Geochemistry
Geophysics
Geosystems
$G^3$

Published by AGU and the Geochemical Society

*Article*
Volume 14, Number 4
30 April 2013
doi:10.1002/ggge.20078
ISSN: 1525-2027

# An adaptive staggered grid finite difference method for modeling geodynamic Stokes flows with strongly variable viscosity

**T. V. Gerya**
*Institute of Geophysics, ETH Zürich, Zürich, Switzerland*

**D. A. May**
*Institute of Geophysics, ETH Zürich, Zürich, Switzerland (dave.may@erdw.ethz.ch)*

**T. Duretz**
*ISTeP, UMR 7193, UPMC Paris 06, CNRS, Paris, France*

[1] Here we describe a new staggered grid formulation for discretizing incompressible Stokes flow which has been specifically designed for use on adaptive quadtree-type meshes. The key to our new adaptive staggered grid (ASG) stencil is in the form of the stress-conservative finite difference constraints which are enforced at the "hanging" velocity nodes between resolution transitions within the mesh. The new ASG discretization maintains a compact stencil, thus preserving the sparsity within the matrix which both minimizes the computational cost and enables the discrete system to be efficiently solved via sparse direct factorizations or iterative methods. We demonstrate numerically that the ASG stencil (1) is stable and does not produce spurious pressure oscillations across regions of grid refinement, which intersect discontinuous viscosity structures, and (2) possesses the same order of accuracy as the classical nonadaptive staggered grid discretization. Several pragmatic error indicators that are used to drive adaptivity are introduced in order to demonstrate the superior performance of the ASG stencil over traditional nonadaptive grid approaches. Furthermore, to demonstrate the potential of this new methodology, we present geodynamic examples of both lithospheric and planetary scales models.

**Components:** 12,300 words, 12 figures, 3 tables.

**Keywords:** AMR; geodynamics; staggered grid; Stokes flow; variable viscosity.

**Index Terms:** 1952 Modeling; 4255 Numerical modeling; 3225 Numerical approximations and analysis.

**Received** 28 September 2012; **Revised** 30 January 2013; **Accepted** 31 January 2013; **Published** 30 April 2013.

Gerya, T. V., D. A. May, and T. Duretz (2013), An adaptive staggered grid finite difference method for modeling geodynamic Stokes flows with strongly variable viscosity, *Geochem. Geophys. Geosyst.*, *14*, 1200–1225, doi:10.1002/ggge.20078.

## 1. Introduction

### 1.1. Computational Geodynamics

[2] Over geological time scales, the dynamics of the Earth's mantle and lithosphere can be described by conservation equations applicable to highly viscous, creeping fluids. Fluid motion is largely driven by the heat produced from the Earth's core and from radiogenic heat sources released from the Earth's mantle. Given our

present-day understanding of the Earth's rheological and compositional structure with depth, it is apparent that this coupled system is driven by inherently multiscale processes. Considering only the compositional layering, we observe a wide range of relevant length scales. For example, sedimentary and volcanic processes lead to the formation of centimetric to kilometric lithological units, localization of deformation can occur from the millimetric up to the kilometric scale, topographic variations occur on the kilometric scale, tectonic motions involve plates of several thousands of kilometers separated by quasi-discrete plate boundaries, length scale of mantle heterogeneities may be on the order of thousands of kilometers, and processes such as core formation take place at planetary scale.

[3] It is well established in the Earth science community that the use of numerical models to study the long-term geodynamics is a powerful tool to further our understanding of the dynamics of coupled thermomechanical systems. Furthermore, regional (lithospheric) scale numerical models have demonstrated that the inclusion of an upper/lower crust, topography and surface processes can have a profound influence on the dynamics of the lithosphere [*Beaumont et al.*, 2001; *Burov et al.*, 2001; *Gerya et al.*, 2000; *Gorczyk et al.*, 2007]. Although the importance of lithosphere-asthenosphere coupling on geodynamics processes is well accepted [*Burov et al.*, 2001; *van Hunen et al.*, 2000; *Gerya et al.*, 2004], many of these modeling studies focused on the deformation within the lithosphere and replaced part of the asthenosphere with various types of boundary conditions (i.e., Winkler) to mimic the effect of mantle flow. Such approximations were necessary to ensure that the essential physics could be resolved using the currently available numerical methods and computational hardware. Specifically, the aforementioned studies that utilized thermomechanical models which employed structured grids, and consequently, simultaneously resolving the entire mantle-lithosphere, the crust, and the topographic variations in a consistently coupled manner, were computationally intractable.

## 1.2. Discretization Techniques

[4] Across the large time scales associated with geologic processes, the underlying rocks (or material lithology) are subject to severe mixing and stirring. To discretize both the viscous flow equations and simultaneously represent and follow the evolution of such large deformations, computational geodynamists frequently advocate the use of a combined mesh-marker approach [*Weinberg and Schmeling*, 1992; *Poliakov and Podladchikov*, 1992; *Zaleski and Julien*, 1992; *Fullsack*, 1995; *Babeyko et al.*, 2002; *Gerya and Yuen*, 2003, 2007; *Moresi et al.*, 2003, 2007]. In such methods, the flow equations are discretized on the mesh, whilst Lagrangian markers are utilized to discretize the material lithology and history variables.

[5] The use of the classical staggered finite difference (SGFD) scheme [*Harlow and Welch*, 1965] to solve the creeping flow equations (Stokes flow) for lithospheric-scale geodynamic problems has been demonstrated to be both (i) practical, in terms of the computational resources required [e.g., *Gerya and Yuen*, 2003, 2007; *Gerya*, 2010; *Petersen et al.*, 2010], and (ii) reliable, with respect to the quality of the numerical solution [*Duretz et al.*, 2011]. The practicality of the discretization stems from the restriction of a coordinate-aligned, structured grid and the low-order nature of the stencil which is used to discretize velocity (linear) and pressure (constant). The low-order discretization results in second-order accurate solutions for velocity and pressure in the $L_1$ norm when the viscosity is a smooth function and a first-order accurate method (in both velocity and pressure, measured in $L_1$) when large discontinuous viscosity variations intersect the pressure control volume [*Duretz et al.*, 2011]. The latter scenario is of most relevance in both regional and global geodynamic simulations involving mobile lithospheric plates [*Tackley*, 2000; *Stadler et al.*, 2010; *Crameri et al.*, 2012a].

[6] The simplicity of the staggered grid discretization is also arguably its biggest weakness. The ability to deform the mesh, for example, to conform to an evolving topography is not permitted in the classical staggered grid formulations. This shortcoming has been addressed by a number of different methods which permit zero normal stress boundary conditions (or approximate boundary conditions) to be applied on the upper surface [*Harlow and Welch*, 1965; *Matsumoto and Tomoda*, 1983; *Leveque and Li*, 1997; *Fedkiw et al.*, 1999; *Chern and Shur*, 2007; *Suckale et al.*, 2010; *Crameri et al.*, 2012b].

## 1.3. Toward an Adaptive Finite Difference Stencil for Stokes Flow

[7] To incorporate local variations in spatial resolution throughout the computational domain, for

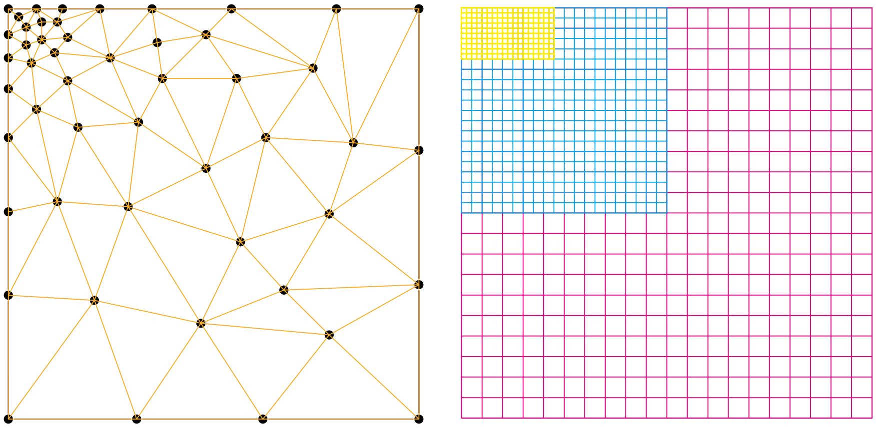

**Figure 1.** (left) Unstructured triangular and (right) block-structured meshes with an increased cell resolution in the upper left corner of the square domain.

example, if we wished to resolve both the mantle and the upper/lower crust, several different types of meshes can be considered. For example, a set of quadrilaterals (hexahedra in 3-D) with an unstructured element connectivity or a set of unstructured triangular (tetrahedra in 3-D) could be employed. Having such versatility within the choice of mesh enables one to (i) easily track the free surface and, if utilizing unstructured meshes, (ii) locally decrease the element size $h$ in regions of interest. Another alternative mesh permitting local mesh refinement is a block-structured mesh which can be conveniently described via a quadtree (octree in 3-D). See Figure 1 for examples of unstructured triangular and block-structured meshes with local refinement.

[8] One of the major arguments to utilize the finite element (FE) method for adaptive mesh calculations is that this discretization naturally permits a large degree of geometric flexibility in defining the mesh. However, the stencil associated with a stable finite element discretization, $Q_2P_1$ (for example) possess far more degrees of freedom than the SGFD stencil. For this reason, much research has focused on generalizing the SGFD methodology to facilitate the use of unstructured adaptive meshes.

[9] Fully unstructured generalizations of the staggered discretization are indeed possible [*Hirt et al.*, 1974; *Rhie and Chow*, 1983; *Reggio and Camarero*, 1986; *Rodi et al.*, 1989] and fall under the class of finite volume methods. Whilst such methods alleviate the limitations of having a coordinate-aligned structured grid, the resulting stencils are larger than the classical SGFD stencil—thus increasing the storage requirements.

[10] To circumvent the cost and programming complexity associated with generalized staggered grid formulations, numerous velocity-pressure arrangements were developed. The most attractive being a co-located (cell-centered or vertex-based) discretization in which both velocity components were defined at the same point in space. Co-located arrangements however suffer from nonphysical oscillations in the pressure field [*Patankar*, 1980]. Extensive comparisons of results obtained with staggered and non-staggered grid arrangements have been conducted [*Perič et al.*, 1988; *Miller and Schmidt*, 1988; *Shih et al.*, 1989; *Armfield*, 1991; *Aksoy and Chen*, 1992; *Melaaen*, 1992a, 1992b; *Choi et al.*, 1994a, 1994b]. Remedies to circumvent the spurious pressure oscillations associated with co-located discretizations have been developed [*Rhie and Chow*, 1983; *Miller and Schmidt*, 1988; *Majumdar*, 1988; *Miller and Schmidt*, 1988; *Papageorgakopoulos et al.*, 2000].

[11] We note that all comparisons conducted between different finite difference stencils and the stabilization techniques developed for co-located formulations focus on constant viscosity, Navier-Stokes equations. The time dependence introduced in the Navier-Stokes equations permits certain time-splitting discretizations (or decoupling) to be employed which further relax the need for properly coupled velocity-pressure discretizations. In geodynamic applications, we specifically require methods which are robust for fluids possessing highly spatially variable viscosity and which are in the steady state Stokes regime (non-inertial). In our experience, the most robust and reliable finite difference discretizations for variable

viscosity Stokes flow are the fully staggered formulations [*Gerya*, 2010].

[12] An alternative approach to using an unstructured staggered grid was the fully adaptive, block-structured orthogonal staggered grid finite difference method of *Albers* [2000] which was developed specifically for studying mantle convection, in which the flow problem possessed smooth variations in viscosity. The methodology developed in this work is simple and maintains a compact stencil. At the transition between different grid resolutions, such meshes produce "hanging" nodes—i.e., the nodes which are only contained within the region of finer resolution. At these hanging nodes, the stencil adopted in *Albers* [2000] employed direct velocity and pressure interpolation. In general, this method cannot be applicable for modeling lithospheric problems with sharply variable viscosity since direct velocity interpolation does not ensure conservation of stresses across resolution boundaries with large viscosity changes. For this common geodynamic modeling situation, a stress-conservative discretization is required [*Gerya and Yuen*, 2003, 2007].

### 1.4. Present Work

[13] To address the need of a low-order, robust, and practical adaptive discretization for computational geodynamics, we extend the methodology presented in *Albers* [2000]. To this end, we develop a fully adaptive, block-structured orthogonal staggered grid finite difference stencil. The crucial component of the method developed here is the use of stress-conservative finite difference schemes across split-cell faces. This proves to be of fundamental importance to eliminate spurious pressure oscillations across cells with split faces. Most importantly, we examine the stability and order of accuracy of the stress-conservative adaptive staggered grid finite difference stencil for a range of viscosity structures.

[14] The outline of the paper is as follows. In section 2, we describe the governing equations of creeping flow and define how to construct staggered grid stencils on quadtree-based adaptive meshes. In section 3, we demonstrate the stability of the adaptive staggered grid discretization and demonstrate the order of accuracy of the method using several analytic solutions for variable viscosity Stokes flow in section 3.2. Furthermore, we highlight the computational advantage obtained using adaptive grids compared to the classical, nonadaptive staggered grids in section 4. Practical examples of lithospheric and planetary scales models are presented in section 5. Lastly, in sections 6 and 7, we summarize the adaptive staggered grid formulation and indicate the future directions and possibilities of this methodology.

## 2. Adaptive Staggered Grid Stencil

### 2.1. Governing Equations

[15] For simplicity in introducing the adaptive staggered grid formulation, here we consider two-dimensional viscous flow problems in a domain $\Omega$. The conservation of mass for an incompressible fluid is given by

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} = 0, \tag{1}$$

where $x, y$ are spatial coordinates and $\boldsymbol{v} = (v_x, v_y)$ are, respectively, the horizontal and vertical velocity components. The 2-D Stokes equations for creeping flow take the form:

$$\frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} - \frac{\partial P}{\partial x} = -\rho(x, y) g_x, \tag{2}$$

$$\frac{\partial \tau_{yx}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} - \frac{\partial P}{\partial y} = -\rho(x, y) g_y, \tag{3}$$

where $\tau_{ij}$ are components of the deviatoric stress tensor, $P$ is the pressure, $\rho(x, y)$ is the density which depends on the spatial coordinates, and $g_x$ and $g_y$ are components of the gravitational acceleration. The deviatoric stress components are given by

$$\tau_{xx} = 2\eta(x, y)\dot{\epsilon}_{xx}, \tag{4}$$

$$\tau_{xy} = 2\eta(x, y)\dot{\epsilon}_{xy}, \tag{5}$$

$$\tau_{yx} = \tau_{xy}, \tag{6}$$

$$\tau_{yy} = 2\eta(x, y)\dot{\epsilon}_{yy}, \tag{7}$$

in which $\eta(x, y)$ is the fluid viscosity and $\dot{\epsilon}_{ij}$ are components of strain rate tensor defined as

$$\dot{\epsilon}_{xx} = \frac{\partial v_x}{\partial x}, \tag{8}$$

$$\dot{\epsilon}_{xy} = \frac{1}{2}\left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}\right), \tag{9}$$

$$\dot{\epsilon}_{yy} = \frac{\partial v_y}{\partial y}. \tag{10}$$

[16] We denote the boundary of the volume $\Omega$ via $\partial\Omega$. The conservation of momentum (equations (2) and (3)) is subjected to a Dirichlet boundary condition

$$\boldsymbol{v} = \boldsymbol{w} \tag{11}$$

along $\partial\Omega_D$ and a Neumann boundary condition

$$(\boldsymbol{\tau} - p\boldsymbol{I})\boldsymbol{n} = \boldsymbol{s} \tag{12}$$

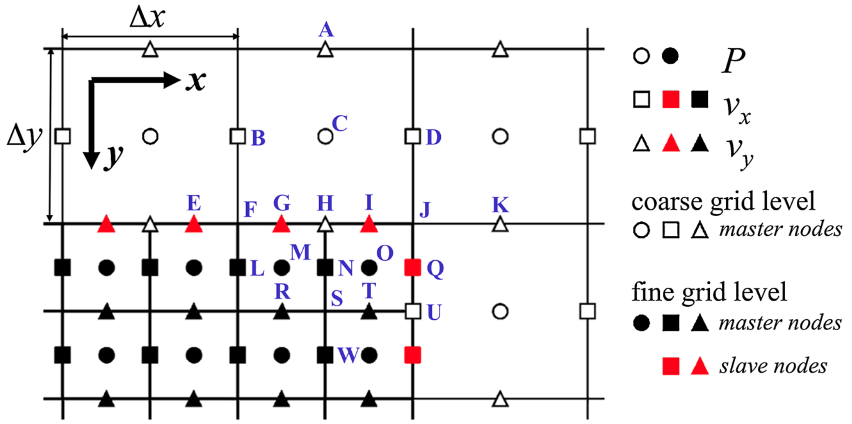

**Figure 2.** Spatial location of horizontal velocity (squares), vertical velocity (triangles), and pressure (circles) on a block-structured, adaptive staggered grid.

along $\partial\Omega_N$, where $\boldsymbol{w}$ is the imposed velocity, $\boldsymbol{s}$ is the imposed stress, $\boldsymbol{n}$ is the outward pointing normal to the boundary $\partial\Omega$, and $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$.

