Highly Scalable Geodynamic Simulations with HyTeG

Ponsuganth Ilangovan[1], Nils Kohl[1], and Marcus Mohr[1]

[1]Geophysics, Dept. of Earth and Environmental Sciences, LMU Munich

Correspondence: Ponsuganth Ilangovan (p.ilango@lmu.de)

Abstract. High-resolution geodynamic simulations of mantle convection are essential to quantitatively assess the complex physical mechanisms driving the large-scale tectonic processes that shape Earth's surface. Accurately capturing small-scale features such as unstable thermal boundary layers requires global resolution on the order of 1 km. This renders traditional sparse matrix methods impractical due to their prohibitively high memory demands and low arithmetic intensity. Matrix-free methods offer a scalable alternative, enabling the solution of large-scale linear systems efficiently. In this work, we leverage the matrix-free Finite Element framework HyTeG to conduct large-scale geodynamic simulations that incorporate realistic physical models. We validate the framework through a combination of convergence studies of the Finite Element approximations against analytical solutions and through geophysical community benchmarks. The latter include test cases with temperature-dependent and nonlinear rheologies. Our scalability studies demonstrate excellent performance, scaling up to problems with about 100 billion (10^{11}) unknowns in the Stokes system.

1 Introduction

Major geological activity such as earthquakes and volcanoes can ultimately be attributed to convection processes in the Earth's mantle. The latter gets heated from below by the Earth's core generating density differences that trigger movement of the rocky material in the mantle over geological timescales and in turn drive geological events (Davies, 1999). In addition, heat production resulting from the decay of radioactive materials and from frictional stresses due to the movement of material also contribute to the thermal state of the mantle.

Creating an accurate model of the Earth's mantle representing this convective process is an essential component for an improved understanding of the geological processes that shape the surface of our planet, such as the formation of mountain chains and oceanic trenches through plate tectonics, or the release of accumulated stresses during inter-plate earthquakes. Retrodiction of past mantle flow, see e.g. Colli et al. (2018), allows us to reconstruct dynamic topography, and is essential for determining the sedimentation record and to better constrain the mantle's rheology which is known only qualitatively.

On the large timescales involved, the mantle can be modelled as a fluid. The high viscosities and small velocities lead to an extremely small Reynolds number, which allows one to neglect the inertial terms in the Navier-Stokes equations and assume an instantaneous force balance. Hence, mantle convection models are usually based on a generalised form of the Stokes equations. Historically, simulations employed the Boussinesq approximation (Oberbeck, 1879; Boussinesq, 1903) to describe the convective process. However, refined models have been developed that can also take the compressibility of mantle material into

account, such as the (truncated) anelastic liquid approximation ((T)ALA), see e.g. Gassmöller et al. (2020) for an overview. A general framework for simulating mantle convection should provide the means to implement various formulations depending on the predilection of its users.

Since the 1980s various codes for the simulation of mantle convection in 2D and 3D settings have been designed in the community. See Baumgardner (1985); Blankenbach et al. (1989); Moresi and Solomatov (1995); Tan et al. (2006) for examples. Modern variants include ASPECT (Bangerth et al., 2023), built on top of the FE (Finite Element) framework deal.II, and approaches employing Firedrake as underlying FE framework (Davies et al., 2022; Ham et al., 2023).

Simulation of mantle convection is a computationally extremely challenging endeavour. The width of the Earth's mantle is on the order of $3,000\,\mathrm{km}$, while features of interest, such as the thermal boundary layers below the surface and above the Core-Mantle Boundary (CMB), short-wavelength asthenosphere dynamics (Brown et al., 2022), or rising plumes and sinking slabs are small-scale features of a few km. Resolving them, hence, requires high resolution, leading to large linear systems of equations. These often need to be solved multiple times per time-step, of which one often has to perform on the order of thousands as one wants to simulate millions of years of Earth's history. Mantle convection codes, thus, have been candidates to be run on supercomputers early on, see e.g. (Bunge and Baumgardner, 1995).

However, not only the mantle's sheer size is problematic; the strong local viscosity variations that develop in mantle convection models also pose a challenge for the solution process, and finding optimal iterative solvers is an ongoing research topic, see e.g. Clevenger and Heister (2021).

Classical FE analysis relies on assembling the system matrix of the linear system of equations and then working in terms of numerical linear algebra. However, such traditional matrix-based methods are not well-suited for truly large-scale high-performance computing, due to their immense memory requirements and performance-penalizing low arithmetic intensity. Instead, matrix-free methods, which do not assemble the system matrix explicitly, but rather provide a routine to apply the system matrix to a given vector, have been shown to be a viable alternative for large-scale simulations, see e.g. Kronbichler and Kormann (2012); Gmeiner et al. (2016). The memory savings are obvious (the system matrices stemming from the generalised Stokes equations using standard Taylor-Hood elements may have approximately 100-200 non-zero entries per row (Kohl and Rüde, 2022)), and the computational performance of matrix-free methods paired with low- to medium-order discretisations is at least on par with, or even better than that of standard sparse matrix methods, see, e.g., Kronbichler and Kormann (2012); Böhm et al. (2025).

In this work, we present studies performed with the mantle convection code TerraNeo, see Bauer et al. (2020); Ilangovan et al. (2024), which is designed to run models at extremely high resolution. TerraNeo is based on the matrix-free FE framework HyTeG (Kohl et al., 2019, 2024) and the associated automatic code-generator HOG (HyTeG Operator Generator) (Böhm et al., 2025). Our aim here is threefold. We are going to present results for various benchmark settings that are commonly employed in the community to assess the usability and correctness of mantle convection codes (Zhong et al., 2008; Davies et al., 2013, 2022; Euen et al., 2023), to show that these problems can successfully be solved with TerraNeo. As part of this we will also evaluate certain model simplifications and adaptations that can potentially improve performance in an HPC (High

Performance Computing) context. The third aspect is to demonstrate the scalability of TerraNeo when simulating at a peak global resolution of 3 km, resulting in $\simeq 10^{11}$ degrees of freedom (DoFs).

