_1 - Q_1$. The point is also true when you discretize $\nabla \cdot u = 0$. The type of conservation property you have places restrictions on the type of transport discretization which can be used if you wished to couple the discrete Stokes flow solution (u_h) with the transport of a material property say χ (representing for example rock-type or lithology),

$$\partial\chi / \partial t + \nabla \cdot (u_h\chi) = 0, \quad (3)$$

or even with the conservation of energy (i.e. evolution of temperature), i.e.

$$\partial T / \partial t + \nabla \cdot (u_h T) = \nabla \cdot (\kappa \nabla T) + Q. \quad (4)$$

I refer to Dawson et al. (2004) for an in-depth discussion of compatible transport schemes. By way of illustrating the point, based on Dawson et al. (2004), the u_h obtained from $Q_2 - P_{-1}$ could

be used to solve (4) with a finite-volume (FV) scheme, SUPG or the entropy viscosity method. In contrast, since the u_h obtained from $Q_2 - Q_1$ does not satisfy a local conservation property, you cannot use FV to solve (4), however usage of SUPG and or the entropy viscosity method would be valid.
Given the ubiquitous nature of including equations such as Eqs. (3) and (4) in geodynamic modelling, the type of conservation you obtain from a given mixed finite-element type is important to highlight and discuss.

Indeed, we should have mentioned this issue. We have added a paragraph to the conclusions section about local conservation and the impact on the advection discretization.
We considered adding a larger section about this, but in the end refrained because we thought that the issue is too large to be addressed in a meaningful way in the paper as it stands. Specifically, issues one would have to consider include that "real" geodynamic models are actually compressible, and one would have to investigate how that affects what one can or cannot do with the resulting velocity field when it comes to the advection equation. Moreover, even for incompressible models, one can relatively easily postprocess a velocity field that is not locally conservative into one that actually is. Given these considerations, we believe that we cannot provide a comprehensive discussion of the pros and cons of locally conservative schemes within the constraints of the current paper, and that the added paragraph in the conclusions needs to suffice.

2.2 Error estimates for $Q_2 - P_{-1}$

The error estimates you have stated in equation (4) (in the submitted manuscript) do not apply in general for the $Q_2 - P_{-1}$ element pair. I refer you to Boffi & Gastaldi (2002) and Matthies & Tobiska (2002) for further details. The P_{-1} function space has at least two possible representations, either it is expressed in the global coordinates (x, y, z), or in the element-local coordinates (ξ, η, ζ) - the latter referred to as the "mapped coordinates" in Boffi and Matthies' papers. When defining $Q_2{-}P_{-1}$ spaces on non-coordinate aligned meshes, the "mapped" representation of the P_{-1} space will result in sub-optimal convergence with respect to your estimate in equation (4).
This point does not affect any of the results you have presented in this submission, but it is important to be aware of in general as any practitioner who follows your study and attempts to extend the results to a more general mesh may find that equation (4) is not valid.

That is an interesting observation we were not aware of. The element we use is of the "mapped" kind, but as you correctly point out, this does not matter for the experiments we have performed. We added a footnote to the introduction that mentions the distinction.

2.3 High-order geometry

When using $Q_2 - Q_1$ in spatial dimensions d, the only representation of the element geometry you can use is Q_1. Hence the geometry of your element face must be defined by a Q_1 space in $d - 1$ dimensions. The arguments for why this is true are similar to those discussed in Boffi & Gastaldi (2002) and Matthies & Tobiska (2002). It is a disappointing reality that when using a mixed element with two continuous spaces (Q_k and Q_{k-1}) the geometric representation of the

element is limited to Q_{k-1}. In practice this means that any Lagrangian or ALE formulations you might wish to use with $Q_2 - Q_1$ need to respect this geometric restriction.
In contrast, if you use Q_2-P_{-1} with the P_{-1} space represented in the global coordinates, you do not have this geometric restriction and your element geometry can be defined in Q_2. The reason this is valid is because you only have one iso-parametric mapping in your mixed finite-element space (i.e. that related to velocity).
Geometric flexibility and the ability to model curved surfaces is of importance for providing high accuracy representations of topography in regional models, and also to facilitate more accurate approximations of the sphere in global models (or regional cap type models). (I appreciate that approximating the sphere by piece wise Q_k patches is not the only way to achieve a spherical or spherical cap model.)