[17] In the remainder of this paper, we consider flow problem in which we either (i) apply Dirichlet conditions along the entire boundary or (ii) we impose a "free slip" condition along $\partial\Omega$ in which we prescribe zero-flow normal to boundary, $\boldsymbol{v}.\boldsymbol{n} = 0$ and zero shear stress tangential to the wall, $\boldsymbol{\tau}\,.\boldsymbol{t} = 0$, where $\boldsymbol{t}$ is the tangent vector to $\partial\Omega$.

[18] With each of these choices for the boundary condition, the pressure is only defined up to an arbitrary constant. In the implementation employed in this work, we define a unique pressure by prescribing four Dirichlet boundary conditions for the pressure in each corner of the domain.

## 2.2. Spatial Discretization

[19] Here we describe in detail the formulation of a 2-D staggered grid stencil suitable for discretizing incompressible Stokes flow on a block-structured, adaptive mesh. Our stencil is specifically designed for block-structured meshes (such as that shown in Figure 1, right) in which transition in resolution occurs via one level of cell bisection. For instance in Figure 1 (right), cells with blue boundaries have side lengths which are half of those of the cells with pink boundaries.

[20] Figure 2 shows the geometry of the adaptive staggered grid used for the formulation of the momentum and continuity equations in 2-D. The main principle of constructing this type of adaptive grid consists of (i) a uniform structure of computational cells at all levels of resolution and (ii) using stress-conservative finite differences both within and between resolution levels. Each computational cell has five staggered nodal points (two $v_x$ nodes, two $v_y$ nodes, and one $P$ node) where various types of equations are formulated. Equations (1)–(3) are discretized at $P$, $v_x$, and $v_y$ master nodes, respectively. Discretization of these equations within the same resolution level is based on stress-conservative finite differences suggested for the uniform rectangular staggered grid [e.g., *Gerya and Yuen*, 2003, 2007; *Gerya*, 2010]. At the interface between changes in the grid resolution, our block-structured adaptive grid formulation introduces "hanging" velocity nodal points (see the red symbols in Figure 2). Across different resolution levels, modified staggered grid stencils are introduced, which utilize values of velocity defined at the hanging nodes. At hanging nodes, we do not discretize the momentum equations, rather we introduce constraints which ensure that mass flux and momentum are conserved. For this reason, we prefer to refer to the hanging nodes as "slave" nodes. Below we explain the modified stencil and define the constraint equations introduced at the slave nodes.

[21] Equation (1) is discretized at the $P$ node of each cell according to the standard finite difference formula (Figure 2):

$$\text{coarse grid level:}\quad \frac{v_x^{\mathrm{D}} - v_x^{\mathrm{B}}}{\Delta x} + \frac{v_y^{\mathrm{H}} - v_y^{\mathrm{A}}}{\Delta y} = 0, \tag{13}$$

$$\text{fine grid level:}\quad \frac{v_x^{\mathrm{N}} - v_x^{\mathrm{L}}}{\frac{1}{2}\Delta x} + \frac{v_y^{\mathrm{R}} - v_y^{\mathrm{G}}}{\frac{1}{2}\Delta y} = 0, \tag{14}$$

where capital letters at $v_x^{\alpha}$ and $v_y^{\beta}$ denote geometrical points (see blue letters in Figure 2) associated with the horizontal and vertical velocity components, respectively. The following example explains the logic of an "across-resolution level" stress-conservative discretization of equation (3) for coarser-level $v_y$ node H (see blue letter H in Figure 2). The discrete form of the $y$ momentum equation at node H is given by

$$\frac{\tau_{xy}^{\mathrm{J}}-\tau_{xy}^{\mathrm{F}}}{\Delta x}+\frac{\tau_{yy}^{\mathrm{N}}-\tau_{yy}^{\mathrm{C}}}{\frac{3}{4}\Delta y}-\frac{P^{\mathrm{N}}-P^{\mathrm{C}}}{\frac{3}{4}\Delta y}=-g_y\rho^{\mathrm{H}}, \tag{15}$$

where superscript letters denote respective geometrical points (see blue letters in Figure 2). The deviatoric stress components and pressure within equation (15) are given by

$$\tau_{yy}^{\mathrm{N}}=\frac{1}{2}\left(\tau_{yy}^{\mathrm{M}}+\tau_{yy}^{\mathrm{O}}\right), \tag{16}$$

$$P^{\mathrm{N}}=\frac{1}{2}\left(P^{\mathrm{M}}+P^{\mathrm{O}}\right), \tag{17}$$

$$\tau_{xy}^{\mathrm{F}}=2\eta^{\mathrm{F}}\dot{\epsilon}_{xy}^{\mathrm{F}}, \tag{18}$$

$$\tau_{xy}^{\mathrm{J}}=2\eta^{\mathrm{J}}\dot{\epsilon}_{xy}^{\mathrm{J}}, \tag{19}$$

$$\tau_{yy}^{\mathrm{C}}=2\eta^{\mathrm{C}}\dot{\epsilon}_{yy}^{\mathrm{C}}, \tag{20}$$

$$\tau_{yy}^{\mathrm{M}}=2\eta^{\mathrm{M}}\dot{\epsilon}_{yy}^{\mathrm{M}}, \tag{21}$$

$$\tau_{yy}^{\mathrm{O}}=2\eta^{\mathrm{O}}\dot{\epsilon}_{yy}^{\mathrm{O}}, \tag{22}$$

and the strain rate components via

$$\dot{\epsilon}_{xy}^{\mathrm{F}}=\frac{1}{2}\left(\frac{v_x^{\mathrm{L}}-v_x^{\mathrm{B}}}{\frac{3}{4}\Delta y}+\frac{v_y^{\mathrm{G}}-v_y^{\mathrm{E}}}{\frac{1}{2}\Delta x}\right), \tag{23}$$

$$\dot{\epsilon}_{xy}^{\mathrm{J}}=\frac{1}{2}\left(\frac{v_x^{\mathrm{U}}-v_x^{\mathrm{D}}}{\Delta y}+\frac{v_y^{K}-v_y^{\mathrm{H}}}{\Delta x}\right), \tag{24}$$

$$\dot{\epsilon}_{yy}^{\mathrm{C}}=\frac{v_y^{\mathrm{H}}-v_y^{\mathrm{A}}}{\Delta y}, \tag{25}$$

$$\dot{\epsilon}_{yy}^{\mathrm{M}}=\frac{v_y^{\mathrm{R}}-v_y^{\mathrm{G}}}{\frac{1}{2}\Delta y}, \tag{26}$$

$$\dot{\epsilon}_{yy}^{\mathrm{O}}=\frac{v_y^{\mathrm{T}}-v_y^{\mathrm{I}}}{\frac{1}{2}\Delta y}. \tag{27}$$

This discretization essentially means that both $\tau_{yy}$ and $P$ for point N are obtained by averaging (interpolating) quantities from the points M and O where they are naturally formulated using standard staggered grid schemes [e.g., *Gerya and Yuen*, 2007; *Gerya*, 2010]. Also we note that the $\tau_{xy}$ shear stress component discretization is naturally based on both coarser- and finer-level velocity nodes. We note that the coarse-level velocity nodes should be used within the $\dot{\epsilon}_{xy}^{\mathrm{J}}$ stencil at the point J where three coarse cells are in contact with one fine cell. This ensures symmetry of the respective shear strain rate stencil, which allows for better accuracy of shear stress representation in a corner separating regions of differing cell resolution. In all other circumstances, preference should be given to finer-level slave nodes (see, e.g., $\dot{\epsilon}_{xy}^{\mathrm{F}}$). Similarly, discretization of equation (2) for finer-level $v_x$ node N (see blue letter N in Figure 2) is the following:

$$\frac{\tau_{xx}^{\mathrm{O}}-\tau_{xx}^{\mathrm{M}}}{\frac{1}{2}\Delta x}+\frac{\tau_{xy}^{\mathrm{S}}-\tau_{xy}^{\mathrm{H}}}{\frac{1}{2}\Delta y}-\frac{P^{\mathrm{O}}-P^{\mathrm{M}}}{\frac{1}{2}\Delta x}=-g_x\rho^{\mathrm{N}}, \tag{28}$$

with

$$\tau_{xy}^{\mathrm{H}}=\frac{1}{2}\left(\tau_{xy}^{\mathrm{F}}+\tau_{xy}^{\mathrm{J}}\right), \tag{29}$$

$$\tau_{xy}^{\mathrm{S}}=2\eta^{\mathrm{S}}\dot{\epsilon}_{xy}^{\mathrm{S}}, \tag{30}$$

$$\tau_{xx}^{\mathrm{M}}=2\eta^{\mathrm{M}}\dot{\epsilon}_{xx}^{\mathrm{M}}, \tag{31}$$

$$\tau_{xx}^{\mathrm{O}}=2\eta^{\mathrm{O}}\dot{\epsilon}_{xx}^{\mathrm{O}}, \tag{32}$$

$$\dot{\epsilon}_{xx}^{\mathrm{M}}=\frac{v_x^{\mathrm{N}}-v_x^{\mathrm{L}}}{\frac{1}{2}\Delta x}, \tag{33}$$

$$\dot{\epsilon}_{xx}^{\mathrm{O}}=\frac{v_x^{\mathrm{Q}}-v_x^{\mathrm{N}}}{\frac{1}{2}\Delta x}, \tag{34}$$

$$\dot{\epsilon}_{xy}^{\mathrm{S}}=\frac{1}{2}\left(\frac{v_x^{\mathrm{W}}-v_x^{\mathrm{N}}}{\frac{1}{2}\Delta y}+\frac{v_y^{\mathrm{T}}-v_y^{\mathrm{R}}}{\frac{1}{2}\Delta x}\right). \tag{35}$$

As before in equation (15), the $\tau_{xy}$ shear stress component at H is obtained by averaging (interpolating) quantities from the points J and F where $\tau_{xy}$ is naturally formulated using standard staggered grid schemes [e.g., *Gerya and Yuen*, 2007; *Gerya*, 2010]. It should also be mentioned that the suggested discretization schemes require that the viscosity $\eta$ is defined at both the four vertices and the centroid of each pressure control volume, whereas the density $\rho$ is defined at master velocity nodes.

[22] In order to discretize the velocity at finer-level slave nodes (or "hanging nodes"), two principles are used: (i) conservation of volume flux across resolution levels (i.e., across a split-cell boundary) and (ii) stress-based interpolation of velocity gradients between $\tau_{xy}$ points. Conservation of volume flux across the split-cell boundary implies that volume flux (i.e., the average velocity across the coarser-cell boundary) should not change between resolution levels, for example,

$$v_y^{\mathrm{H}}=\frac{1}{2}\left(v_y^{\mathrm{G}}+v_y^{\mathrm{I}}\right). \tag{36}$$

Stress-based interpolation of velocity gradients implies that velocity gradient between two slave nodes on a coarser-cell boundary should be inversely proportional to the local viscosity similarly to the velocity gradients composing $\tau_{xy}$ stress components

$$\eta^{\mathrm{H}}\left(\frac{\partial v_y}{\partial x}\right)\bigg|_{\mathrm{H}}=\frac{1}{2}\left[\eta^{\mathrm{F}}\left(\frac{\partial v_y}{\partial x}\right)\bigg|_{\mathrm{F}}+\eta^{\mathrm{J}}\left(\frac{\partial v_y}{\partial x}\right)\bigg|_{\mathrm{J}}\right], \tag{37}$$

where

$$\left.\frac{\partial v_y}{\partial x}\right|_{\mathrm{H}} = \frac{v_y^{\mathrm{I}} - v_y^{\mathrm{G}}}{\frac{1}{2}\Delta x}, \tag{38}$$

$$\left.\frac{\partial v_y}{\partial x}\right|_{\mathrm{F}} = \frac{v_y^{\mathrm{G}} - v_y^{\mathrm{E}}}{\frac{1}{2}\Delta x}, \tag{39}$$

$$\left.\frac{\partial v_y}{\partial x}\right|_{\mathrm{J}} = \frac{v_y^{\mathrm{K}} - v_y^{\mathrm{H}}}{\Delta x}. \tag{40}$$

Again, similarly to $\dot{\epsilon}_{xy}^{\mathrm{J}}$, the coarse-level velocity nodes are used for computing $\left.\frac{\partial v_y}{\partial x}\right|_{\mathrm{J}}$. Combining equations (36) and (37) gives the following finite difference formulas for the "slave" (hanging) $v_y$ nodes G and I, respectively:

$$2\eta^{\mathrm{H}}\left(\frac{v_y^{\mathrm{H}} - v_y^{\mathrm{G}}}{\frac{1}{4}\Delta x}\right) - \eta^{\mathrm{F}}\left(\frac{v_y^{\mathrm{G}} - v_y^{\mathrm{E}}}{\frac{1}{2}\Delta x}\right) - \eta^{\mathrm{J}}\left(\frac{v_y^{\mathrm{K}} - v_y^{\mathrm{H}}}{\Delta x}\right) = 0, \tag{41}$$

$$2\eta^{\mathrm{H}}\left(\frac{v_y^{\mathrm{I}} - v_y^{\mathrm{H}}}{\frac{1}{4}\Delta x}\right) - \eta^{\mathrm{F}}\left(\frac{v_y^{\mathrm{G}} - v_y^{\mathrm{E}}}{\frac{1}{2}\Delta x}\right) - \eta^{\mathrm{J}}\left(\frac{v_y^{\mathrm{K}} - v_y^{\mathrm{H}}}{\Delta x}\right) = 0. \tag{42}$$

[23] We tested a number of other forms of the constraint used at the slave nodes. These alternative constraint formulations were found to be less successful, compared to equations (41) and (42), in consideration of the discretization error in velocity and pressure. For completeness, the trial slave node constraints are listed in Appendix A.

[24] The standard staggered grid formulation results in a compact stencil. This produces a sparse matrix, with only a few nonzero entries on each row. The stencil for the discretized momentum equation contains a total of 11 entries—nine components are associated with the velocity degrees of freedom and two with the pressure degrees of freedom. The stencil for the discrete continuity equation contains four entries, all of which are associated with velocity degrees of freedom. The adaptive stencil described above contains slightly more entries in the discrete momentum equations. The total number of entries here is dependent upon how many neighboring cells possess hanging nodes. The worst-case scenario (in terms of stencil size) occurs at node H in Figure 2. Here the adaptive stencil for the momentum equation contains a total of 15 entries: 12 associated with velocity unknowns and 3 associated with pressure unknowns. The continuity equation for the adaptive stencil always contains four entries associated with velocity unknowns. Furthermore, we note that the form of the constraint used at the slave nodes requires a maximum of five entries, each associated with velocity unknowns.

## 3. Stencil Verification

[25] In this section, we examine the behavior of the errors associated with the ASG stencil using three-variable viscosity reference problems with known analytic solutions. First, we define the reference problems (section 3.1), then we numerically determine the order of accuracy of the ASG stencil (section 3.2). Lastly, we examine profiles taken across the error field associated with one of the reference problems (section 3.3) and visually demonstrate that no numerical artifacts or spurious oscillations occur across split faces between resolution levels, even when they are intersected by discontinuous viscosity structures. In addition to these investigations, we also numerically verified the stability of the ASG scheme for variable viscosity problems by an eigenvalue analysis (see Appendix B).

### 3.1. Reference Problems

[26] In order to test and measure the numerical properties of the ASG stencil, three analytical solutions were employed. These solutions correspond to two-dimensional variable viscosity and incompressible steady state flow problems. The viscosity variations (smooth or sharp) that are inherent to these solutions are analogous to actual geodynamic problems. Large and smooth viscosity variations are likely to occur in the asthenosphere [*Stacey et al.*, 1989; *Mitrovica and Forte*, 1998], whereas the rheology of the lithosphere is characterized by material layering [*Burov*, 2011]. We therefore regard these reference problems as robust and representative tests for the ASG stencil in the framework of geodynamics.