The remainder of this paper starts with a description of the model used to represent convection in the Earth's mantle, its spatial and temporal discretisation as well as some simplifications for an HPC setting. This is followed by a brief overview of HYTEG and HOG and the numerical solution process we employed, before we come to the central part of the paper that consists of a section on benchmarking and validation.

2 Model

On geological timescales, the mantle behaves like a highly viscous fluid. Its movement can thus be described by the Navier-Stokes equations. Due to the large values of viscosity ($> 10^{19}$ Pa s), the Ekman number (ratio of viscous to Coriolis forces) becomes very large, while the Reynolds number (ratio of inertial to viscous forces) becomes very small (Schubert et al., 2001). This allows one to neglect Coriolis and inertial forces, and conservation of mass and balance of forces can be modelled by the Stokes equations alone. The driving force of the system is buoyancy, which results from density differences in the mantle. These are mostly due to temperature gradients. Our model can, thus, be written as

$$-\nabla \cdot \boldsymbol{\tau} + \nabla p = \rho \mathbf{g} \tag{1}$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{2}$$

where $\mathbf{u}$ represents velocity, ρ density, p pressure, $\mathbf{g}$ the gravity vector and $\boldsymbol{\tau} = \eta \left(\nabla \mathbf{u} + \nabla \mathbf{u}^\top \right) - \frac{2}{3}\eta (\nabla \cdot \mathbf{u}) \mathbf{I}$ the deviatoric stress tensor including dynamic viscosity η. This problem is posed on a thick spherical shell,

$$\Omega = \Omega_{r_{\text{min}}, r_{\text{max}}} := \left\{ \boldsymbol{x} \in \mathbb{R}^3 \;\middle|\; r_{\text{min}} < \|\boldsymbol{x}\|_2 < r_{\text{max}} \right\} \;,$$

used to represent the geometry of Earth's mantle.

It is common practice in the community to work with a reformulation of the problem which involves a hydrostatic reference state. One then considers deviations from this reference state instead of the total physical quantities. In the following we will denote by $\bar{\square}$ a radial reference profile for quantity $\square$ and by $\square'$ deviations from the latter. We can then e.g. express pressure as $p = \bar{p} + p'$, where the hydrostatic pressure $\bar{p}$ satisfies $\nabla \bar{p} = \bar{\rho} \mathbf{g}$. As a next step we approximate total density by a first order Taylor expansion

$$\rho(p, T) \simeq \bar{\rho}(r) \left[1 + \bar{\kappa}_T(r) p' - \bar{\alpha}(r) T' \right] \;. \tag{3}$$

Here r denotes a point's distance from the center of the problem domain Ω, while $\bar{\kappa}_T$ and $\bar{\alpha}$ denote isothermal compressibility and thermal expansivity. Equation (3) constitutes one part of the so-called anelastic liquid approximation (ALA) (Gough, 1969). The other is to neglect the density variations ρ' in the continuity equation eq. (2), as these can be assumed to be small compared to $\bar{\rho}$. As the latter is constant in time, this also removes the time derivative from the equation, and, thus, also the

potential for sound waves in the model. By further dropping the pressure dependence in eq. (3), one arrives at the so called truncated anaelastic liquid approximation (TALA)

$$\rho(p,T) \simeq \bar{\rho}(r)\left[1-\bar{\alpha}(r)T'\right] \ . \tag{4}$$

With this the Stokes part of our model obtains its final form

$$-\nabla\cdot\boldsymbol{\tau}+\nabla p' = \bar{\rho}\bar{\alpha}T'\hat{\boldsymbol{r}} \tag{5}$$

$$\nabla\cdot(\bar{\rho}\mathbf{u}) = 0 \ , \tag{6}$$

where $\hat{\boldsymbol{r}}$ is a radial unit vector.

In addition to this, conservation of energy in the mantle needs to be considered (McKenzie and Jarvis, 1980). This can be written in terms of temperature T as primary quantity and couples to the Stokes system. Under the assumptions of the TALA, deviations of density from the reference profile $\bar{\rho}$ can be neglected and one obtains

$$\bar{\rho}\bar{c}_p\left(\frac{\partial T}{\partial t}-\nabla\left(k\nabla T\right)\right) = \boldsymbol{\tau}:\dot{\boldsymbol{\epsilon}}+H+\bar{\alpha}T\mathbf{u}\cdot\nabla p \ . \tag{7}$$

Here $\dot{\boldsymbol{\epsilon}}$ denotes the strain rate tensor, $\bar{c}_p$ the specific heat capacity, H the internal heat generation and k the thermal conductivity. The equation is of advection-diffusion type, and the terms on the right-hand side describe, in this order, shear-heating, internal heating (through radioactive decay) and adiabatic heating.

Under the assumption that the pressure deviations from the hydrostatic reference are very small ($\bar{p} \gg p'$) and using again $\nabla\bar{p} = \bar{\rho}\mathbf{g}$, we end up with the energy equation in its final form

$$\frac{\partial T}{\partial t}+\mathbf{u}\cdot\nabla T-\frac{1}{\bar{\rho}\bar{c}_p}\nabla\cdot(k\nabla T) = \frac{1}{\bar{\rho}\bar{c}_p}\boldsymbol{\tau}:\dot{\boldsymbol{\epsilon}}+\frac{\bar{\alpha}}{\bar{c}_p}(\mathbf{u}\cdot\mathbf{g})T+H \tag{8}$$

In most parts of the mantle, advective transport is the dominant process, exceeding the effectiveness of diffusive transport by several orders of magnitude (Peclet number Pe $\simeq 10^3$).