The arguments above are, we believe, only partly true. One *can* use higher order geometries (for example, ASPECT by default uses a Q4 geometry description regardless of the finite element used; in other contexts, we use exact geometries rather than polynomial approximations). Higher-order geometry approximations in those cases simply do not provide better convergence orders than lower-order approximations. But, better geometry descriptions might still (and often do) provide smaller errors, i.e., they affect the constant in error estimates.

We do not disagree with the statements above, but think that their relevance is not sufficient to make the paper more complex. After all, having concluded that the Taylor-Hood-type elements are superior for rectangular geometries is not negated by the observations about curved geometries: The worst that could happen is that the advantage is not quite as large on other geometries; the lower-order geometry approximation mentioned above is no worse than what has to be used for the Q1-Q1 or Q1-P0 elements.

2.4 Solution regularity

In Sec 3.2 you discuss solution regularity and equations (9) and (10) introduce new error estimates. I am not completely convinced these (equations / error estimates) currently add a lot to the paper. I like the discussion and it is certainly valid, however currently the content and message is not (in my opinion) well connected with the remainder of the paper. For instance, after Sec 3.2, solution regularity is not discussed again in the context of any of the experimental results, and is only ever mentioned in the conclusion where it asserts we expect a lack of regularity in typical geodynamic scenarios (without further explanation).
The lack of a connection between solution regularity and the numerical results has the potential to lead to some confusion. For example, based on the introduction of different error estimates, and the order of accuracy reported for SolCx vs SolVi, readers may believe that SolCx possess sufficient regularity (i.e. $q = 2$) since estimates (4) and (8) are satisfied for Q_2-P_{-1}, whilst the SolVi solution lacks regularity (as estimate (4), (8) are violated for Q_2-P_{-1}). Hence, the reader may refer to the estimate in (9) and the discussion of solution regularity to try and understand why SolCx and SolVi differ in the obtained order of accuracy. Of course if you solve SolCx using $Q_2 - P_{-1}$ with a mesh in which the element edges don't align with the viscosity jump, then the order of accuracy in estimate (4) and (8) is not valid. The sub-optimal convergence observed in this case has nothing to do with the regularity of the true solution u*,* p as the physical problem is unchanged, rather all that has changed is the discretization (the mesh) and the resulting discrete

solutions u_h, p_h. Hence a reader trying to understand the sub-optimal convergence in the context of SolCx vs. SolVi is not going to understand the observation from thinking about solution regularity.
I think what could clarify all of this is: (i) extend Sec 3 such that independent of solution regularity, it discusses under what situations the order of accuracy drops to $h^{1/2}$, and why this occurs; (ii) add additional experiments for SolCx which consider the case when the mesh elements are not aligned with the jump in viscosity.

First, see also our answer to comment #8 below for a discussion of where the $h^{1/2}$ convergence rate might be explained theoretically.
Second, we have mentioned in an added paragraph at the end of Section 3 the issue of actually observing reduced error rates. This turns out to be quite difficult in practice. It often requires using scrambled meshes to really observe the *minimal theoretically guaranteed rate*. The SolCx case mentioned in the comment above is just one example of this: symmetry in mesh, specifics of the element under consideration, and many other factors sometimes (often?) produce a rate that is better than the minimum the estimates guarantee for a given regularity.