[27] The analytical solution SOLKZ [*Revenaugh and Parsons*, 1987] is characterized by a large and smooth variation of viscosity in one dimension. The model domain is contained in $\Omega = [0,1]\times[0,1]$ box. The flow is driven by internal variations of density, and assuming that $g_x = 0, g_y = 1$, the body force in the momentum equation $\boldsymbol{f}$ is given by

$$\boldsymbol{f} = (0, -\sigma \sin(k_m y)\cos(k_n x))^{\mathrm{T}}, \tag{43}$$

with the parameters $\sigma = 1$, $k_m = 1.6\pi$, and $k_n = 3\pi$. The viscosity field is continuous and varies exponentially in the $y$ direction according to

$$\eta(x,y) = \exp(2B(1-y)), \tag{44}$$

with the parameter $B = 6.9$. The boundary conditions are free slip on all sides of the domain.

Over the depth of the model domain, this parameter choice produces a maximum viscosity ratio of $\sim 10^6$.

[28] The SOLCX analytical solution [*Zhong*, 1996] exhibits a discontinuity in the viscosity field in the $x$ direction. Similarly to SOLKZ, the domain is defined as $\Omega = [0,1] \times [0,1]$ and its boundaries are free slip. Assuming that $g_x = 0, g_y = 1$, the momentum body force is given by

$$\boldsymbol{f} = (0, -\sigma \sin(k_m y)\cos(k_n x))^{\mathrm{T}}, \tag{45}$$

with the parameters $\sigma = 1$, $k_m = \pi$, and $k_n = \pi$. The viscosity field is defined as

$$\eta(x,y) = \begin{cases} 1 & \text{for } x \leq 0.5 \\ 10^6 & \text{for } x > 0.5 \end{cases}. \tag{46}$$

The source code which evaluates the velocity and pressure fields for SOLCX and SOLKZ is distributed as part of the open source package Underworld (http://underworldproject.org).

[29] The viscous inclusion test [*Schmid*, 2002; *Schmid and Podladchikov*, 2003], termed SOLVI in the remainder of the paper, is characterized by a two-dimensional discontinuous circular viscosity structure. The physical domain is given by $\Omega = [-1,1] \times [-1,1]$. The origin of the circular inclusion is located at $(0,0)$ and has a radius, $R_{\mathrm{inc}} = 0.5$. The viscosity field is prescribed such that $\eta(x,y) = 1$ in the background and $\eta(x,y) = 10^3$ within the inclusion. Along the boundary of the domain, we prescribe $v_x$ and $v_y$ using the analytic solution. The flow is only entirely driven by the Dirichlet boundary conditions, thus the forcing term is zero. The MATLAB source code which generates the pressure and stress fields under various shear boundary conditions is available in *Schmid* [2002].

## 3.2. Order of Accuracy of the ASG Discretization

[30] According to our previous study [*Duretz et al.*, 2011], we expect the standard staggered grid discretization to yield second-order global accuracy of the velocity and pressure in the $L_1$ norm, provided the model domain contains a smooth viscosity field (SOLKZ test). On the other hand, when the viscosity field contains discontinuities, the convergence drops to first order (SOLCX and SOLVI tests). Using the same order of accuracy methodology described in *Duretz et al.* [2011], here we establish whether the discretization errors associated with ASG stencil exhibit the same characteristics, in regard to viscosity structure, as the standard staggered discretization. We refer to Appendix C for the definition of the norms, the a priori error estimate, and the technique used to numerically evaluate the discretization errors. We note that throughout the remainder of this paper, all of the reported errors computed via the $L_1$ norm have been normalized by the volume of the domain such that the error measured possesses the same physical dimensions as the original quantity.

[31] The grid convergence tests were carried out using the arbitrary mesh depicted in Figure 3a. The white region represents the area of highest resolution (Level 1), $h$ being twice smaller than in the shaded region (Level 0). Such an adapted mesh ensures that the grid is not aligned with the geometry of the viscosity structures defined in the reference problems described above. In that way, we avoid any potential superconvergent cases that may occur when viscosity discontinuities are aligned with the grid. This condition reflects a general case in which material boundaries cut across the cells. Hence, the circular refined areas crosscut the sharp viscosity structures that are characteristics of both SOLCX and SOLVI tests. We emphasize that the mesh used in the order of accuracy test is not an optimal mesh, i.e., that which minimizes the discretization error. Rather, it was chosen as a potentially "worst-case scenario" to highlight any numerical artifacts which might arise when discontinuous viscosity structures intersect interfaces between different mesh resolutions within the adaptive grid.

[32] The results of the convergence test are reported in Figure 3. Figures 3b–3d display the reduction of the $L_1$ error with homogenous refinement of the mesh for the different test cases (SOLKZ, SOLCX, and SOLVI). The errors are plotted against the corresponding coarser-mesh resolution, $h_{\mathrm{Level\,0}}$. The employed coarse grids have the same number of nodes $N_{\mathrm{Level\,0}}$ in both directions, and the grid sequence $N_{\mathrm{Level\,0}}$ is defined as $N_{\mathrm{Level\,0}} = \{21, 41, 81, 161\}$. For reference, the figure also displays the convergence of the standard staggered grid using a nonadaptive grid sequence that is equivalent to $N_{\mathrm{Level\,0}}$. The order of accuracy is reported in the top-left corner of Figures 3b–3d. The continuous viscosity test, SOLKZ, yielded to almost second order of accuracy for both velocity and pressure fields (Figure 3b). In case of a one-dimensional discontinuity in the viscosity field, the SOLCX test exhibited first-order velocity and pressure accuracy (Figure 3c). Similarly, SOLVI leads to

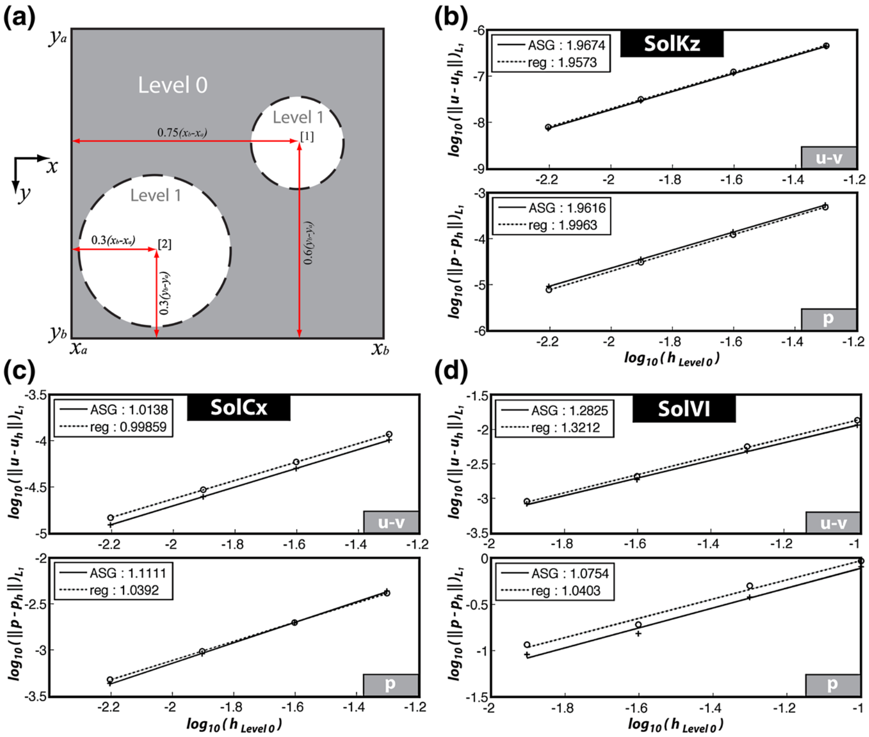

**Figure 3.** Grid convergence of the ASG for variable viscosity Stokes problems. (a) The arbitrary mesh used for the tests. This grid contains two circular areas of refined cells (Level 1). The radii of the circles are $R_1 = 0.15(x_b - x_a)$ and $R_2 = 0.25(x_b - x_a)$. The grid convergence for (b) SOLKZ, (c) SOLCX, and (d) SOLVI. Dashed lines depict results obtained using the standard staggered grid discretization with $h_{\text{Level 0}}$, and solid lines indicate the ASG results obtained using the locally refined mesh in Figure 3a.

first-order velocity accuracy and a pressure accuracy that slightly exceeds first order (Figure 3d).

[33] The results indicate that the ASG stencil conserves the convergence properties of the standard staggered grid. Therefore, the hanging node scheme derived in section 2.2 does not alter the accuracy of the method. This is a crucial point since the use of an inadequate hanging node scheme can severely degrade the convergence of the ASG. We refer to Appendix A where this is illustrated.

### 3.3. Velocity and Pressure Behavior Across Resolution Transitions

[34] The ASG scheme on quadtree meshes implies that the numerical resolution $h$ is not smooth and varies as $h_{\text{Level n}} = 2^{-n} h_{\text{Level 0}}$ at each grid-level transition. Such sharp variation in the grid resolution can potentially generate numerical artifacts (e.g., unphysical solution oscillations in space) in the vicinity of the grid-level transitions. Moreover, numerical artifacts will possibly be amplified at material boundaries since viscosity jumps are likely to produce locally large discretization errors [*Deubelbeiss and Kaus*, 2008].

[35] In order to evaluate the behavior of ASG numerical solution at grid-level transitions which intersect discontinuous viscosity structures, we employed the SOLVI test described above (section 3.1). The mesh structure was chosen such that the resolution increases from the leftside to the rightside of the domain. Two mesh resolution transitions were prescribed at $x = -0.2$ and $x = 0.2$, such that they cut across the material boundary represented by the inclusion. With the purpose of verifying the smoothness of the solution field across grid levels, both pressure and velocity discretization errors were computed cell-wise from the analytical solution. The top row in Figure 4 depicts the spatial distribution of discretization error within the model domain. The transitions between grid levels do not locally generate large discretization errors nor spurious oscillations. The smooth behavior of the error is further illustrated in the bottom row of Figure 4. These plots show profiles of the absolute discretization error probed along the $x$ direction for various values of the $y$ coordinate. The profiles taken at $y = 0.25$ cut across the material boundary and the grid-level transitions. Nevertheless, the variations of the discretization

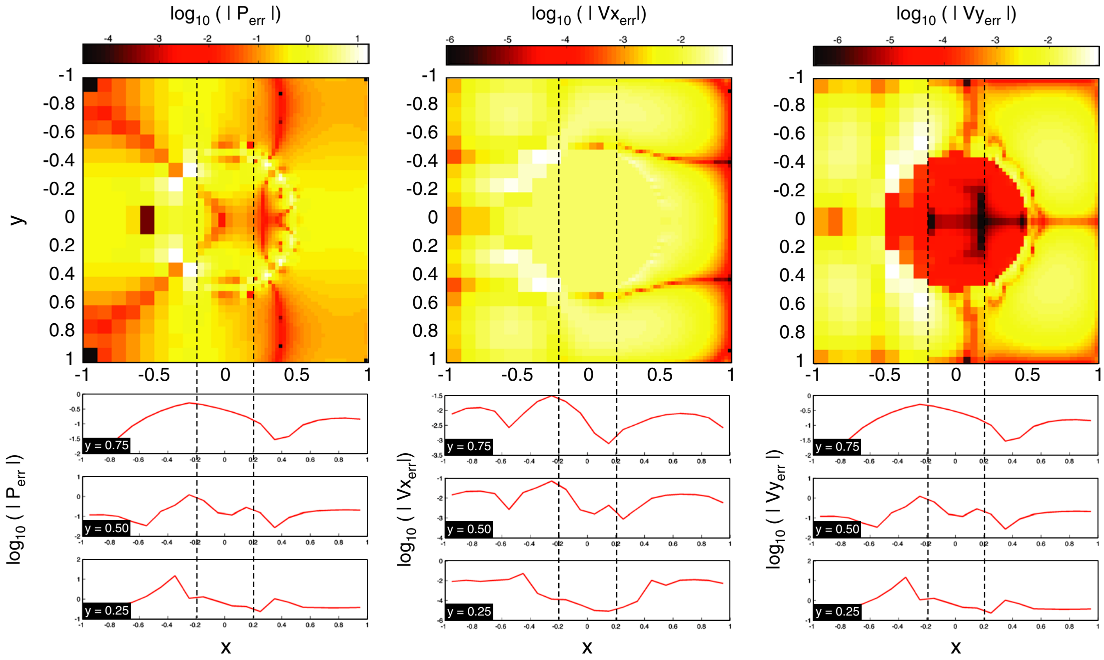

**Figure 4.** Spatial distribution of discretization errors of the ASG numerical solution using three resolution levels. (top row) The absolute value of the pressure and velocity discretization errors for SOLVI. The resolution increases from the left to the right, and grid-level transitions are located at $x$ = –0.2 and $x$ = 0.2 (black dashed line). The pressure in the four corners is defined as Dirichlet values, therefore leading to zero discretization error (black squares).(bottom row) Error profiles at $y$ = 0.25, $y$ = 0.5, and $y$ = 0.75. The red line indicates the ASG error profiles.

error (at $x \sim -0.35$ and $x \sim 0.35$) are related to the viscosity jumps and do not coincide with the mesh refinement. The velocity and pressure solutions provided by the ASG are therefore smooth, and their corresponding discretization error does not locally deteriorate due to transitions in the mesh resolution.

## 4. Performance of the ASG Discretization

[36] The purpose of using an adaptive spatial discretization is that one can locally modify the spatial discretization to minimize a given objective function $\mathcal{L}$. The nature of the objective function is application specific and may relate to the discretization error of the velocity and pressure field (in the case of Stokes flow), or it might be constructed in a manner to measure the accuracy of some quantity derived from the numerical solution, e.g., the total stress around the rim of a clast (c.f. SOLVI).

[37] Using adaptivity, the ASG discretization should deliver a numerical solution which is as accurate (with regard to the given objective function) as a high-resolution standard staggered grid; however, the ASG mesh will contain far fewer cells. Thus, the discrete system will contain less unknowns ($N_{\mathrm{dof}}$) compared to the nonadaptive discretization, which has the advantage of significantly reducing both the CPU time and memory requirements of the calculation.

[38] Adaptive procedures can thus be considered as an optimization problem related to the available computational resource. The optimization problem can be expressed by either of the following statements:

[39] *C1:* Given a tolerance $\epsilon_0$, minimize the computational work $W$ required to obtain a discrete solution $v_h, p_h$ such that $\mathcal{L} < \epsilon_0$.

[40] *C2:* For a fixed amount of computational work $W_0$, find a solution $v_h, p_h$ which minimizes $\mathcal{L}$.

[41] The adaptive procedure for a steady state problem can then be summarized as follows:

[42] 1. Generate an initial mesh, $\mathcal{T}^0$. Specify a maximum number of adaptive iterations $i_{\mathrm{max}}$. Initialize the adaptive cycle counter, $i = 0$.

[43] 2. **Solve:** Solve the discrete problem on the current mesh, $\mathcal{T}^i$.

[44] 3. **Estimate:** Analyze the computed solution. Estimate the error $\mathcal{L}$ and evaluate the error on each cell. If either condition *C1* or *C2* is satisfied $\longrightarrow$ stop.

[45] 4. **Mark:** Based on the chosen refinement indicator(s) (e.g., on local error), identify cells

which should be modified, i.e., coarsened or refined.

[46] 5. **Adapt:** Using the cells marked for adaptation, generate a new mesh, $\mathcal{T}^{i+1}$.

[47] 6. If $i < i_{\max}$, set $i = i + 1$ and go to Step 2, else stop.

## 4.1. Adaptive Procedure

[48] At each iteration $i$, the refinement algorithm identifies the location where resolution needs to be increased. Refinement is usually based on the current value of a given function $\Phi^i$ (i.e., refinement indicator) that can be evaluated in cells' center and/or vertices (e.g., velocity gradient). The upper threshold for refinement $\Phi^{i+1}_{\max}$ is iteratively updated according to

$$\Phi^{i+1}_{\max} = \bar{\Phi}^i + \alpha_\Phi \left(\Phi^i_{\max} - \bar{\Phi}^i\right). \tag{47}$$

The parameter $\Phi^i_{\max} = \max\left(|\Phi^i|\right)$, $\bar{\Phi}^i$ is the average function value defined as $\bar{\Phi}^i = \dfrac{1}{n}\sum_{k=1}^{n}\Phi^i(\boldsymbol{x}_k)$, where $n$ is the number of points (e.g., pressure or velocity nodes) where the discrete field $\Phi^i$ is evaluated, and $\alpha_\Phi \in (0,\ 1]$ is the refinement factor. During the cell refinement cycle, cells in which the value of $\Phi^i$ exceeds $\Phi^{i+1}_{\max}$ are split into four cells. In the following tests, a minimum of two cells separates two grid-level transitions, and cells that are surrounded by at least three higher resolution cells are recursively split. The way $\Phi^i_{\max}$ and $\bar{\Phi}^i$ are computed for each of the refinement indicators employed in the following tests is described in Appendix D.