The model is complemented by choosing initial conditions for temperature, from which one can derive an initial velocity field and initial pressure by solving the Stokes equations for the resulting buoyancy term. Additionally, one needs to impose boundary conditions for $\boldsymbol{u}$ and T on the surface ($r = r_{\text{max}}$) and the CMB ($r = r_{\text{min}}$). Temperature boundary conditions are usually of Dirichlet type, while for velocity different choices are possible. Often in a mantle convection model one requires

$$\begin{aligned}
\boldsymbol{u}\cdot\hat{\boldsymbol{r}} &= 0 && \text{on } \delta\Omega \\
\boldsymbol{t}\cdot\boldsymbol{\tau}\cdot\hat{\boldsymbol{r}} &= 0 && \text{for } r = r_{\text{min}} \\
\boldsymbol{u}\cdot\boldsymbol{t} &= \boldsymbol{u}_{\text{tangential}} && \text{for } r = r_{\text{max}}
\end{aligned} \tag{9}$$

The first of these constitutes a no-outflow condition. In the second condition $\boldsymbol{t}$ represents any vector in the tangential plane of the boundary point. Thus, it requires that at the CMB the tangential shear-stress vanishes. Combined with no-outflow, this constitutes a free-slip condition. Physically this is motivated by the fact that below the CMB lies the outer core, composed of molten iron on which the mantle rocks can freely move. The third and final condition corresponds to fixing the tangential

velocity component on the surface, for which plate-tectonic reconstructions can be used. This is commonly done in simulations for retrodictions to assimilate data, see e.g. Colli et al. (2018). Note that this requires care (such as appropriately scaling velocities) in order to not introduce artificial forcing into the model. Thus, people might prefer to use alternative boundary conditions like e.g. the previous one also on the surface.

2.1 Space Discretization Method

We employ the Finite Element Method (FEM) for determining an approximate solution of the Stokes problem as well as for discretising the spatial derivatives in the energy equation, apart from the advection term, see section 2.2. Let $\mathcal{T}$ be a triangulation of the problem domain Ω, i.e. a splitting into non-overlapping triangles in 2D or tetrahedrons in 3D. Then the discrete version of the weak form of the Stokes problem becomes

$$\int_\Omega \boldsymbol{\tau}(\mathbf{u}_h) : \nabla \mathbf{v}_h d\Omega - \int_\Omega p_h \nabla \cdot \mathbf{v}_h d\Omega = \int_\Omega \mathbf{f} \cdot \mathbf{v}_h d\Omega \tag{10}$$

$$- \int_\Omega \nabla \cdot (\bar{\rho} \mathbf{u}_h)\, q_h d\Omega = 0, \tag{11}$$

Here $\mathbf{f}$ is the right-hand side (RHS) from eq. (5) and $(\boldsymbol{u}_h, p_h)$ are the sought for approximate velocity and pressure solutions, while $(\boldsymbol{v}_h, q_h)$ are test functions. These are chosen from finite-dimensional spaces, commonly e.g. $\boldsymbol{u}_h, \boldsymbol{v}_h \in V_h \subset (H_1(\Omega))^3$ and $p_h, q_h \in Q_h \subset L_2^0(\Omega)$, where the latter denotes the space of all square-integrable functions with zero mean. Note that eq. (10) does not contain any surface integrals for the considered boundary conditions eq. (9). More details on the derivation of the variational formulation including boundary conditions can e.g. be found in Burstedde et al. (2013).

In the case of a free-slip boundary condition, cf. eq. (9), we treat it as a natural boundary condition and only remove the radial velocity component from the test and trial functions to satisfy the no-outflow component. In the context of a curved domain this requires special consideration, hence we follow the approach by Engelman et al. (1982).

A large variety of different mixed FE function pairs have been suggested in the literature for the Stokes problem, see e.g. Terrel et al. (2012). For our experiments in this study, we have opted for the standard $P_2 - P_1$ Taylor-Hood pair. This is not only inf-sup stable, but also one of the two choices advocated for in the comparison done by Thieulot and Bangerth (2022).

2.2 Time Discretization Method

Finding a stable and accurate time discretization scheme for the energy equation is tricky in our setting as it is strongly advection dominated. A common technique that avoids unphysical oscillations in the numerical solution is the Streamline Upwind Petrov Galerkin (SUPG) approach, which adds a consistent artificial diffusion in the direction of the velocity streamline. This resolves the problem, but requires heuristic tuning of additional parameters for weighting the stabilization term. Another alternative would be entropy-viscosity stabilisation, see Euen et al. (2023) for details and a comparison to SUPG.

Here, we instead use an Eulerian-Lagrangian approach, in the sense that we treat the advective part of the equation with the Modified Method of Characteristics (MMOC). The latter is based on the fact that in the case of pure advection temperature

remains constant along the characteristic curves of the PDE. For full details and benchmarks, see Kohl et al. (2022). Here we only give a brief overview.

In the MMOC we initialise virtual particles at the locations of the DoFs of the FE temperature field at time $t+\delta t$. We then use an explicit Runge-Kutta scheme of 4$^{\text{th}}$ order to trace each particle back in time to its departure point $x_{\text{dept}}(t)$ at the previous time-step t. At these positions we evaluate the old temperature field to obtain an advected intermediate temperature $\hat{T}_h(x_{\text{dept}}(t), t+\delta t)$. The intermediate advected temperature is used in further time discretization of the FE weak form of the energy equation. We discretise the temperature T_h in space with a quadratic P_2 FE function, hence by choosing an appropriate test function s_h, we can write the weak form with an implicit Euler time discretization as,

$$\int_\Omega \frac{T_h(t+\delta t) - \hat{T}_h(t)}{\delta t} s_h d\Omega + \int_\Omega k \nabla T_h(t+\delta t) \cdot \nabla s_h d\Omega - \int_\Omega \alpha\,(\mathbf{u}_h \cdot \mathbf{g})\, T_h(t+\delta t) s_h d\Omega = \int_\Omega f_T(t+\delta t) s_h d\Omega. \tag{12}$$

Here the term f_T conflates the internal heating and shear-heating parts of eq. (8). Note that the latter depends on velocity, which we take to be $\mathbf{u}_h(t)$ for practical reasons. This also applies to the adiabatic heating term. The linear system resulting from eq. (12) can easily be solved with the help of a Krylov subspace solver. In our experiments we use either Conjugate Gradient or GMRES for this.