3 Comments / corrections

1. [lines. 50-55] The wording "..in which the pressure is discontinuous and of (total) polynomial degree $k - 1$, but missing the shape functions that distinguish the space Q_k on quadrilaterals..." is not clear (and actually misleading). It is more precise to talk about the underlying basis and not refer to shape-functions. P_{-1} has a basis of $\{1, \xi, \eta\}$ (or $\{1, x, y\}$) whilst the basis for Q_1 is $\{1, \xi, \eta, \xi\eta\}$ (or $\{1, x, y, xy\}$). You cannot define the shape-functions for P_{-1} (in practice or in code) by simply removing shape-functions associated with your Q_1 implementation.

Fair enough. We have made the corresponding change in the manuscript.

2. [lines. 50-55] You need to add a reference for the method mentioned here "Another variation is to enrich the pressure space by a constant shape function on each cell."

Yes, good point. This turned out to be a bit more work because it seems difficult to track down where exactly this element was first proposed. We *think* that it is in Gresho et al. in a conference proceedings in 1980, but that does not appear to be available electronically. We think we have found a good reference in Boffi et al. (2011).

3. [pg. 23] Above line 475, you wrote "Our interpretation of this experiment is that the inability of the $Q_1 \times Q1$ element ...". The last "1" next to "Q" should appear as a subscript.

Fixed.

4. [Fig. (3)] Stating "Number of FGMRES solver iterations as a function of the mesh size h" is only meaningful if we know what preconditioner was used for the Stokes problem. Without a preconditioner the iteration count will always increase as h decreases. I didn't find in the text or caption any statement (or reference to the preconditioner used in ASPECT) specifying that you are preconditioning FGMRES and what preconditioner you used.

That's a fair point, and we have added the reference to the preconditioner used as requested. That said, the point of the figure is not to show the *absolute* number of iterations, but how it *changes* – or doesn't – with the mesh size for the different discretizations. That fact that the Q1-Q0 element leads to matrices whose conditioning clearly blows up is a reflection of the instability of the element, as described in the text.

5. [lines. 175-185] The norm $||\nabla(u - u_h)||_2$ is the H^1 semi-norm as opposed to the H^1 norm given by

$$H^1(\Omega) := \{u : \Omega \to R | u, \nabla u \in L_2(\Omega)\} .$$

You haven't explicitly defined what H^k is (I think you are using the semi-norm), could you please either define it with words or an equation to avoid confusion.

The statements given in the text are not actually wrong. It is true that the first of the norms in eq. (8) is a *seminorm* of the velocity (though it is a norm of the gradient of the velocity), but both the statements in (8) and the discussion following are correct.
We have added a reference and an informal definition of these function spaces in a footnote.

6. The font size of the palette labels, tick numbers in Figures 1, 4, 6, 8, 11, 13 is too small and should be increased for improved legibility.

This comment was also made by reviewer #1 and we have redone all figures accordingly.

7. In several places the writing infers that elements $Q_2 - Q_1$ and $Q_2 - P_{-1}$ are part of the same finite element family (referred to as "Taylor-Hood") – they are not. The elements are distinct and the writing should reflect this point. Some instances I came across: line 205, ..."the Taylor-Hood elements..."; line 4, "or more recently the stable Taylor-Hood element with . . . discontinuous ($Q_2 \times P_{-1}$) pressure". Taylor-Hood is to be understood as mixed elements given by function spaces $Q_k - Q_{k-1}$ (quads/hexes) and $P_k - P_{k-1}$ (triangles/tets) for $k \geq 2$.

Strictly speaking, the term "Taylor-Hood" should probably only apply to quadrilaterals, and even there only to the Q2-Q1 case (not higher orders, or 3d), as that is what Taylor and Hood proposed in their 1973 paper. In fact, not even that is true: The original proposition was to use the 8-node serendipity element for the velocity, rather than the full Q2 element. Yet, even the first paper showing stability of the cubic-quadratic case (Franco Brezzi and Richard S. Falk, "Stability of Higher-Order Hood-Taylor Methods", SIAM Journal on Numerical Analysis, Vol. 28, No. 3 (Jun., 1991), pp. 581-590) did not find it necessary to state that that work was on triangles whereas the original paper was for quadrilaterals – they just referred to the triangular case by the term "Taylor-Hood", in the title of the paper nonetheless.