## 4.2. Adaptive Grid Performance Metric

[49] As shown in section 3.2, the standard staggered grid and the ASG method exhibit the same velocity and pressure order of accuracy in $L_1$ error norms. The methods attain second-order accuracy in presence of continuous viscosity field and degenerate to first order in the presence of viscosity jumps. For the standard two-dimensional staggered grid, assuming $N_x = N_y$ and identical grid spacing $h$ in both dimensions, the number of degrees of freedom scales as $N_{\mathrm{dof}} = 3N_xN_y - 2N_x - 2N_y + 1 \propto {N_x}^2$, and $h$ scales as $h \propto L\,N_{dof}^{-\frac{1}{2}}$, in which the domain $\Omega$ is given by $L \times L$. Therefore, from equation (C6), the standard staggered grid velocity and pressure errors are

$$\|v_{\mathrm{err}}\|_{L_1(\Omega)} = \|\boldsymbol{v} - \boldsymbol{v}_h\|_{L_1(\Omega)} \le D_1 {N_{\mathrm{dof}}}^{-\frac{k_v}{2}},$$
$$\|p_{\mathrm{err}}\|_{L_1(\Omega)} = \|p - p_h\|_{L_1(\Omega)} \le D_2 {N_{\mathrm{dof}}}^{-\frac{k_p}{2}}, \tag{48}$$

where $D_1 = C_1L^{k_v}$ and $D_2 = C_2L^{k_p}$ are constants, and $k_v$ and $k_p$ are the order of accuracy of the velocity and pressure discretization. The velocity and pressure error reduction rates, $R_v$, $R_p$, are, respectively, equal to $\frac{k_v}{2}$ and $\frac{k_p}{2}$, and equation (48) can be rewritten as

$$\|v_{\mathrm{err}}\|_{L_1(\Omega)} \le D_1 {N_{\mathrm{dof}}}^{-R_v},$$
$$\|p_{\mathrm{err}}\|_{L_1(\Omega)} \le D_2 {N_{\mathrm{dof}}}^{-R_p}. \tag{49}$$

Accordingly, for the standard, nonadaptive staggered grid discretization, the $L_1$ errors scale as $(N_{\mathrm{dof}})^{-1}$ if the viscosity field is smooth $(R_v, R_p \sim 1)$ and as $(N_{\mathrm{dof}})^{-1/2}$ in the presence of viscosity jumps $\left(R_v, R_p \sim \frac{1}{2}\right)$.

[50] One advantage of using a spatially adaptive discretization is the ability to locally reduce cell sizes close to viscosity discontinuities, thereby locally reducing the discretization error and globally reducing the $L_1$ error norms. We can therefore expect that for a given number of degrees of freedom, adaptivity will lead to lower integrated discretization errors than the nonadaptive grid solution.

[51] In the following sections, we test various refinement indicators and evaluate how the $L_1$ discretization error scales with respect to the number of degrees of freedom. In all experiments, the initial mesh ($\mathcal{T}^0$) is quite coarse, with a uniform grid resolution employing $N_x = N_y = 21$ vertices in the $x$ and $y$ directions, respectively. Thus, at each stage of the adaptive cycle, the mesh is primarily modified via refinement. A maximum number of adaptive cycles were chosen as $i_{\max} = 25$. Adaptive cycles were terminated if $N_{\mathrm{dof}} > 70{,}000$. As in section 3.2, we employ the SolCx, SolKz, and SolVI tests and obtain the discretization errors from the analytical solutions. A complete summary of the convergence rates for $R_v, R_p$ is provided in Table 1.

## 4.3. Exact Discretization Error-Based Adaptivity

[52] Here we use the value of the exact discretization error to drive adaptivity. At each adaptive cycle, the flow and pressure fields along with their associated errors were evaluated. The refinement criteria were updated according to equation (47) for either $\alpha_{v_{\mathrm{err}}}$, $\alpha_{p_{\mathrm{err}}}$, or both. Subsequently, the mesh geometry was modified in order to satisfy the current refinement criteria and the conditions listed in section 4.1. The refinement factors, $\alpha_{v_{\mathrm{err}}}$ and $\alpha_{p_{\mathrm{err}}}$, were varied between 0.1 and 0.9 for every test. For each refinement test, we measured the decrease rate of discretization error ($R_v$ and $R_p$) with increasing number of degrees of freedom in logarithmic space;

**Table 1.** Convergence of the Volume-Weighted $L_1$ Velocity and Pressure Errors With Increasing Number of Degrees of Freedom[a]

| | SolKz | | SolCx | | SolVI | |
|---|---|---|---|---|---|---|
| Refinement Indicator ($\Phi$) | $R_v$ | $R_p$ | $R_v$ | $R_p$ | $R_v$ | $R_p$ |
| $v_{\mathrm{err}}$ | 0.924 | 0.437 | 0.472 | 0.405 | 0.983$^{*}$ | 0.939$^{*}$ |
| $p_{\mathrm{err}}$ | 0.763 | 0.416 | 0.959$^{*}$ | 0.705$^{*}$ | 0.970$^{*}$ | 0.945$^{*}$ |
| $v_{\mathrm{err}}, p_{\mathrm{err}}$ | 0.974$^{*}$ | 0.491 | 0.746$^{*}$ | 0.730$^{*}$ | 1.022$^{*}$ | 0.966$^{*}$ |
| $\Delta_v$ | 0.714 | 0.326 | 0.678$^{*}$ | 0.655$^{*}$ | 0.927$^{*}$ | 0.724$^{*}$ |
| $\Delta_p$ | 0.501 | 0.566 | 0.376 | 0.265 | 0.983$^{*}$ | 0.939$^{*}$ |
| $\Delta_v$, $\Delta_p$ | 0.981$^{*}$ | 0.625 | 0.641$^{*}$ | 0.581$^{*}$ | 1.047$^{*}$ | 1.094$^{*}$ |
| $\Delta_\rho$ | 0.922 | 0.554 | 0.448 | 0.448 | — | — |
| $\Delta_{\log_{10}\eta}$ | — | — | 0.906$^{*}$ | 0.944$^{*}$ | 1.040$^{*}$ | 0.998$^{*}$ |
| $\Delta_\rho$, $\Delta_{\log_{10}\eta}$ | — | — | 0.910$^{*}$ | 0.918$^{*}$ | — | — |
| $\Delta_v$, $\Delta_\eta$ | — | — | 1.037$^{*}$ | 0.870$^{*}$ | 0.914$^{*}$ | 0.724$^{*}$ |
| $\Delta_p$, $\Delta_\eta$ | — | — | 0.897$^{*}$ | 0.893$^{*}$ | 1.028$^{*}$ | 0.985$^{*}$ |
| $\Delta_v$, $\Delta_\rho$ | 0.982$^{*}$ | 0.558 | 0.650$^{*}$ | 0.598$^{*}$ | — | — |
| $\Delta_p$, $\Delta_\rho$ | 0.916 | 0.990 | 0.528$^{*}$ | 0.525 | — | — |
| Uniform mesh | 0.949 | 1.000 | 0.492 | 0.536 | 0.665 | 0.524 |

[a] The refinement strategies tested were based on exact discretization error ($v_{\mathrm{err}}$, $p_{\mathrm{err}}$), solution contrasts ($\Delta_v$, $\Delta_p$), material properties contrasts ($\Delta_\rho$, $\Delta_{\log_{10}\eta}$), or a combination of them. Results from reference problems which did not include either density variations or constant/linear viscosity variations in logarithmic space are excluded. The values of $R_v$, $R_p$ obtained using a uniform mesh are shown in the last row. Values marked with (*) indicate a rate which is larger than that obtained using a uniform mesh.

these results are compiled in Table 1 and Figure 6. The final mesh structures obtained using the exact discretization error-based adaptivity (both $v_{\mathrm{err}}$ and $p_{\mathrm{err}}$) are depicted in Figure 5.

[53] For SolCx, refinement based on $v_{\mathrm{err}}$ leads to results that are less efficient than nonadaptive solutions, namely $R_v, R_p \leq 0.5$. On the other hand, refinement based on $p_{\mathrm{err}}$, leads to a pressure discretization error decrease which is more efficient than nonadaptive solution with $R_v, R_p \sim 1 > 0.5$. We noticed that for $\alpha_\Phi > 0.8$, velocity and pressure error decrease rates can grow larger than 1 using refinement based on pressure error. The combination of both refinement criteria produced refinement close to the material boundary and in the weak convecting material which are, respectively, the sources of the largest pressure and velocity errors. This combination of refinement criteria produced an efficient decrease of both integrated velocity and pressure error, yielding to $R_v, R_p \sim 0.75 > 0.5$.

[54] The inclusion problem (SolVI) yielded results that are much more accurate than nonadaptive meshes of a similar number of degrees of freedom. For refinement based on either $v_{\mathrm{err}}$ or $p_{\mathrm{err}}$, the values of $R_v$ and $R_p$ are greater than 0.9. The combination of both $v_{\mathrm{err}}$ or $p_{\mathrm{err}}$ for refinement leads to aggressive refinement around the inclusion rim and in the surrounding fluid. Values of $R_v$ and $R_p$ reaching $\sim 1$ imply that the errors scale as $(N_{\mathrm{dof}})^{-1}$, whereas nonadaptive solutions only yielded $(N_{\mathrm{dof}})^{-\frac{1}{2}}$ in the presence of a viscosity jump.

[55] The SolKz test produced mediocre results on adaptive meshes. The use of $v_{\mathrm{err}}$, $p_{\mathrm{err}}$, or both as refinement criteria leads to results that are less efficient than nonadaptive meshes ($R_v, R_p \leq 1$) for problems with smooth viscosity fields. For this reference problem, the pressure error reduction rates are particularly low with $R_p \leq 0.5$. The main reason for obtaining low-error reduction rate is related to sharp changes in grid resolution (doubling) across resolution level, which makes finite difference stencil notably asymmetric. Stencil asymmetry lowers the accuracy of numerical solution for pressure at the resolution boundaries (typically 2–3 times higher errors are observed compared to the uniform nonadaptive grid), with contribution to global error being proportional to the total length of the resolution boundaries. Since the length of resolution boundaries grows during the refinement process, pressure error reduction rates are notably lowered. The problem of lowered accuracy of strongly asymmetric staggered grid stencils is well known [*Sundqvist and Veronis*, 1970; *Crowder and Dalton*, 1971; *de Rivas*, 1972]. Additional calculations performed with SolKz test have shown the same character and magnitude of pressure error

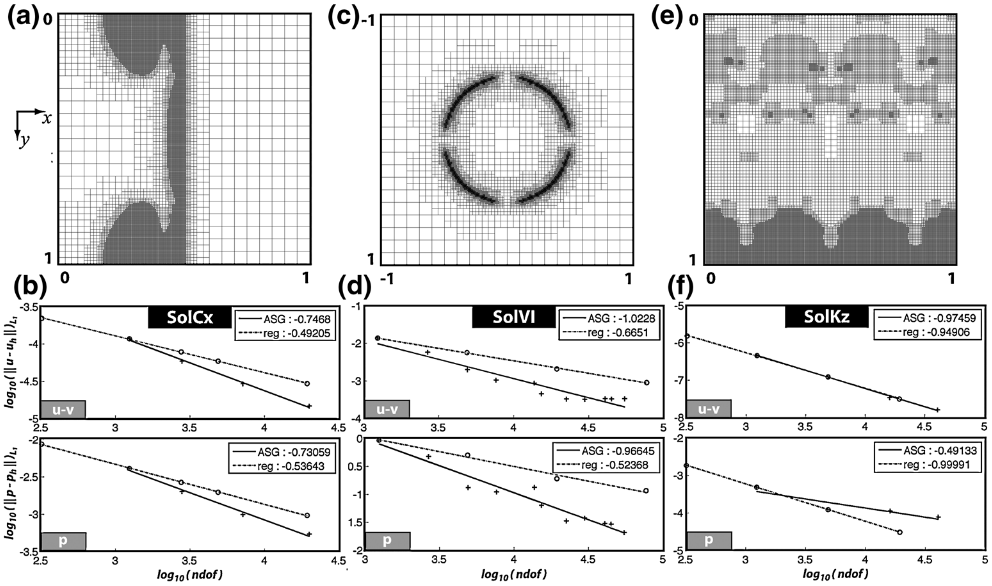

**Figure 5.** Mesh geometry and discretization errors using adaptivity based on the true discretization error. All tests employed both $v_{\mathrm{err}}$ and $p_{\mathrm{err}}$ as refinement criteria. (a, c, and e) The adaptive mesh geometry at the end of each test (up to eight levels). (b, d, and f) The decrease of $L_1$ discretization error with increasing $N_{\mathrm{dof}}$. For comparison, nonrefined results using the standard staggered grid (reg) are shown for comparison (dashed lines). Figures 5a and 5b relate to SolCx, Figures 5c and 5d relate to SolVI, and Figures 5e and 5f relate to SolKz.

increase for the case of a simple nonuniform rectangular grid in which vertical resolution is doubled in a limited $y$ interval. We conclude therefore that this problem is not specific to ASG and further effort may be needed in the future to develop more accurate finite difference schemes on staggered grids with sharply variable grid spacing. In general, adaptive strategies based on the true discretization error are not practical, i.e., without knowledge of the analytical solution of the given problem or using a posteriori error estimates. Thus, in order to provide a more complete test of the ASG scheme, refinement tests based on the discrete solution and material properties contrasts were performed.

## 4.4. Field-Based Adaptivity

[56] Velocity contrasts ($\Delta_v$) are likely to develop in shear layers or in the vicinity of material boundaries. Large pressure contrasts ($\Delta_p$) are representative of discontinuous viscosity structures. One might thus expect $\Delta_v$ and $\Delta_p$ based refinement to behave comparably to refinement based on velocity and pressure errors.

[57] For SolCx, refinement based on the velocity contrast leads to efficient results ($R_v, R_p \sim 0.7 > 0.5$), whereas pressure difference refinement behaved less efficiently than nonadaptive grid solutions. Combining the two criteria leads to locally high resolution at the material jump and in the low viscosity material and results in improved values of $R_v, R_p$ compared to nonadaptive mesh calculations. Solution-contrast-based refinement leads to efficient results when applied to the SolVI test. The values of $R_v$ and $R_p$ reached 1 when combining $\Delta_v$ and $\Delta_p$. Regarding the SolKz test, none of the tested solution-contrast-based refinement schemes yielded efficient error reduction rates. The obtained values of $R_v$ and $R_p$ ranged between 0.3 and 1 (see Table 1 and Figure 6a) which are less, or equal to the results obtained with nonadaptive grids.