Once the temperature at time $(t+\delta t)$ is obtained, we use it to set up the buoyancy term and solve the Stokes system, thereby obtaining the corresponding velocity $\mathbf{u}(t+\delta t)$.

An aspect that we have not mentioned so far is that solving the individual ODEs for the trajectories of the virtual particles requires to provide the velocity field $\mathbf{u}_h$ at the micro time-steps of the Runge-Kutta method in $[t, t+\delta t]$. In practice this is not known. We use a predictor-corrector approach in our implementation. In a first step, we simply use the temperature from the previous time-step (constant extrapolation). Once we have solved for the new temperature and velocity at $(t+\delta t)$, we redo the step by linearly interpolating between $\boldsymbol{u}_h(t)$ and $\boldsymbol{u}_h(t+\delta t)$, while also updating the velocity dependent heating terms, and computing new values first for T_h and then for $\boldsymbol{u}_h$ at $(t+\delta t)$. In the case of a more pronounced coupling, e.g. in the case of a velocity dependent viscosity, we can repeat this process multiple times.

When using an implicit time-discretisation scheme one is not necessarily constraint by the Courant-Friedrichs-Lewy (CFL) condition when choosing time-step sizes. Nevertheless, in our experiments we still use the CFL condition as a guide for selecting the latter as it couples spatial and temporal resolution. Given a threshold $C_{\text{CFL}} > 0$, the maximal magnitude of velocity field in the domain u_{max}, and the maximal FE element length h, we chose as length for the next time-step

$$dt_{\text{CFL}} = C_{\text{CFL}} \frac{h}{u_{\text{max}}} \quad . \tag{13}$$

2.3 Simplifying Formulations for HPC

The most computationally intensive part of the simulation process is solving the Stokes system. Hence, any simplifying approximation or modification is desirable as long as it is consistent with the physics of the simulation. In this section, we give a brief overview on some of such methods, which we will then test in section 4.2.1.

2.3.1 Frozen velocity

Using a compressible flow formulation, the density field in the continuity equation (2) varies spatially. This then spoils the symmetry of the resulting saddle point linear system arising from the FE discretization, as the bilinear forms for the gradient and divergence part are no longer identical. Following Tan et al. (2011); Heister et al. (2017), by applying the product rule of the divergence operator, one can, in a first step, re-write eq. (6) as

$$\nabla \cdot \mathbf{u} = -\frac{\nabla \rho}{\rho} \cdot \mathbf{u}, \tag{14}$$

and modify the weak form of the TALA formulation, specifically eq. (11) to become

$$\int_\Omega (\nabla \cdot \mathbf{u}_h)\, q_h d\Omega = -\int_\Omega \frac{\nabla \bar{\rho}}{\bar{\rho}} \cdot \mathbf{u}_h q_h d\Omega. \tag{15}$$

When solving the Stokes system at time $t_{n+1} = t + \delta t$, one can, as a second step, "freeze" the velocity on the right-hand side to be $u_h(t_n)$, which retains symmetry between the two bilinear forms.

2.3.2 Frozen divergence

The full stress tensor $\boldsymbol{\tau}$ contains a divergence part $\frac{2}{3}\eta(\nabla \cdot \mathbf{u})\boldsymbol{I}$ which is non-zero in the compressible cases and, when the divergence of $\boldsymbol{\tau}$ is taken, results in a grad-div term. Additionally, this also increases the complexity for the core kernel functions which perform matrix-vector operations. One can treat this term similar to the frozen velocity case, i.e. by moving it to the RHS of momentum equation eq. (1) and evaluating it with the velocity field from the previous time-step:

$$-\nabla \cdot \left[\eta\left(\nabla \mathbf{u}(t_{n+1}) + \nabla \mathbf{u}^\top(t_{n+1})\right)\right] + \nabla p(t_{n+1}) = \mathbf{f} - \frac{2}{3}\nabla\left[\eta \nabla \cdot \mathbf{u}(t_n)\right] \tag{16}$$

Hence, we obtain an extra forcing term on the right-hand side, but retain the stress tensor in the form of the incompressible setting. In this work we only use this approach for the benchmark with the compressible Stokes system on a unit square in section 4.2.1. More experiments and mathematical analysis will be required to fully validate the effectiveness of the approach and its influence on accuracy.

3 Finite Element Framework

Given the desired global resolution of $\simeq 1$ km, holding the data required to solve the resulting linear systems in memory becomes challenging. If we consider a FE coefficient vector with about a trillion (10^{12}) entries (which approximately corresponds to the number of elements in such a mesh), then the size of the FE coefficient vector is about 8 TB when working in double precision. Comparing this to the total amount of about 1 PB of main memory of the Hawk supercomputer (66th in Top500 as of Nov '24) that was used for experiments in this paper, we can see that we can fit only 125 such vectors in memory. Since we typically cannot access the entire machine at once, and since this does not include the memory required for the system matrices yet (which will require storing 10-100 or more non-zeros per row) such simulations can only be performed with significantly more main memory or with numerical methods with reduced memory demands.

Since memory access also is relatively slow, and thus often a performance-limiting factor, a better solution is to not form the matrices explicitly and employ matrix-free methods (Brown, 2010; Kronbichler and Kormann, 2012; May et al., 2015; Kohl and Rüde, 2022). This, of course, restricts the set of solvers that can be implemented to those that work with matrix-free compute kernels, but enables the solution of linear systems at the extreme scale. In this work, we build on the framework HYTEG for parallel, matrix-free FEM simulations and the associated code generation framework HOG for architecture-optimized matrix-free compute kernels. Both together have been shown to allow efficient solution of systems at tera scale ($\simeq 10^{12}$ DoFs) (Kohl and Rüde, 2022; Böhm et al., 2025).