This observation underlines the fact that the *common use* of the term "Taylor-Hood" is "an element for which the velocity has one higher polynomial degree than the pressure". As a consequence, both element families are referred to as "Taylor-Hood" or "Taylor-Hood-like" in street language, and we do not wish to make a distinction where most scientists would not.

We have, however, made this point clear in a footnote in the introduction and a comment after introducing the variations of the element.

8. I think Section 3 would be more complete and improved if it also included a discussion (and references) which also cover the case when the order of accuracy drops to $h^{1/2}$. This, in addition to the comments about solution regularity, would further re-enforce why high-order (here meaning $k > 2$) may not be useful in geodynamic modelling contexts.

We don't actually know why we get the specific rate $h^{1/2}$. Presumably, this is related to the regularity q of the solution, but as mentioned in this section, determining the regularity of a solution is a mathematically rather difficult problem which we are not equipped to undertake (nor would it be useful to the readers of this paper). More generally, however, it is not very difficult to *construct* solutions with a specific regularity. One way to do that is to consider the *L*-shaped domain and play with the opening angle since the opening angle determines the strength of the singularity of the solution at the re-entrant corner. Another approach is to consider a square domain and consider the Laplace equation with a coefficient that is constant in all four quadrants of the square, but discontinuous where these quadrants come together. In that case, the solution will have a singularity at the center of the square, and the strength of the singularity depends on the values of the coefficients; that is, the solution will be the sum of a smooth part and a part of the form r^a where r is the distance from the center and a is an exponent that depends on the coefficients. If a is less than some critical value, then the solution will not be in $H^2, H^3, \ldots$ and consequently one should not expect (and will not observe) the optimal convergence rate for elements of polynomial degree k. Instead, if the solution is only in H^s with $s<k+1$ (where s will be related to a), then one is only guaranteed convergence in the energy norm proportional to h^{s-1}, though in practice one often sees something better due to local cancellation, symmetries, alignment of the mesh, etc. Observing the minimal guaranteed convergence order typically requires using a completely unstructured grid, for example.

As a consequence of these sorts of considerations, coming up with a *theory* of why we get convergence rate ½ is not useful, nor in our view feasible with the level of mathematical detail that would be appropriate for this paper and its intended audience. Instead, we have chosen to simply mention at the end of this section that we will observe a reduced convergence order with the cases discussed in Sections 5.2 and 5.3.

(As an aside, however, it is worth noting that PDE theory only guarantees that $u \in H^1$ in the most general case – only if the coefficients and right hand side are sufficiently smooth and the domain is convex may we hope for anything better. If that is not the case, H^1 regularity is all we have for the velocity, and L^2 for the pressure. Inspecting estimate (9) then reveals that we cannot actually expect *any* kind of convergence rate for the gradient of the velocity and for the pressure; going through the derivation of the L^2 convergence of the velocity – a technique that uses duality arguments – will show that the rate provided in the middle formula of (9) will then also not hold. In other words, if the solution really has no regularity beyond the bare minimum, the finite element solution will not converge at any rate h^s at all, though with substantially more work, one can still show that it converges – for example proportional to $log(1/h)$ or something similarly slow.

So why do we observe $h^{1/2}$ for the SolCx at all? It may be that some finer analysis shows that the solution is marginally more smooth than just H^1, for example it may be in $H^{3/2}$. That would be testable because we know the exact solution and can compute which Sobolev spaces it is in. Or it may be that the symmetry of the mesh yields some cancellation in error terms that produces this rate. A 1970s or 1980s numerical analyst interested in these sorts of quite theoretical questions could probably have spent a few weeks or months on finding the precise answer to the question, but little of practical value is gained by it and we are quite sure that the audience for this paper would not appreciate the detail necessary to convey the findings.)