## 4.5. Material Property-Based Adaptivity

[58] Refinement based on material properties contrasts enables the mesh to adapt in accordance with the evolution of the properties of a fluid. For a Stokes problem, the contrasts of interest are the viscosity and density contrasts ($\Delta_{\log_{10}\eta}$ and $\Delta\rho$). Theses quantities are likely to represent reasonable choices for refinement since they often represent material interfaces, which are responsible for localizing large discretization errors. The test SolCx is characterized by viscosity variations, which were captured using refinement factor values equal to $\alpha_{\log_{10}\eta} = \alpha_\rho = 0.2$. Although $\Delta\rho$ based refinement did not produce efficient results, both $\Delta_{\log_{10}\eta}$ and the combination of $\Delta_{\log_{10}\eta}$ and $\Delta\rho$ as a refinement

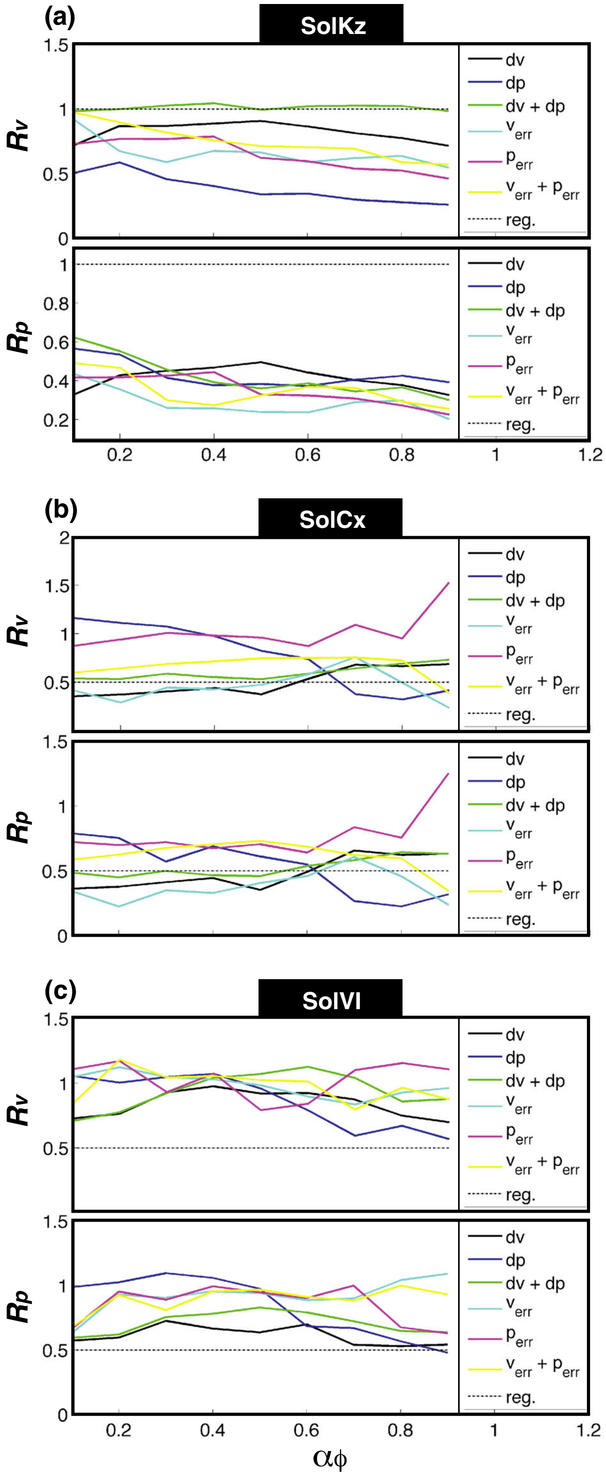

**Figure 6.** Error reduction rates for as a function of $\alpha_\Phi$ (equation (47)). Color lines represent different refinement indicators ($\Phi = \{\Delta_v\ \Delta_p\ v_{\text{err}}\ p_{\text{err}}\}$). The dashed lines correspond to the nonadaptive solution $L_1$ error reduction rate which is constant for each given problem. (a) SolKz, (b) SolCx, and the inclusion test (c) SolVI.

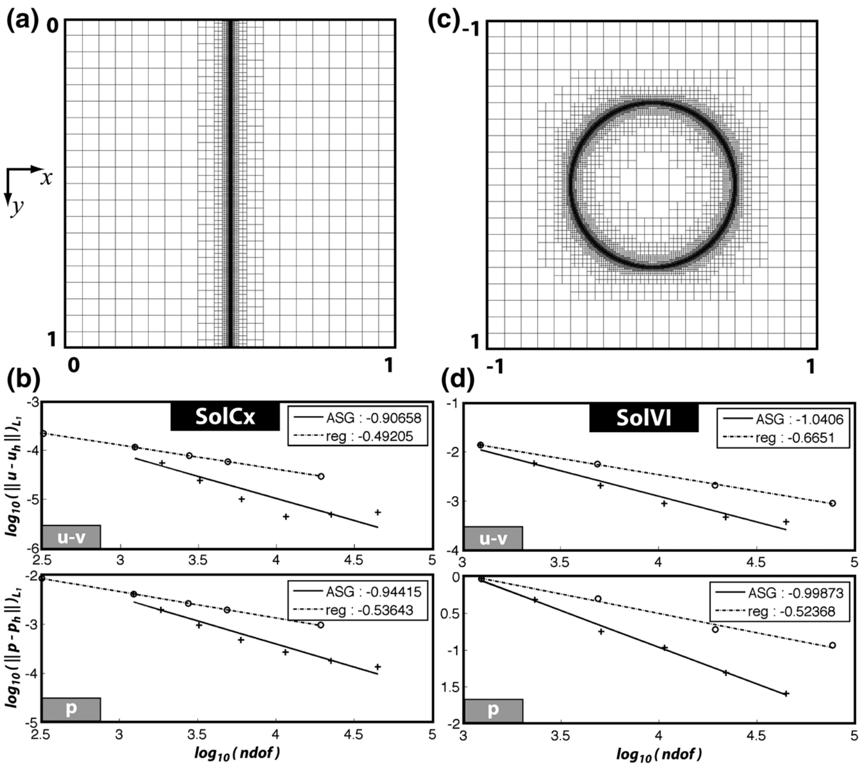

**Figure 7.** The mesh structure obtained using viscosity contrasts based refinement for (a) SolCx and (c) SolVI. The decrease of $L_1$ pressure and velocity discretization error with increasing $N_{\text{dof}}$ for (b) SolCx and (d) SolVI. Results for SolKz are not presented here. This reference problem includes a linear variation of viscosity in logarithmic space, which is not captured by our viscosity-based refinement scheme.

criteria yielded error reduction rates of $R_v, R_p > 0.9$. The grid structure achieved with viscosity contrast refinement is depicted in Figure 7a. SolVI is purely driven by kinematic boundary conditions. Since no buoyancy contrast exists, only $\Delta_{\log_{10}\eta}$ based refinement was tested. Similarly, as for the refinement indicators discussed in sections 4.3 and 4.4, the results were successful with $R_v, R_p > 1$. Both velocity and pressure error reduction rates scaled as $(N_{\text{dof}})^{-1}$ (Figure 7b). Conversely, SolKz does not include logarithmic viscosity contrasts which could be captured. The results obtained for $\Delta\rho$ based refinement (Table 1) are, once again, less efficient than the nonadaptive mesh solution.

## 5. Geodynamic Examples

[59] In this section, we will display simplified examples of geodynamic applications of ASG for (1) spontaneous localization of deformation during extension of a viscoplastic material and (2) deformation of a self-gravitating planetary object associated with movement of an iron diapir. The main purpose of these examples is to demonstrate potential of ASG for solving realistic regional and global numerical modeling problems with complex geometries and nonlinear rheology under conditions of large deformation. In both examples, ASG stencil is used in combination with standard Marker-And-Cell (MAC) technique [*Gerya and Yuen*, 2003, 2007; *Gerya*, 2010] to advect and visualize material properties.

### 5.1. Extension of Viscoplastic Media

[60] Numerical modeling of extension and shortening of viscoplastic materials poses significant computational challenges [*Buiter et al.*, 2006; *Gerya*, 2010] which require that any numerical code is able to (1) calculate large deformations along spontaneously forming narrow shear zones, (2) represent complex boundary conditions, including free surfaces, and (3) include a complex rheology involving both viscous and frictional/plastic materials. These challenges directly reflect the state-of-the-art requirements for numerical modeling of lithospheric-scale tectonic processes. We conducted an extensional numerical experiment of a simplified model resembling the sandbox benchmark conducted by *Buiter et al.* [2006] (Figure 8).

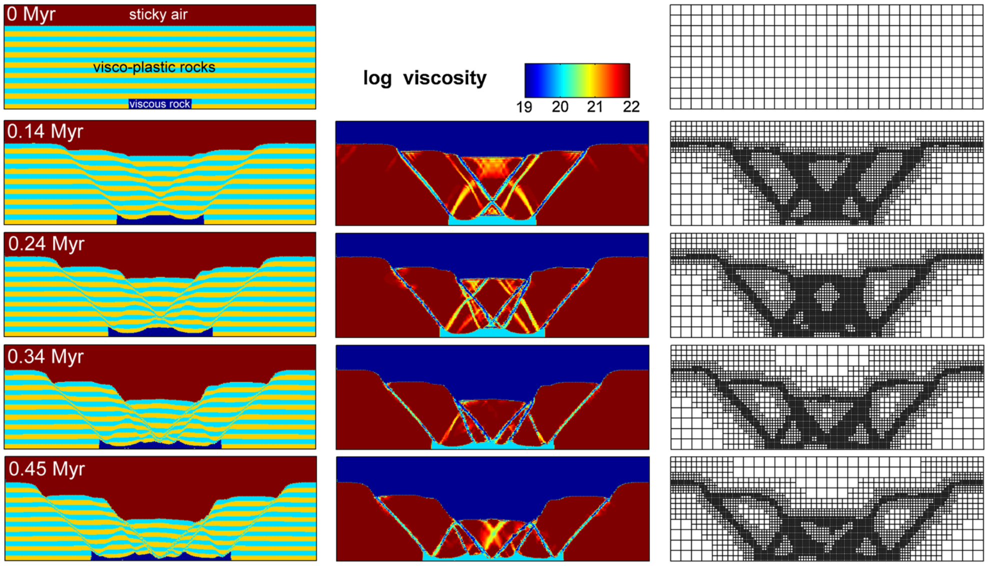

**Figure 8.** Initial conditions and results for the extension experiment (see text for details). Evolution of the (left) lithological structure, (middle) viscosity field, and (right) grid structure.

[61] An Eulerian numerical modeling domain of $330 \times 10$ km in size (Figure 8) is employed. Initially a 2 km thick sticky air layer (density 1 kg/m$^3$, viscosity $10^{19}$ Pa s) is used to simulate the free surface condition [e.g., *Schmeling et al.*, 2008; *Crameri et al.*, 2012b]. The remaining 8 km of the initial model cross section is composed of viscoplastic rocks (density 3000 kg/m$^3$, viscosity $10^{22}$ Pa s) with brittle strength ($\sigma_{yield}$) defined via a Drucker-Prager yield criterion with strain weakening [e.g., *Lavier et al.*, 2000; *Huismans and Beaumont*, 2002; *Buiter et al.*, 2006] which is given by

$$\sigma_{\text{yield}} = C + \phi P, \tag{50}$$

in which we use

$$\phi = \begin{cases} 0.6 - 0.05\,\gamma & \text{for } \gamma \le 2 \\ 0.5 & \text{for } \gamma > 2 \end{cases};$$
$$C = \begin{cases} (3 - 0.5\,\gamma) \times 10^7 & \text{for } \gamma \le 2 \\ 2 \times 10^7 & \text{for } \gamma > 2 \end{cases}, \tag{51}$$

where $P$ is dynamic pressure (Pa), $\phi$ is friction coefficient, $C$ is the rock strength at $P = 0$ (Pa), and $\gamma$ is the integrated plastic strain. A weak viscous $6 \times 1$ km inclusion (density 3000 kg/m$^3$, viscosity $10^{20}$ Pa s) is prescribed at the bottom of the model to initiate localized deformation [*Buiter et al.*, 2006]. The initial grid resolution is $30 \times 10$ cells with 76,800 regularly distributed Lagrangian markers used for tracing of material properties. Four levels of resolution of ASG are used in this model with the finest grid corresponding to an effective resolution of $240 \times 80$ cells. Logarithmic viscosity contrast $\Delta_{\log_{10}\eta}$ is used as a grid refinement ($\Delta_{\log_{10}\eta} \ge 0.5$) and coarsening ($\Delta_{\log_{10}}\eta \le 0.2$) indicator. Figure 8 shows the evolution of the model cross section through time. Viscoplastic deformation of layered rocks is localized along spontaneously forming narrow shear zones where the effective viscosity of rocks is strongly reduced. ASG resolution is changing with time and closely following the rock deformation pattern. Maximal grid resolution is always achieved along the active shear zones and at the free surface. This model example thus shows that the combination of ASG with MAC can serve as an efficient tool for modeling of complex lithospheric deformation with nonlinear rheology and spontaneous localization phenomena.

## 5.2. Planetary Deformation

[62] Numerical modeling of a self-gravitating planetary body requires coupling of a gravitational field and interior deformation, with a freely evolving planetary surface [*Gerya and Yuen*, 2007; *Lin et al.*, 2009; *Golabek et al.*, 2009]. These requirements can be satisfied with the use of "spherical-Cartesian" approach (Figure 9) allowing the computation of a self-gravitating body of arbitrary form on the Cartesian grid with an approximate free surface condition (see details and tests in *Gerya and Yuen* [2007], *Lin et al.* [2009], and *Gerya* [2010]). We summarize the modeling procedure below:

[63] 1. The body is surrounded by "sticky air" (density 1 kg/m$^3$, viscosity $10^{20}$ Pa s) allowing

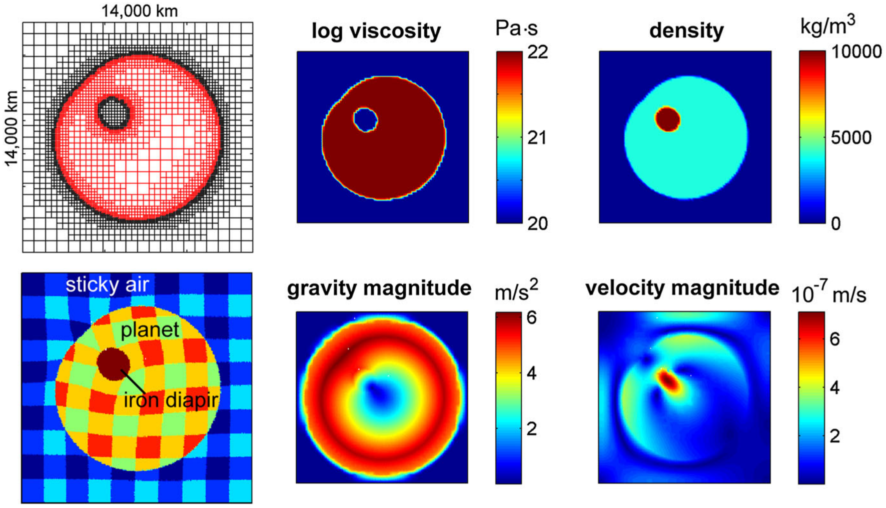

**Figure 9.** Results for the planetary deformation experiment (see text for details).

a high (100) viscosity contrast at the planetary surface which closely approximates the free surface condition.

[64] 2. The gravity field is computed by solving the Poisson equation for the gravitational potential associated with the density field defined by the current configuration of the markers at each time step.

[65] 3. During the solution of the momentum equations (equations (2) and (3)), the components of gravitational acceleration vector $(g_x, g_y)$ are computed locally from the gravitational potential at the corresponding nodal points.

[66] A test experiment shown in Figure 9, applied this methodology in combination with ASG for the case of planetary deformation associated with the propagation of a large iron diapir (density 10,000 kg/m$^3$, viscosity $10^{20}$ Pa s) toward the center of a silicate planet (density 4000 kg/m$^3$, viscosity $10^{22}$ Pa s). The gravity potential is defined in the centers of cells (i.e., in pressure points) where the Poisson equation is formulated and discretized [*Gerya and Yuen*, 2007; *Lin et al.*, 2009; *Gerya*, 2010]. The initial grid resolution for the $14,000 \times 14,000$ km computational domain is $20 \times 20$ cells, with 102,400 randomly distributed Lagrangian markers used for tracing of material properties. Four levels of resolution of ASG are used in this model with the finest grid corresponding to an effective resolution of $160 \times 160$ cells. Logarithmic viscosity contrast $\Delta_{\log_{10}\eta}$ is used as a grid refinement $(\Delta_{\log_{10}\eta} \geq 1)$ and coarsening $(\Delta_{\log_{10}\eta} \leq 0.5)$ indicator. The results of the experiment show that the spontaneously deforming planetary surface is numerically stable under conditions of large viscosity contrasts and ongoing internal deformation inside the planet. The highest ASG resolution focuses around the propagating iron diapir and at the planetary surface. The test thus demonstrates that the combination of ASG with MAC can potentially be extended for planetary scale models [e.g., *Golabek et al.*, 2009; *Stadler et al.*, 2010] which account for self-gravitation, freely evolving planetary surface, and large viscosity contrasts.