3.1 HYTEG and HOG

HYBRID TETRAHEDRAL GRIDS (HYTEG) builds upon an unstructured coarse mesh composed of triangles in 2D and tetrahedrons in 3D that captures the problem's basic geometry. The elements of the coarse mesh are split into their geometric primitives (vertices, edges, faces and cells), which we denote as macro-primitives. On these we perform structured refinement to reach a desired mesh resolution, thereby generating micro-elements. Data associated with micro-entities is stored in the corresponding macro-entities, which act as containers and also form the smallest units for parallelisation in HYTEG, which is based on the distributed memory paradigm and implemented via the Message Passing Interface (MPI). The refinement process results in a hierarchy of meshes and is, thus, particularly well-suited for the development of efficient geometric multigrid methods. The block-structuredness of the resulting mesh allows using data-structures without indirections, which is beneficial in multiple aspects, e.g. memory access, and avoids the need to store coordinate and topology information on micro-entities, as these can be computed on-the-fly. Additionally, it supports implementation of various optimisations, such as e.g. cache-friendly loop strategies within the matrix-free compute kernels, which are not possible on fully unstructured FE meshes. In order to ease performance portability and testing of new optimisations, the core compute kernels such as operator-application (i.e. the analogue of a matrix-vector multiplication in a matrix-free setting) or computation of inverse-diagonals are not implemented manually, but auto-generated via the HYTEG OPERATOR GENERATOR (HOG). This also significantly simplifies addition of new PDE problems and their bilinear forms. For more details on HYTEG and HOG we refer the reader to Kohl et al. (2019, 2024); Böhm et al. (2025).

3.2 Mesh refinement

HYTEG can import arbitrary tetrahedral meshes in GMSH format and also supports inline mesh generators for standard domains. Among these is also one for our primary target domain, the thick spherical shell. Construction of a coarse mesh for the latter is based on the icosahedral meshing approach, see Baumgardner and Frederickson (1985); Davies et al. (2013). Here an icosahedron is mapped onto the unit sphere, resulting in 20 spherical triangles. To these one applies $(n_{\text{tan}} - 1)$ steps of recursive midpoint refinement with geodesic arcs. Once this tangential surface mesh is fine enough, it is radially extended to cover $\Omega_{r_{\text{min}}, r_{\text{max}}}$ with n_{rad} radial layers. This results in spherical prismatoids, each of which can be split into three tetrahadrons. To this tetrahedral base mesh the standard refinement process of HYTEG is applied, which recursively splits each triangle into four sub-triangles and each tetrahedron into eight sub-tetrahedra.

However, this standard refinement will not improve the domain approximation, due to the curvature of the latter. To resolve this issue, HYTEG employs a blending map, which maps the micro-elements of the refined triangulation onto the actual problem domain. We denote the former as computational domain Ω_{comp} and the latter as physical domain Ω_{phys}, see Bauer et al. (2017) for further details. In this setting, if we specify n_{rad} radial layers to the icosahedral mesh generator, then we get n_{rad} spherical surfaces in the resulting mesh which also corresponds to $(n_{\text{rad}} - 1)$ FE elements in the radial direction. From an FE perspective HYTEG in such a case employs two mappings. One from the reference element to the micro-element on the computational domain, which is then mapped onto the computational domain. This has to be accounted for in the element integrals and the gradients of the shape functions. Denoting by J_A the Jacobian of the affine mapping onto the computational domain and by J_B that of the blending map we obtain e.g.

$$\nabla\psi_i = \mathcal{J}_{\text{B}}^{-\top}\mathcal{J}_{\text{A}}^{-\top}\nabla\phi_i \tag{17}$$

for an ansatz function ψ_i and its associated shape function ϕ_i on the reference element. On every micro-element the blending map needs to be at least a diffeomorphism to support the usual integral transformations. This can be achieved by making it a diffeomorphism on every macro-element. However, globally it suffices for the blending map to be only homomorphic, which eases its construction. In the case of the icosahedral mesh for the thick spherical shell such a blending map is known explicitly which also holds for its inverse. This is of technical importance for the MMOC approach (Kohl et al., 2022) and constitutes an advantage over using a quadratic isoparametric mapping.

3.3 Linear System Solvers

The linear system of equations that arises from the weak form of the Stokes system is the saddle point problem

$$\begin{bmatrix} A & B^\top \\ B & 0 \end{bmatrix} \begin{pmatrix} U \\ P \end{pmatrix} = \begin{pmatrix} f_U \\ f_P \end{pmatrix}. \tag{18}$$

In a matrix-free setting, iterative solvers form the only possible choice for computing solutions. HYTEG offers different approaches for solvers and preconditioners, and due to the hierarchy of levels generated through the structured refinement it naturally supports geometric multigrid approaches. Multigrid methods, either algebraic or geometric, feature in most approaches to solve the Stokes system. This is related to the A-block, which results from the discretisation of an elliptic operator. Gmeiner et al. (2016) e.g. compared three different iterative solvers, all involving multigrid in some form, for the isoviscous case and found a monolithic multigrid method with inexact Uzawa smoothing to offer the best performance. The smoother can be written as consecutive velocity and pressure system iterations, see e.g. Gaspar et al. (2014) for details. In Kohl and Rüde (2022) it was shown that using this combination in HYTEG makes it possible to achieve extreme scalability and solve for a trillion DoFs in less than a minute.

While the monolithic multigrid approach proves to be extremely scalable, we found it to have robustness issues in cases where viscosity is sharply varying. As this is often the case in geodynamic scenarios we opted for this paper to instead employ a Krylov subspace solver with multigrid preconditioning. To be more precise, we use a Flexible GMRES (FGMRES) method in

combination with the following inexact adjoint Uzawa preconditioner

$$p_{k+1} = p_k - \hat{S}^{-1}\Big[g - Bu_k\Big]$$
$$u_{k+1} = u_k + \hat{A}^{-1}\left(f - Au_k - B^\top p_{k+1}\right)$$

Here $\hat{S}$ approximates the system's Schur complement. A common choice for $\hat{S}$ is to use the pressure mass matrix scaled by the inverse of viscosity, as this is spectrally equivalent to the actual Schur complement, as long as viscosity is not too heterogeneous, see Rudi et al. (2017) and the references therein. The approximation $\hat{S}$ can be constructed as

$$\hat{S}_{ij} = \int_\Omega \mu^{-1}\phi_i\phi_j d\Omega \ . \tag{19}$$

Note that an alternative for truly strong viscosity variations is to use BFBT approaches for approximating the inverse Schur complement, see e.g. Rudi et al. (2017); Burkhart et al. (2025) and the references therein.