## 6. Discussion

[67] One of the desirable properties of the adaptive staggered grid stencil developed here is that it does not introduce significantly more entries into the stencil compared to the standard nonadaptive staggered grid stencil. The continuity stencil used in the adaptive formulation requires exactly the same number of entries as the nonadaptive stencil. At the interface between regions of refinement, the stencil for the momentum equations grows from 11 entries to a maximum of 15 entries. We note that this slight increase in stencil size occurs only at the transitions in the mesh resolution—away from the transition, the stencil used reverts to the standard staggered grid formulation. At hanging (slave) velocity nodes, the stencil used to define the velocity constraint only contains five entries. Given that the interface between mesh resolutions is of one spatial dimension lower than the physical problem (2-D), the adaptive stencil will produce only a slight increase in the total storage requirements. Maintaining a compact stencil is important to preserve matrix sparsity. Sparsity is desirable as it reduces the

memory requirements needed to store the matrix, the assembly of the matrix can be performed efficiently as only a small number of nonzero entries have to be computed per stencil, and the sparsity is a desirable property for both sparse direct factorization algorithms and iterative methods.

[68] As already stated, the ASG introduced here has the benefit that it is a low-order method and preserves a compact stencil. This results in a "cheap" computational method, thereby enabling high-resolution computations to be performed with only modest computational resources. Furthermore, we have demonstrated that the quality of the discrete solution obtained for velocity and pressure for variable viscosity (smooth and discontinuous) is comparable to the standard staggered grid formulation. On the consideration of memory requirements alone, the ASG discretization method is superior to any stable finite element method applied on an orthogonal block-structured mesh. If we consider a conforming finite element space applied to the velocity field on a structured mesh, and we *only* consider the size of the stencil associated with the gradient of the deviatoric stress (i.e., we only consider velocity-velocity coupling in the discrete problem), we find for $Q_1$ (bilinear) elements that the stencil contains 18 entries and for $Q_2$ (biquadratic), the stencil contains 50 entries. Thus, even without considering the contribution of the pressure degrees of freedom, the finite element stencil is larger than that obtained with the ASG method. A careful and thorough comparison which focuses on the accuracy of the velocity/pressure fields obtained using staggered grid methods and the finite element method has yet to be performed with sufficient rigor. Such a comparison must examine the discretization errors obtained using structured grid formulations which are coupled with suitable coefficient interpolants to map marker properties (viscosity and density) on to the mesh (such as those defined in *Duretz et al.* [2011]). To be a fair comparison, equivalent error measures must be employed. Answering the question "are finite differences better than finite elements for block-structured adaptive meshes?", in the context of marker-and-cell-based approaches, requires that such a study be performed.

[69] Having demonstrated the robustness of this new ASG formulation, many new avenues for future research utilizing this technique exist. In particular, these include the development of robust posteriori error estimators, extension of the adaptive stencil formulation to three dimensions, and the development of a multilevel preconditioner (geometric or algebraic) which is efficient and robust for this particular ASG discretization.

## 7. Conclusion

[70] In this paper, we developed a new finite difference staggered grid stencil for discretizing the incompressible Stokes flow equations on block-structured adaptive meshes. The development of this particular discretization was spawned by the proven robustness of the classical, nonadaptive staggered grid discretization for studying lithospheric dynamics and the clear demonstration that adaptive spatial discretizations are an extremely useful technique within the domain of computational geodynamics.

[71] The adaptive staggered grid (ASG) discretization introduced here was rigorously demonstrated to be both numerically stable and possesses discretization error characteristics which closely resemble those of the classical nonadaptive staggered grid method. Specifically, we showed that the adaptive stencil does not introduce any spurious pressure oscillations when transitions in grid resolution occur, and in particular when such transitions intersect discontinuous viscosity structures. Furthermore, we verified using several variable viscosity reference problems (with known analytic solutions) that the adaptive stencil is second-order accurate in velocity and pressure (in the $L_1$ norm) for cases with continuous viscosity structures and reduces to first-order accuracy in both variables when strongly discontinuous viscosity structures are present. A number of pragmatic error indicators were examined to demonstrate the superiority of the ASG method over the nonadaptive staggered grid approach.

## Appendix A: Hanging Node Schemes Tested

[72] As mentioned in section 2.2, several variants of the constraint used at slave nodes were tested. In order to identify which scheme yielded the most accurate results, we have measured the grid convergence obtained using each of the different constraint formulations. The results of these tests are depicted in Figure A1 and enabled us to conclude that the combination of equations (41) and (42) with an arithmetic viscosity interpolation (Figure A1a) produced the most reliable results. The different forms of the constraints are labeled from A to F and their implementation is explained below.

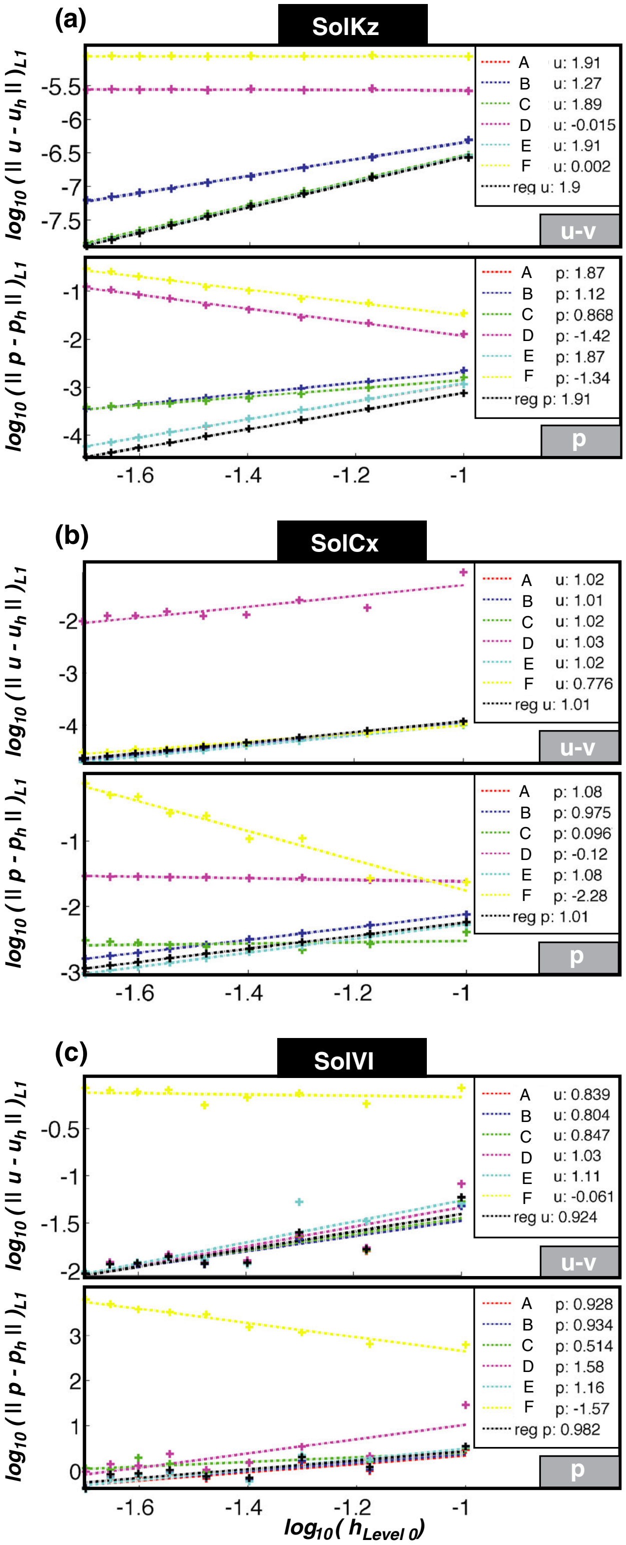

**Figure A1.** Grid convergence tests for the different hanging treatments. The color dashed lines (labeled A to F within each panel) correspond the hanging nodes schemes listed in Appendix A, and the black dashed line represent the nonadaptive mesh case. These tests were carried out for the (a) SolKz, (b) SolCx, and (c) SolVI. The adaptive grid described in Figure 3a was employed.

[73] 1. Arithmetic viscosity mean:
Define the average viscosity as

$$\hat{\eta} := \left(\eta^{\mathrm{F}} + \eta^{\mathrm{J}}\right) \tag{A1}$$

and utilize the following constrain equations,

$$\hat{\eta}\left(\frac{v_y^{\mathrm{H}} - v_y^{\mathrm{G}}}{\frac{1}{4}\Delta x}\right) - \eta^{F}\left(\frac{v_y^{\mathrm{G}} - v_y^{\mathrm{E}}}{\frac{1}{2}\Delta x}\right) - \eta^{J}\left(\frac{v_y^{\mathrm{K}} - v_y^{\mathrm{I}}}{\frac{3}{4}\Delta x}\right) = 0, \tag{A2}$$

$$\hat{\eta}\left(\frac{v_y^{\mathrm{I}} - v_y^{\mathrm{H}}}{\frac{1}{4}\Delta x}\right) - \eta^{F}\left(\frac{v_y^{\mathrm{G}} - v_y^{\mathrm{E}}}{\frac{1}{2}\Delta x}\right) - \eta^{J}\left(\frac{v_y^{\mathrm{K}} - v_y^{\mathrm{I}}}{\frac{3}{4}\Delta x}\right) = 0. \tag{A3}$$

[74] 2. Harmonic viscosity mean:
Define the average viscosity as

$$\hat{\eta} := \frac{2}{\frac{1}{2}\left(1/\eta^{F} + 1/\eta^{J}\right)} \tag{A4}$$

and use equations (A2) and (A3) to define the constraints.

[75] 3. Geometric viscosity mean:
Define the average viscosity as

$$\hat{\eta} := 2\exp\left[\frac{1}{2}\left(\ln(\eta^{\mathrm{F}}) + \ln(\eta^{\mathrm{J}})\right)\right] \tag{A5}$$

and use equations (A2) and (A3) to define the constraints.

[76] 4. Direct velocity gradient interpolation:

$$2\left(\frac{v_y^{\mathrm{H}} - v_y^{\mathrm{G}}}{\frac{1}{4}\Delta x}\right) - \left(\frac{v_y^{\mathrm{G}} - v_y^{\mathrm{E}}}{\frac{1}{2}\Delta x}\right) - \left(\frac{v_y^{\mathrm{K}} - v_y^{\mathrm{I}}}{\frac{3}{4}\Delta x}\right) = 0, \tag{A6}$$

$$2\left(\frac{v_y^{\mathrm{I}} - v_y^{\mathrm{H}}}{\frac{1}{4}\Delta x}\right) - \left(\frac{v_y^{\mathrm{G}} - v_y^{\mathrm{E}}}{\frac{1}{2}\Delta x}\right) - \left(\frac{v_y^{\mathrm{K}} - v_y^{\mathrm{I}}}{\frac{3}{4}\Delta x}\right) = 0. \tag{A7}$$

[77] 5. No velocity gradient:

$$v_y^{\mathrm{H}} - v_y^{\mathrm{G}} = 0, \tag{A8}$$

$$v_y^{\mathrm{I}} - v_y^{\mathrm{H}} = 0. \tag{A9}$$

[78] 6. Local viscosity interpolation (no volume flux control between the resolution levels):
Making the following definitions

$$\eta_{\alpha} := \left(\frac{5}{4}\eta^{\mathrm{F}} + \frac{3}{4}\eta^{\mathrm{J}}\right) \tag{A10}$$

$$\eta_{\beta} := \left(\frac{3}{4}\eta^{\mathrm{F}} + \frac{5}{4}\eta^{\mathrm{J}}\right) \tag{A11}$$

we have

$$\eta_{\alpha}\left(\frac{v_y^{\mathrm{H}} - v_y^{\mathrm{G}}}{\frac{1}{4}\Delta x}\right) - \frac{5}{4}\eta^{\mathrm{F}}\left(\frac{v_y^{\mathrm{G}} - v_y^{\mathrm{E}}}{\frac{1}{2}\Delta x}\right) - \frac{3}{4}\eta^{\mathrm{J}}\left(\frac{v_y^{\mathrm{K}} - v_y^{\mathrm{I}}}{\frac{3}{4}\Delta x}\right) = 0, \tag{A12}$$

$$\eta_{\beta}\left(\frac{v_y^{\mathrm{I}} - v_y^{\mathrm{H}}}{\frac{1}{4}\Delta x}\right) - \frac{3}{4}\eta^{\mathrm{F}}\left(\frac{v_y^{\mathrm{G}} - v_y^{\mathrm{E}}}{\frac{1}{2}\Delta x}\right) - \frac{5}{4}\eta^{\mathrm{J}}\left(\frac{v_y^{\mathrm{K}} - v_y^{\mathrm{I}}}{\frac{3}{4}\Delta x}\right) = 0. \tag{A13}$$

## Appendix B: Solvability and Stability of the Discretization

[79] Following the application of ASG discretization, the discrete form of the Stokes equations can be expressed as

$$\begin{bmatrix} A & G \\ D & 0 \end{bmatrix}\begin{bmatrix} v \\ p \end{bmatrix} = \begin{bmatrix} F \\ 0 \end{bmatrix} \longrightarrow \mathcal{A}x = \mathcal{F}, \tag{B1}$$

where $A$ is the discrete gradient of the stress tensor, $G$ is the discrete gradient operator, $D$ is the discrete divergence operator, and $F$ is the discrete forcing term. We note that the discrete operators $A, G, D$ and the right-hand side $F$ have been modified to enforce the velocity Dirichlet boundary conditions.

[80] Under the assumption that $A$ is nonsingular, it follows from a block decomposition of equation (B1) that $\mathcal{A}$ is nonsingular if and only if the Schur complement, $S = DA^{-1}G$, is also invertible. In the context of a spatial discretization for Stokes flow, simply demonstrating that $\mathcal{A}$ is nonsingular is not sufficient to guarantee that the discrete velocity and pressure solution is physically meaningful. To ensure the discrete solution is well posed (i.e., stable), it is required that $\mathcal{A}$ is "uniformly invertible" as the discretization parameter $h$ (i.e., the cell size) approaches zero. This implies that the distribution of eigenvalues (or some measure of an appropriate condition number) remains bounded as $h \rightarrow 0$. Consequently, this further implies $S$ must also be uniformly invertible as $h \rightarrow 0$. If the discretization is not stable, artifacts in the pressure field will be observed in the form of spurious pressure modes—or the so called "checkerboard" instability. Hence, verifying that the ASG scheme is stable is essential to establish the robustness of this spatial discretization for Stokes flow.

[81] In the context of a finite element method, the stability of the discretization applied to Stokes equations is given by the Ladyženskaja-Babuška-Brezzi (LBB) or inf-sup condition [*Babuška*, 1973; *Brezzi*, 1974]. The inf-sup condition has been formally proven for the classical staggered grid finite difference method [*Shin and Strikwerda*, 1997]. Demonstration of the LBB condition can be achieved via a formal mathematical proof, or via numerical computations [*Malkus*, 1981; *Chapelle and Bathe*, 1993; *Bathe*, 2001]. Given the complexity of the grid structure employed in the adaptive finite difference stencil, here we adopt a numerical approach to verify that the LBB condition is satisfied.

**Table B1.** Extremal Eigenvalues $\lambda$ for a Sequence of Different Adaptive Meshes[a]

| #pdof | $\min(\mathcal{R}e(\lambda))$ | $\max(\mathcal{R}e(\lambda))$ | $\min(\mathcal{I}m(\lambda))$ | $\max(\mathcal{I}m(\lambda))$ |
|---|---|---|---|---|
| 400* | 9.999009311391e-06 | 2.000000000000e+01 | –4.140150165238e-15 | 4.140150165238e-15 |
| 640 | 9.341960940568e-06 | 2.000000000000e+01 | –1.009197067266e-02 | 1.009197067266e-02 |
| 1162* | 9.338930025513e-06 | 2.000000000000e+01 | –5.747449164543e-02 | 5.747449164543e-02 |
| 2092 | 9.338346298167e-06 | 2.000000000000e+01 | –6.398934274588e-02 | 6.398934274588e-02 |
| 2284 | 9.338346299333e-06 | 2.000000000000e+01 | –6.442340005421e-02 | 6.442340005421e-02 |
| 2584 | 9.338346307213e-06 | 2.000000000000e+01 | –6.442329265390e-02 | 6.442329265390e-02 |
| 3094 | 9.338346300710e-06 | 2.000000000000e+01 | –4.922797015270e-02 | 4.922797015270e-02 |
| 3550 | 9.338346303195e-06 | 2.000000000000e+01 | –3.582211897156e-02 | 3.582211897156e-02 |
| 3820* | 9.338346305461e-06 | 2.000000000000e+01 | –5.874622706523e-02 | 5.874622706523e-02 |

[a] Meshes marked with (*) are displayed in Figure B1.