Application of $\hat{A}^{-1}$ corresponds to solving a linear system with A by geometric multigrid with Chebychev smoothing. Strongly variable coefficients, as in our case viscosity, pose a challenge for geometric multigrid, as without proper handling the convergence may degrade significantly. In a matrix-free setting, the construction of suitable coarse grid operators, e.g. by using a Galerkin Coarse Grid Approximation (GCA), see Trottenberg et al. (2001), is generally difficult or at least expensive (Knapek, 1998).

For this study we have opted for the following approach. We employ a Direct Coarse Grid Approximation (DCA), sometimes also referred to as rediscretisation, in combination with a homogenization approach for the viscosity. For the latter we first convert viscosity into an element-wise constant, i.e. P_0, function on the finest level by computing an arithmetic average. This P_0 viscosity is then restricted to coarser mesh levels by recursively computing an average of the viscosity over its child elements (Clevenger and Heister, 2021). We use an arithmetic average here in most models. For the case of nonlinear rheology in section 4.2.2 we found that the harmonic average provided a better multigrid convergence behaviour. Besides its use in the multigrid part, note that this averaging also has another positive effect. If viscosity exhibits sharp variations in the domain, pressure becomes inherently discontinuous. Using a Taylor-Hood approach the approximate pressure solution is forced to be continuous. This results in spikes in the solution, which get damped when one employs this averaging approach (Heister et al., 2017).

3.4 Rigid Body Modes

In some benchmarks and models free-slip boundary conditions are imposed on both the inner and outer boundary of either the thick spherical shell or the annulus. In both cases this results in a non-trivial kernel for the A block of the linear system eq. (18) that is composed of rigid body modes. This can lead to convergence problems with the iterative solver or the quality of the approximate solution. To avoid such issues one can employ a nullspace removal algorithm that filters out the rotational mode components from the residuals and solution vectors during every step of the iterative solver. This, however, will increase the number of dot products and require additional MPI communication, which in turn can affect the run-time performance of the solver.

Thus, we follow a different approach and penalize rigid body modes in the velocity field by extending the weak form of the momentum equation eq. (10) by the term

$$\sum_{d=\{x,y,z\}} \int_\Omega c_{\text{rot}} \left(\mathbf{e}^d_{\text{rigid}} \cdot \mathbf{u}_h\right) \left(\mathbf{e}^d_{\text{rigid}} \cdot \mathbf{v}_h\right) d\Omega. \tag{20}$$

Here, $\mathbf{e}^d_{\text{rigid}}(\mathbf{x})$ denotes the unit vector describing an elementary rotation around the d-axis, e.g. in the case of the z-axis we have $\mathbf{e}^z_{\text{rigid}} = [-y/r, x/r, 0]^\top$ with $r = \sqrt{x^2+y^2}$. The penalty value c_{rot} is chosen as a fixed global value and tuned manually. The extra term can easily be incorporated into the code and optimized by using the HOG code generator. The benefit here is that no extra MPI communication is needed, since the handling of the rigid body modes happens as part of the compute kernel for operator application.

3.5 Stress and Heat Flow Computation

After having performed a mantle convection simulation important quantities which one often wants to derive from the primary FE variables in a post-processing step are dynamic topography, the geoid or heat flow. Doing so requires accurate computation of the gradient of a FE field, e.g. velocity or temperature. For Finite Elements that are only H_1-conforming computing gradients inside each local element will result in a field that is usually discontinuous at the cell interfaces, especially also at the nodes. In this work, we consider two approaches for deriving a continuous gradient field. The first one is to compute gradients in an element-local fashion, as mentioned above, and then map the result into a selected space W_h^3 of globally continuous functions using a standard $\mathcal{L}_2$ projection. As an example, let $T_h \in V_h$ be a scalar FE function. Then $[\nabla T]_h$, the unknown approximation of its gradient field in W_h^3, needs to satisfy

$$\int_\Omega [\nabla T]_h \cdot \mathbf{v}_h d\Omega = \int_\Omega \nabla T_h \cdot \mathbf{v}_h d\Omega \tag{21}$$

for all $\mathbf{v}_h \in W_h^3$. It can, thus, easily be found by solving a linear system with the corresponding mass matrix. We will use $W^h = P_2$ for our tests.

The other method we consider is the Consistent Boundary Flux (CBF) method (Mizukami, 1986; Gresho et al., 1987) which is well known in the geophysics community through the work of Zhong et al. (1993), and its implementation and testing in CitComS and ASPECT (Zhong et al., 2008; Liu and King, 2019; Williams et al., 2024). In the CBF method we consider the weak form of the Stokes momentum equation eq. (10) assuming that the test function $\mathbf{v}_h$ is non-zero on the boundary $\partial\Omega$ of the domain Ω. Rearranging the integrals results in the following equation

$$\int_\Omega \mathbf{f} \cdot \mathbf{v}_h d\Omega - \int_\Omega \boldsymbol{\tau}_h : \nabla \mathbf{v}_h d\Omega + \int_\Omega p_h \left(\nabla \cdot \mathbf{v}_h\right) d\Omega = -\int_{\partial\Omega} \left(\boldsymbol{\tau}_h \mathbf{n}\right) \cdot \mathbf{v}_h d\Gamma + \int_{\partial\Omega} \left(p_h \mathbf{n}\right) \cdot \mathbf{v}_h d\Gamma = -\int_{\partial\Omega} \boldsymbol{\sigma}_h \mathbf{n} \cdot \mathbf{v}_h d\Gamma \tag{22}$$

Here $(\boldsymbol{\sigma}_h \mathbf{n})$ denotes the vector component of the full stress tensor with respect to the plane normal to $\mathbf{n}$. Considering the latter as an unknown FE function on $\partial\Omega$ we can compute it by solving a system with the boundary mass matrix. Note that in the case of a free-slip boundary only the radial component of this stress vector is non-zero. We will compare computation of radial

stress with the $\mathcal{L}_2$ projection and the CBF method in section 4.1.1 for the smooth forcing function case. This will demonstrate that, while the $\mathcal{L}_2$ projection allows to derive the gradient field in the whole domain Ω, the CBF method is superior in terms of accuracy in providing the gradient field on the boundary $\partial\Omega$.

The same strategy can be applied to the energy equation to compute the heat flux through the boundary $\partial\Omega$. Considering the weak form of the energy equation for a test function s_h not vanishing on the boundary, we obtain

$$\int_{\partial\Omega} (\nabla T_h \cdot \mathbf{n})\, s_h d\Gamma = \int_{\Omega} \frac{\partial T}{\partial t} s_h d\Omega + \int_{\Omega} (\mathbf{u}_h \cdot \nabla T_h)\, s_h d\Omega + \int_{\Omega} \nabla T_h \cdot \nabla s_h d\Omega \ . \tag{23}$$

Fixing again the terms on the right-hand side and taking $(\nabla T_h \cdot \mathbf{n})$ to be an unknown FE function $(\nabla T \cdot \mathbf{n})_h$ living on $\partial\Omega$ it can also be obtained by solving a system with a boundary mass matrix. This is the approach we employed for computing the Nusselt numbers reported in table 1.

4 Benchmarking and Validation

In this section, we benchmark HyTeG against various geophysical problems. Initially, we consider mathematical convergence studies against analytical solutions of the Stokes problem on the thick spherical shell. We then move to time-dependent community benchmarks including setups with compressible flow, temperature-dependent and nonlinear rheologies. This is a continuation of the work done in Ilangovan et al. (2024) where several simple benchmark experiments had been conducted with HyTeG.

4.1 Stationary Benchmarks

A common approach to verify numerical algorithms and their implementation is based on the method of manufactured solutions (Roache, 2001), which can be employed to check the order of convergence of the FE approximation to a known solution. However, this is typically restricted to comparatively simple setups and not actual application scenarios which usually involve complications such as e.g. discontinuities (or sharp variations) of material parameters. For incompressible Stokes flow on the annulus and spherical shell (semi-)analytical solutions were derived in e.g. Blinova et al. (2016); Horbach et al. (2019); Kramer et al. (2021). In a similar vein, solutions based on the propagator matrix method were used in Zhong et al. (2000); Choblet et al. (2007); Zhong et al. (2008); Tan et al. (2011) to benchmark mantle convection frameworks. In section 4.1.1 we first study HyTeG subject to the setup from Kramer et al. (2021), where the authors derive solutions for cases where the forcing function is in the form of a spherical harmonic and the same imposed with a Dirac delta function. To keep the setup closer to geodynamic models similar to Colli et al. (2018); Brown et al. (2022) (where Dirichlet boundary conditions are considered on the outer surface), the boundary conditions chosen here are of mixed type (noslip on the outer boundary, free-slip on the inner boundary). As a baseline to evaluate the δ-function case we consider the mixed case boundary condition setup with a smooth forcing function. In addition, we also show the order of convergence of the computed radial stress on the free-slip boundary with respect to the analytical solution for the smooth forcing, mixed boundary condition case. Here, we compare the accuracy of the radial stress computed with an $\mathcal{L}_2$ projection of the gradient to the CBF method. We conclude this subsection

by showing topography kernels computed at the surface and CMB for free-slip conditions. These kernels describe the response when placing a spherical density anomaly given by a specific spherical harmonic at various depths in the mantle. Next, in section 4.1.2, we demonstrate convergence for the pure free-slip case when using the penalty approach to treat rigid body modes proposed in section 3.4. For this we consider again the smooth forcing function.

4.1.1 Smooth and δ-function forcing

Consider the incompressible, isoviscous Stokes system

$$-\nabla \cdot \left(\nabla \mathbf{u} + (\nabla \mathbf{u})^{\top} \right) + \nabla p = f\mathbf{g} \ , \tag{24}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{25}$$

posed on a thick spherical shell with $r_{\text{min}} = 1.22$ and $r_{\text{max}} = 2.22$. The radii are chosen such that we keep the aspect ratio of the domain close to that of Earth's mantle $\frac{r_{\text{min}}}{r_{\text{max}}} \simeq 0.55$ while having a unit thickness. We impose the no-slip conditions on the outer and free-slip on the inner boundary. The forcing functions f_δ and f_{smooth} are given by

$$f_{\text{smooth}} = \frac{r^k}{R_{\text{max}}^k} \mathcal{Y}_{\ell m}(\theta, \phi) \ , \quad f_\delta = \delta\left(r - r_d\right) \mathcal{Y}_{\ell m}(\theta, \phi) \ , \tag{26}$$

where $\mathcal{Y}_{\ell m}$ is the spherical harmonic function of order ℓ degree m, and $\delta(\cdot)$ denotes the Dirac delta function. The analytical solution for the δ-function case will be a radial function piecewise defined on the two branches $r < r_d$ and $r > r_d$, whilst satisfying an appropriate interface condition. Based on (Kramer et al., 2021) we derived the analytical solution for the mixed boundary case in (Ilangovan et al., 2024) and found the expected convergence orders for a $P_2 - P_1$ pairing. Here we repeat this test using this time the FGMRES solver preconditioned with geometric multigrid described in section 3.3. The coarse base mesh $\mathcal{T}_0$ uses $(n_{\text{rad}}, n_{\text{tan}}) = (5, 3)$ resulting in 486 nodes and 1,920 tetrahedrons. Its radial layers are placed at the outer and inner surface and the middle one at $r_d = 1.72$, where we center the δ-function.