[82] We verified the stability of the ASG discretization numerically by examining whether $S$ is uniformly invertible. To this end, we computed the eigenvalues of $S$ associated with a number of different adaptive meshes used to solve the variable viscosityreference problem SolCx. The spectrums obtained and the meshes used for each test are shown in Figure B1. From Figure B1, we observed that the eigenvalue spectrum on the adaptive meshes is largely independent of $h$. The extremal eigenvalues are summarized in Table B1. From these results, we see that both the minimum and maximum real eigenvalues are bounded away from zero and independent of the mesh used. We also observe that the uniform mesh (#pdof = 400) has eigenvalues with a small complex part which is $\sim O(10^{-15})$, whilst the adaptive meshes have eigenvalues with a larger complex part $\sim O(10^{-2})$. Nevertheless, the magnitude of the complex part of the eigenvalues does not

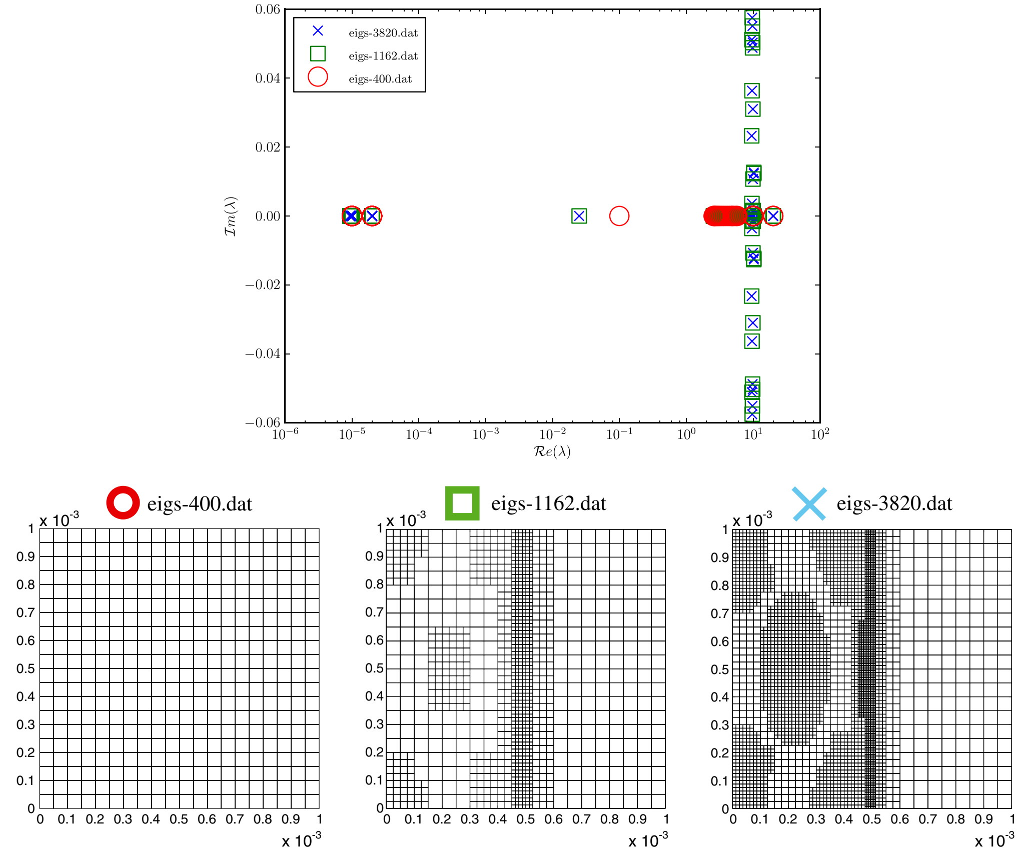

Figure B1. Eigenvalue spectrum for the Schur complement $DA^{-1}G$ for a number of different adaptive meshes.

**Table B2.** Extremal Singular Values $s$ for a Sequence of Different Adaptive Meshes[a]

| #pdof | min($s$) | max($s$) |
|---|---|---|
| 400* | 9.999009318452e-06 | 2.002518522834e+01 |
| 640 | 9.194141720418e-06 | 2.000360552629e+01 |
| 1162 | 8.147117276074e-06 | 2.000988329822e+01 |
| 2092 | 6.483747171417e-06 | 2.004073213973e+01 |

[a] Meshes marked with (*) are displayed in Figure B1.

fluctuate significantly on the different adaptive meshes examined.

[83] In general, the adaptive staggered grid discretization will give rise to a nonsymmetric Schur complement. Thus, to ascertain whether $S$ is uniformly invertible, it is of interest to examine the singular values of $S$. In Table B2, we summarize the singular values of $S$ obtained from the same sequence of adaptive meshes used in Table B1. We observe that the maximum singular value is bounded independent of $h$; however, we note that the minimum singular value is slightly decreasing with decreasing cell size $h$. The approximately mesh-independent behavior of both the eigenvalues and singular values of the Schur complement indicates that the ASG discretization results in a uniformly invertible operator $\mathcal{A}$, and thus, the discretization is stable in the regime of practical interest.

## Appendix C: A Priori Error Estimates

[84] The theoretical error estimates for the order of accuracy of the ASG discretization applied to Stokes flow are defined in terms of the $L_2$ norm, which for scalar function $q$ and vector function $\boldsymbol{w}$ are given by

$$\|q\|^2_{L_2(\Omega)} = \int_\Omega q^2 \, \mathrm{d}V, \qquad \|\boldsymbol{w}\|^2_{L_2(\Omega)} = \int_\Omega \left(w_x^2 + w_y^2\right) \mathrm{d}V. \tag{C1}$$

In addition we also require the $H^1$ norm, for a vector valued function $\boldsymbol{v}$ which is given by

$$\begin{aligned} \|\boldsymbol{v}\|^2_{H^1(\Omega)} = {} & \|v_x\|^2_{L_2(\Omega)} + \|v_y\|^2_{L_2(\Omega)} \\ & + \left\|\frac{\partial v_x}{\partial x}\right\|^2_{L_2(\Omega)} + \left\|\frac{\partial v_x}{\partial y}\right\|^2_{L_2(\Omega)} \\ & + \left\|\frac{\partial v_y}{\partial x}\right\|^2_{L_2(\Omega)} + \left\|\frac{\partial v_y}{\partial y}\right\|^2_{L_2(\Omega)}. \end{aligned} \tag{C2}$$

[85] Under the condition that the solution is sufficiently smooth, the optimal error bounds for the discrete Stokes problem are

$$\|\boldsymbol{v} - \boldsymbol{v}_h\|_{H^1(\Omega)} + \|p - p_h\|_{L_2(\Omega)} \le Ch, \tag{C3}$$

where $h$ is the control volume cell size, $\boldsymbol{v}, p$ represent the exact velocity and pressure solution, $\boldsymbol{v}_h, p_h$ represent the approximate velocity and pressure, and $C$ is a constant independent of the cell size $h$. We refer to $C$ for the definition of the $L_2$ and $H^1$ norms. Using the Aubin-Nitsche method, equation (C3) can be expressed in terms of $L_2$ norms

$$\|\boldsymbol{v} - \boldsymbol{v}_h\|_{L_2(\Omega)} \le C_1 h^2, \qquad \|p - p_h\|_{L_2(\Omega)} \le C_2 h. \tag{C4}$$

[86] For comparison with *Duretz et al.* [2011], we rewrite equation (C4) in terms of the $L_1$ norms which for scalar function $w$ and vector function $\boldsymbol{w}$ are given by

$$\|\phi\|_{L_1(\Omega)} = \int_\Omega |\phi| \, \mathrm{d}V, \qquad \|\boldsymbol{w}\|_{L_1(\Omega)} = \int_\Omega (|w_x| + |w_y|) \, \mathrm{d}V. \tag{C5}$$

Applying of Schwarz inequality to equation (C4) yields

$$\begin{aligned} \|\boldsymbol{v} - \boldsymbol{v}_h\|_{L_1(\Omega)} \le \|\boldsymbol{v} - \boldsymbol{v}_h\|_{L_2(\Omega)} \le C_3 h^2, \\ \|p - p_h\|_{L_1(\Omega)} \le \|p - p_h\|_{L_2(\Omega)} \le C_4 h, \end{aligned} \tag{C6}$$

where $C_1, C_2, C_3, C_4$ are constants independent of the grid size $h$. We re-emphasize that throughout this work, all of the reported errors computed via the $L_1$ norm have been normalized by the volume of the domain such that the error measured possesses the same physical dimensions as the original quantity.

[87] Here we describe how the integrals in equations (C6) were evaluated over the domain discretized via the adaptive staggered grid stencil. All integrals are evaluated using a one-point quadrature rule, e.g., given a scalar function $f(\boldsymbol{x})$, we make the following approximation

$$\int_\Omega f \, \mathrm{d}V \approx \sum_k f^k \, \Delta\Omega^k \tag{C7}$$

where $k$ indicates a particular grid point in the finite difference stencil, $f^k$ is the value of $f$ at $\boldsymbol{x}_k$, and $\Delta\Omega^k$ is the volume associated with grid point $k$. Each

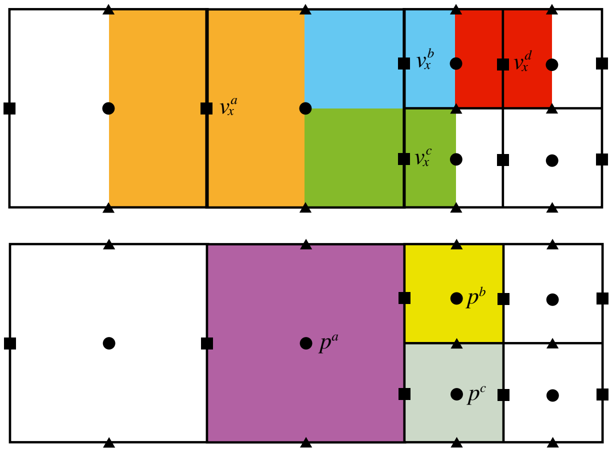

**Figure C1.** Integration domains ($\Delta\Omega_k$, denoted by colors) associated with different stencil points within adaptive staggered grid discretization. (top) Integration domains for the $x$ component of the velocity. (bottom) Integration domains for the pressure unknowns.

discretization point for $u_x, u_y$, and $p$ is located at different points in space; thus, the definition of $\Delta\Omega^k$ is different for each component of the velocity field and pressure.

[88] To define integrals involving the pressure space (cell centered), $\Delta\Omega_k$ is defined as the volume of the cell associated with pressure degree of freedom (see Figure C1, bottom). Components of vector quantities in the staggered grid formulation are discretized at different locations. For example, given $\boldsymbol{w} = (w_x, w_y)$, the $x$ components are located at the vertical cell faces and the $y$ components are located at the horizontal cell faces (see Figure C1, top). Thus, the individual components of a vector are integrated using a different set of volumes. We denote this via

$$\int_\Omega w_x \, \mathrm{d}V \approx \sum_{\alpha=1}^{N_u} w_x^\alpha \, \Delta\Omega_x^\alpha, \qquad \int_\Omega w_y \, \mathrm{d}V \approx \sum_{\beta=1}^{N_v} w_y^\beta \, \Delta\Omega_y^\beta, \tag{C8}$$

where $N_u, N_v$ are the number of grid points used to discretize the $x, y$ components of the vector, respectively. In Figure C1 (top), we denote via different colors, the domains $\Delta\Omega_x^\alpha$ for four different stencil points. We note that the integration domains $\Delta\Omega_y^\beta$ associated with the $y$ component of velocity are defined in an analogous manner. Hence, from the definition of the $L_1$ norm of a vector quantity $\boldsymbol{w}$, we approximate $\|\boldsymbol{w}\|_{L_1(\Omega)}$ according to

$$\|\boldsymbol{w}\|_{L_1(\Omega)} \approx \sum_{\alpha=1}^{N_u} \left| w_x^\alpha \right| \, \Delta\Omega_x^\alpha \quad + \quad \sum_{\beta=1}^{N_v} \left| w_y^\beta \right| \, \Delta\Omega_y^\beta. \tag{C9}$$

# Appendix D: Adaptivity Indicator Definitions

[89] For each of the adaptation strategy employed in section 4, adaptivity evolves according to equation (47). However, the manner in which $\Phi_{\rm max}$ and $\bar{\Phi}$ are computed differs depending on the nature of the field which is selected to drive adaptivity.

[90] When adaptivity is driven by velocity ($\Phi \equiv v_{\rm err}$) and pressure ($\Phi \equiv p_{\rm err}$) errors, each velocity component and pressure error are calculated cell-wise. The absolute value of the velocity error is evaluated on all cell edges. On each pressure cell $k$, we compute

$$\left(\bar{v}_x^{\rm err}\right)_k = \frac{1}{n_u} \sum_{\alpha=1}^{n_u} \left\| \left(v_x^{\rm err}\right)_\alpha \right\|; \quad \left(\bar{v}_y^{\rm err}\right)_k = \frac{1}{n_v} \sum_{\alpha=1}^{n_v} \left\| \left(v_y^{\rm err}\right)_\alpha \right\|, \tag{D1}$$

where $n_u, n_v$ are the number of edges associated with cell $k$ for the $x, y$ velocity components, respectively, $v_i^{\rm err} = v_i^{\rm exact} - v_i^{\rm approx}$, and $\left(v_i^{\rm err}\right)_\alpha$ are the velocity error components associated with cell $k$. The maximum value is defined as

$$\Phi_{\rm max} := \max \left[ \max_{k=1,n_p} \left[ \left(\bar{v}_x^{\rm err}\right)_k \right], \max_{k=1,n_p} \left[ \left(\bar{v}_y^{\rm err}\right)_k \right] \right], \tag{D2}$$

where $n_p$ is the number of pressure control volumes in the mesh. We define the mean value as

$$\bar{\Phi} := \frac{1}{n_p} \sum_{k=1}^{n_p} \left(\bar{v}_x^{\rm err}\right)_k + \left(\bar{v}_y^{\rm err}\right)_k . \tag{D3}$$

The pressure error $p^{\rm err} = p^{\rm exact} - p^{\rm approx}$ is evaluated directly at cell centers $\left(p_k^{\rm err}\right)$ and its maximum is thus defined such that

$$\Phi_{\rm max} := \max_{k=1,n_p} \left\| p_k^{\rm err} \right\|, \tag{D4}$$

with a mean value given by

$$\bar{\Phi} := \frac{1}{n_p} \sum_{k=1}^{n_p} \left\| p_k^{\rm err} \right\| . \tag{D5}$$

[91] Adaptivity governed by velocity differences ($\Phi \equiv \Delta_v$) uses values which are computed at cell centers and cell vertices. To this end, we define velocity differences on a cell $k$ via

$$\left(\bar{\Delta} v_x\right)_k = \left| v_x^{\rm E}(k) - v_x^{\rm W}(k) \right|; \qquad \left(\bar{\Delta} v_y\right)_k = \left| v_y^{\rm S}(k) - v_y^{\rm N}(k) \right|, \tag{D6}$$

where the superscripts E, W, S, and N indicate the directions east, west, south, and north, respectively. We construct nodal values for velocity $(\Delta v_x)_i, (\Delta v_y)_i$ at each vertex $i$ by computing the absolute value of the difference of velocity components on the edges connected to vertex $i$. Then the maximum is defined as

$$\Phi_{\max} := \max\left[\max_{k=1,n_p}[(\bar{\Delta} v_d)_k],\ \max_{i=1,n_{\mathrm{vert}}}[(\Delta v_d)_i]\right], \quad \text{for } d = \{x, y\}, \tag{D7}$$

and the mean by

$$\bar{\Phi} := \frac{1}{n_p + n_{\mathrm{vert}}}\left(\sum_{k=1}^{n_p}\left((\bar{\Delta} v_x)_k + (\bar{\Delta} v_y)_k\right) + \sum_{i=1}^{n_{\mathrm{vert}}}\left((\Delta v_x)_i + (\Delta v_y)_i\right)\right), \tag{D8}$$

where $n_{\mathrm{vert}}$ corresponds to the number of vertices in the mesh.

[92] Pressure differences were also employed to control mesh adaptivity ($\Phi \equiv \Delta_p$). In this case, the absolute value of the pressure differences was calculated at each cell edge $e$, using the two neighboring cells $c_1, c_2$, which share this edge. Thus, we define the edge pressure difference as

$$\Delta p_e = \|p_{c_1} - p_{c_2}\|. \tag{D9}$$

Subsequently the maximum is defined via

$$\Phi_{\max} := \max_{e=1,n_{\mathrm{edge}}}[\Delta p_e], \tag{D10}$$

where $n_{\mathrm{edge}}$ stands for the total number of edges in the mesh. The mean values of the absolute pressure differences are given by

$$\bar{\Phi} := \frac{1}{n_{\mathrm{edge}}}\sum_{e=1}^{n_{\mathrm{edge}}}\Delta p_e. \tag{D11}$$

[93] For adaptivity based on material properties, we used both the density ($\Phi \equiv \Delta_\rho$) and viscosity ($\Phi \equiv \Delta_{\log_{10}\eta}$) field. Absolute density differences are calculated for each cell and characterize density differences existing between the cell center $k$ and the four vertices $i = 1, \ldots, 4$. The density difference between cell $k$ and the its four vertices is denoted via

$$\Delta\rho_{k,i} = \|\rho_k - \hat{\rho}_i\|, \tag{D12}$$

where $\rho_k$ is the density at cell $k$ and $\hat{\rho}_i$ is the density defined at the $i$th node associated with cell $k$. The maximum value is calculated as

$$\Phi_{\max} := \max_{\substack{k=1,n_p\\ i=1,4}}\left[\Delta\rho_{k,i}\right] \tag{D13}$$

and the mean value such as

$$\bar{\Phi} := \frac{1}{4n_p}\sum_{k=1}^{n_p}\sum_{i=1}^{4}\Delta\rho_{k,i}. \tag{D14}$$

Similarly, viscosity variations are evaluated between cell centers and cell vertices using

$$\Delta\eta_{k,i} = \left\|\log_{10}\left(\frac{\eta_k}{\hat{\eta}_i}\right)\right\| \tag{D15}$$

where $\eta_k$ is the viscosity at cell $k$ and $\hat{\eta}_i$ is the viscosity defined at the $i$th node associated with cell $k$. The maximum and mean values are obtained by replacing $\Delta\rho_{k,i}$ with $\Delta\eta_{k,i}$ in equations (D13) and (D14), respectively.

## Acknowledgments

Author D.A.M thanks Matt Knepley for the informative discussion related to the inf-sup stability condition associated with finite difference stencils. Marcin Dabrowski and Alik Ismail-Zadeh are thanked for providing in-depth reviews of this work.

## References

Aksoy, H., and C.-J. Chen (1992), Numerical solution of Navier-Stokes equations with non-staggered grids using finite analytic method, *Numer. Heat Transfer, Part B: Fundam.*, *21*(3), 287–306.

Albers, M. (2000), A local mesh refinement multigrid method for 3-D convection problems with strongly variable viscosity, *J. Comput. Phys.*, *160*(1), 126–150.

Armfield, S. (1991), Finite difference solutions of the Navier-Stokes equations on staggered and non-staggered grids, *Comput. Fluids*, *20*(1), 1–17.

Babeyko, A. Y., S. V. Sobolev, R. B. Trumbull, O. Oncken, and L. L. Lavier (2002), Numerical models of crustal-scale convection and partial melting beneath the Altiplano-Puna plateau, *Earth Planet. Sci. Lett.*, *199*, 373–388.

Babuška, I. (1973), The finite element method with Lagrangian multipliers, *Numer. Math.*, *20*, 179–192.

Bathe, K.-J. (2001), The inf-sup condition and its evaluation for mixed finite element methods, *Comput. Struct.*, *79*, 243–252.

Beaumont, C., R. A. Jamieson, M. Nguyen, and B. Lee (2001), Himalayan tectonics explained by extrusion of a low-viscosity crustal channel coupled to focused surface denudation, *Nature*, *414*, 738–742.

Brezzi, F. (1974), On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers, *R. A. I. R. O., Anal. Numer.*, *R2*, 129–151.

Buiter, S. J. H., A. Y. Babeyko, S. Ellis, T. V. Gerya, B. J. P. Kaus, A. Kellner, G. Schreurs, and Y. Yamada (2006), The numerical sandbox: Comparison of model results for a shortening and an extension experiment, in *Analogue and Numerical Modelling of Crustal-Scale Processes, Special Publication—Geological Society of London*, vol. 253, Geological Society of London, London, 29–64.

Burov, E., L. Jolivet, L. Le Pourhiet, and A. Poliakov (2001), A thermomechanical model of exhumation of high pressure (HP) and ultra-high pressure (UHP) metamorphic rocks in Alpine-type collision belts, *Tectonophysics*, *342*, 113–136.

Burov, E. B. (2011), Rheology and strength of the lithosphere, *Mar. Pet. Geol.*, *28* (8), 1402–1443, doi:10.1016/j.marpetgeo.2011.05.008.

Chapelle, D., and K.-J. Bathe (1993), The inf-sup test, *Comput. Struct.*, *47*(4/5), 537–545.

Chern, I. L., and Y. C. Shur (2007), A coupling interface method for elliptic interface problems), *J. Comput. Phys.*, *225*(2), 2138–2174.

Choi, S. K., H. Y. Nam, and M. Cho (1994a), Use of staggered and nonstaggered grid arrangements for incompressible flow calculations on nonorthogonal grids, *Numer. Heat Transfer, Part B: Fundam.*, *25*(2), 193–204.

Choi, S. K., H. Y. Nam, and M. Cho (1994b), Systematic comparison of finite-volume calculation methods with staggered and non-staggered grid arrangements, *Numer. Heat Transfer, Part B: Fundam.*, *25*(2), 205–221.

Crameri, F., P. J. Tackley, I. Meilick, T. V. Gerya, and B. J. P. Kaus (2012a), A free plate surface and weak oceanic crust produce single-sided subduction on Earth, *Earth Planet. Sci. Lett.*, *39*, L03,306.

Crameri, F., H. Schmeling, G. J. Golabek, T. Duretz, R. Orendt, S. J. H. Buiter, D. A. May, B. J. P. Kaus, T. V. Gerya, and P. J. Tackley (2012b), A comparison of numerical surface topography calculations in geodynamic modelling: An evaluation of the “sticky air” method, *Geophys. J. Int.*, *189*(1), 38–54.

Crowder, H. J., and C. Dalton (1971), Errors in the use of nonuniform mesh systems, *J. Comput. Phys.*, *7*(1), 32–45, doi:10.1016/0021-9991(71)90047-7.

de Rivas, E. (1972), On the use of nonuniform grids in finite-difference equations, *J. Comput. Phys.*, *10* (2), 202–210, doi:10.1016/0021-9991(72)90060-5.

Deubelbeiss, Y., and B. J. P. Kaus (2008), Comparison of Eulerian and Lagrangian numerical techniques for the stokes equations in the presence of strongly varying viscosity, *Phys. Earth Planet. Inter.*, *171* (1-4), 92–111, doi:10.1016/j.pepi.2008.06.023.

Duretz, T., D. A. May, T. V. Gerya, and P. J. Tackley (2011), Discretization errors and free surface stabilization in the finite difference and marker-in-cell method for applied geodynamics: A numerical study, *Geochem. Geophys. Geosyst.*, *12*, Q07,004.

Fedkiw, R. P., T. Aslam, B. Merriman, and S. Osher (1999), A non-oscillatory Eulerian approach to interfaces in multimaterial flows (the ghost fluid method), *J. Comput. Phys.*, *152*(2), 457–492.

Fullsack, P. (1995), An arbitrary Lagrangian-Eulerian formulation for creeping flows and its application in tectonic models, *Geophys. J. Int.*, *120*, 1–23.

Gerya, T. V. (2010), *Introduction to Numerical Geodynamic Modelling*, 345 pp., Cambridge University Press The Edinburgh Building, Cambridge CB2 8RU, UK.

Gerya, T. V., and D. A. Yuen (2003), Charaterictics-based marker method with conservative finite-difference schemes for modeling geological flows with strongly variable transport properties, *Phys. Earth Planet. Inter.*, *140*(4), 293–318.

Gerya, T. V., and D. A. Yuen (2007), Robust characteristics method for modelling multiphase visco-elasto-plastic thermo-mechanical problems, *Phys. Earth Planet. Inter.*, *163*, 83–105.

Gerya, T. V., L. L. Perchuk, D. D. van Reenen, and C. A. Smit (2000), Two-dimensional numerical modeling of pressure-temperature-time paths for the exhumation of some granulite facies terrains in the Precambrian, *J. Geodyn.*, *29*, 17–35.

Gerya, T. V., D. A. Yuen, and W. V. Maresch (2004), Thermomechanical modeling of slab detachment, *Earth Planet. Sci. Lett.*, *226*, 101–116.

Golabek, G. J., T. V. Gerya, B. J. P. Kaus, R. Ziethe, and P. J. Tackley (2009), Rheological controls on the terrestrial core formation mechanism, *Geochem. Geophys. Geosyst.*, *10*, Q11,007.

Gorczyk, W., T. V. Gerya, J. A. Connolly, and D. A. Yuen (2007), Growth and mixing dynamics of mantle wedge plumes, *Geology*, *35*, 587–590.

Harlow, F. H., and E. Welch (1965), Numerical calculation of time-dependent viscous flow of fluid with free surface, *Phys. Fluids*, *8*(12), 2182–2189.

Hirt, C. W., A. A. Amsden, and J. L. Cook (1974), An arbitrary Lagrangian-Eulerian computing method for all flow speeds, *J. Comput. Phys.*, *14*(3), 227–253.

Huismans, R. S., and C. Beaumont (2002), Asymmetric lithospheric extension: The role of frictional plastic strain softening inferred from numerical experiments, *Geology*, *30*, 211–214.

Lavier, L. L., W. R. Buck, and A. N. B. Poliakov (2000), Factors controlling normal fault offset in an ideal brittle layer, *J. Geophys. Res.*, *105*, 23,431–23,442.

Leveque, R. J., and Z. Li (1997), Immersed interface methods for Stokes flow with elastic boundaries or surface tension, *SIAM J. Sci. Comput.*, *18*(3), 709–735.

Lin, J.-R., T. V. Gerya, P. Tackley, and D. Yuen (2009), Numerical modeling of protocore destabilization during planetary accretion: Methodology and results, *Icarus*, *204*, 732–748.

Majumdar, S. (1988), Role of underrelaxation in momentum interpolation for calculation of flow with nonstaggered grids, *Numer. Heat Transfer*, *13*(1), 125–132.

Malkus, D. S. (1981), Eigenproblems associated with the discrete LBB condition for incompressible finite elements, *Int. J. Engng. Sci.*, *19*, 1299–1310.

Matsumoto, T., and Y. Tomoda (1983), Numerical simulation of the initiation of subduction at the fracture zone, *J. Phys. Earth*, *31*, 183–194.

Melaaen, M. C. (1992a), Calculation of fluid flows with staggered and non-staggered curvilinear non-orthogonal grids—The theory, *Numer. Heat Transfer, Part B: Fundam.*, *21*(1), 1–19.

Melaaen, M. C. (1992b), Calculation of fluid flows with staggered and non-staggered curvilinear non-orthogonal grids—A comparison, *Numer. Heat Transfer, Part B: Fundam.*, *21*(1), 21–39.

Miller, T. F., and F. W. Schmidt (1988), Use of a pressure-weighted interpolation method for the solution of the incompressible Navier-Stokes equations on a non-staggered grid system, *Numer. Heat Transfer*, *14*(2), 213–233.

Mitrovica, J. X., and A. M. Forte (1998), New insights obtained from joint inversions for the radial profile of mantle viscosity, *Phys. Earth Planet. Inter.*, *23*(9-10), 857–863, doi:10.1016/S0079-1946(98)00105-0.

Moresi, L., F. Dufour, and H.-B. Mühlhaus (2003), A Lagrangian integration point finite element method for large deformation modeling of viscoelastic geomaterials, *J. Comput. Phys.*, *184*, 476–497.

Moresi, L., S. Quenette, V. Lemiale, C. Meriaux, B. Appelbe, and H. B. Mühlhaus (2007), Computational approaches to studying non-linear dynamics of the crust and mantle, *Phys. Earth Planet. Inter.*, *163*, 69–82.

Papageorgakopoulos, J., G. Arampatzis, D. Assimacopoulos, and N. C. Markatos (2000), Enhancement of the momentum interpolation method on non-staggered grids, *Int. J. Numer. Methods Fluids*, *33*, 1–22.

Patankar, S. V. (1980), *Numerical Heat Transfer and Fluid Flow*, 197 pp., Hemisphere, New York, U.S.A.

Perič, M., R. Kessler, and G. Scheuerer (1988), Comparison of finite-volume numerical methods with staggered and colocated grids, *Comput. Fluids*, *16*(4), 389–403.

Petersen, K. D., S. B. Nielsen, O. R. Clausen, R. Stephenson, and T. Gerya (2010), Small-scale mantle convection produces stratigraphic sequences in sedimentary basins, *Science*, *329*, 827–830.

Poliakov, A., and Y. Podladchikov (1992), Diapirism and topography, *Geophys. J. Int.*, *109*, 553–564.

Reggio, M., and R. Camarero (1986), Numerical solution procedure for viscous incompressible flows, *Numer. Heat Transfer*, *10*(2), 131–146.

Revenaugh, J., and B. Parsons (1987), Dynamic topography and gravity anomalies for fluid layers whose viscosity varies exponentially with depth, *Geophys. J. R. Astron. Soc.*, *90*, 349–368.

Rhie, C. M., and W. L. Chow (1983), Numerical study of the turbulent flow past an airfoil with trailing edge separation, *AIAA*, *21*(11), 1525–1532.

Rodi, W., S. Majumdar, and B. Schönung (1989), Finite volume methods for two-dimensional incompressible flows with complex boundaries, *Comput. Methods Appl. Mech. Engrg.*, *75*(1-3), 369–392.

Schmeling, H., et al. (2008), A benchmark comparison of spontaneous subduction models—Towards a free surface, *Phys. Earth Planet. Inter.*, *171*(1-4), 198–223.

Schmid, D. W. (2002), Finite and infinite heterogeneities under pure and simple shear, Ph.D. thesis, ETH Zurich, Zurich, Switzerland.

Schmid, D. W., and Y. Y. Podladchikov (2003), Analytical solutions for deformable elliptical inclusions in general shear, *Geophys. J. Int.*, *155*, 269–288.

Shih, T. M., C. H. Tan, and B. C. Hwang (1989), Effects of grid staggering on numerical schemes, *Int. J. Numer. Methods Fluids*, *9*(2), 193–212.

Shin, D., and J. C. Strikwerda (1997), Inf-sup conditions for finite difference approximations of the Stokes equations, *J. Austral. Math. Soc. Ser. B.*, *39*, 121–134.

Stacey, F. D., F. Rong-Shan, and S. Spiliopoulos (1989), Viscosity structure implied by mantle convection, *Phys. Earth Planet. Inter.*, *55* (1-2), 1–9, doi: 10.1016/0031-9201(89)90228-8.

Stadler, G., M. Gurnis, C. Burstedde, L. C. Wilcox, L. Alisic, and O. Ghattas (2010), The dynamics of plate tectonics and mantle flow: From local to global scales, *Science*, *329*, 1033–1038.

Suckale, J., J.-C. Nave, and B. H. Hager (2010), It takes three to tango: 1. Simulating buoyancy-driven flow in the presence of large viscosity contrasts, *J. Geophys. Res.*, *115*, B07,409.

Sundqvist, H., and G. Veronis (1970), A simple finite-difference grid with non-constant intervals, *Tellus*, *22* (1), 26–31, doi:10.1111/j.2153-3490.1970.tb01933.x.

Tackley, P. J. (2000), Self-consistent generation of tectonic plates in time-dependent, three-dimensional mantle convection simulations Part 1: Pseudo-plastic yielding, *Geochem. Geophys. Geosyst.*, *1*, 2000GC000,036.

van Hunen, J., A. van den Berg, and N. J. Vlaar (2000), A thermomechanical model of horizontal subduction below an overriding plate, *Earth Planet. Sci. Lett.*, *182*, 157–169.

Weinberg, R. F., and H. Schmeling (1992), Polydiapirs: Multiwave length gravity structures, *J. Struct. Geol.*, *14*, 425–436.

Zaleski, S., and P. Julien (1992), Numerical simulation of Rayleigh-Taylor instability for single and multiple salt diapirs, *Tectonophysics*, *206*(1-2), 55–69.

Zhong, S. (1996), Analytic solutions for Stokes' flow with lateral variations in viscosity, *Geophys. J. Int.*, *124*, 18–28.