Exploiting the definition of the δ-function the volume integral on the RHS of the discrete weak form reduces to a surface integral over of spherical layer at the corresponding depth r_d

$$\int_{\Omega} \delta(r - r_d) \mathcal{Y}_{\ell m} \phi_i \mathbf{g} \, d\Omega = \int_{\Gamma_{r_d}} \mathcal{Y}_{\ell m} \phi_i \mathbf{g} \, d\Gamma \ . \tag{27}$$

We study convergence going from $\mathcal{T}_2$ (545,400 DoFs for velocity and pressure) up to $\mathcal{T}_4$ (33,300,936 DoFs). The coarsest level for the multigrid solver is always set to $\mathcal{T}_2$. The $\mathcal{L}_2$ error (e) on the domain Ω between the FE solution $\mathbf{u}_h^{\text{FE}}$ and the analytical solution $\mathbf{u}_h^{\text{analytical}}$ evaluated on the nodes of the FE mesh is computed as,

$$e = \sqrt{\left(\int_{\Omega} \| \mathbf{u}_h^{\text{FE}} - \mathbf{u}_h^{\text{analytical}} \|_2^2 d\Omega \right)} . \tag{28}$$

The behaviour of the error, shown in fig. 1 (left), demonstrates that the FE solution converges to the actual analytical solution with $\mathcal{O}(h^3)$ for velocity and $\mathcal{O}(h^2)$ for pressure in the case of smooth forcing (denoted by SF). This is the theoretically

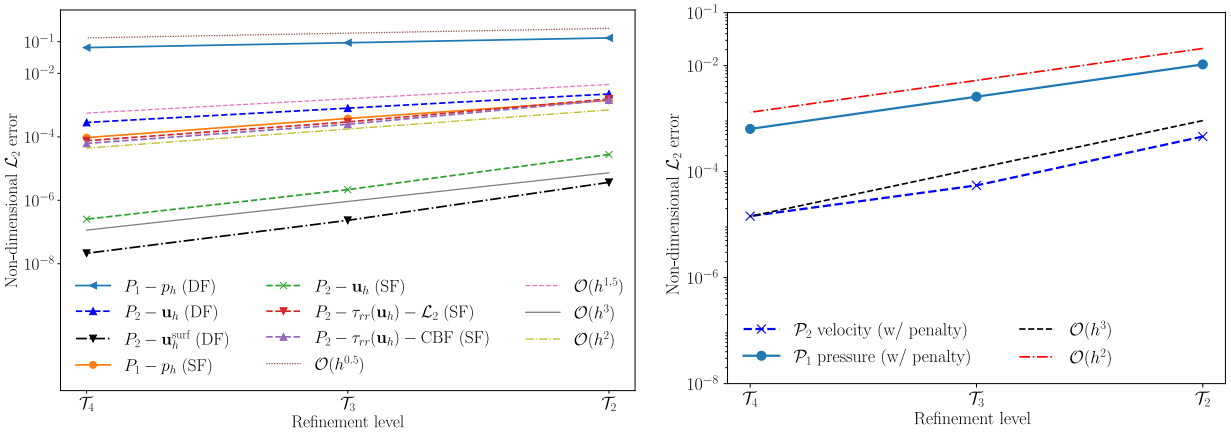

Figure 1. (Left) Order of convergence of the $P_2 - P_1$ FE solution to analytical solution on the Spherical Shell with noslip condition on outer and free-slip on the inner surface (mixed case) under the δ-forcing (denoted by DF) function case and smooth-forcing (denoted by SF) function case for ($\ell = 2, m = 2$, for SF case $k = 2$); (Right) Order of convergence of the $P_2 - P_1$ FE solution to analytical solution on the Spherical Shell with free-slip imposed on both surfaces under the smooth forcing case and tuned rotational mode penalty for ($\ell = 2, m = 2, k = 2$)

expected order of convergence for a $P_2 - P_1$ Taylor Hood element. In addition, fig. 1 (left) compares the errors in the radial stress computed on the inner free-slip boundary for the SF case, when using the $\mathcal{L}_2$-projection approach and the CBF method. Both achieve approximately a convergence order of $\mathcal{O}(h^2)$. However, the overall error with the CBF method is consistently lower showing its superiority.

In the δ-function forcing case (denoted by DF), convergence of both the velocity and pressure solution deteriorates to $\mathcal{O}(h^{1.5})$ and $\mathcal{O}(h^{0.5})$ respectively, as can be seen from fig. 1 (left). This is a known phenomenon, when considering the error over the complete domain, and can, supposedly, be attributed to the fact that, while the analytical pressure solution is discontinuous, we force it to be continuous by representing it as a P_1 function. However, when only considering the error on the free-slip surface one observes an $\mathcal{O}(h^3)$ convergence of velocity for r_d in the interior. See (Kramer et al., 2021) for a more detailed discussion. As a final result, we present in fig. 2 dynamic topography kernels for the surface and CMB. These are the responses resulting from spherical density perturbations placed at varying depths in the mantle. These perturbations are modeled again by f_δ from eq. (26). Radial stresses are computed from the resulting velocity field using the CBF method and are used to extract the spherical harmonic coefficient corresponding to the spherical harmonic (ℓ, m). Extracting this value for every depth, gives the surface and CMB kernel functions for that spherical harmonic (ℓ, m). These are plotted in fig. 2, where the kernel functions derived from the FE solution are overlaid with the ones derived from the analytical radial stresses on the same surface grid. The test uses the parameter set $(\ell, m) = \{(2,2), (5,4), (8,4), (12,8), (15,8)\}$, although the kernels are only dependent on the degree ℓ (Ribe, 2007). From the plots, we see excellent agreement in the values from spherical harmonic degrees $\ell = 2$ to $\ell = 15$.

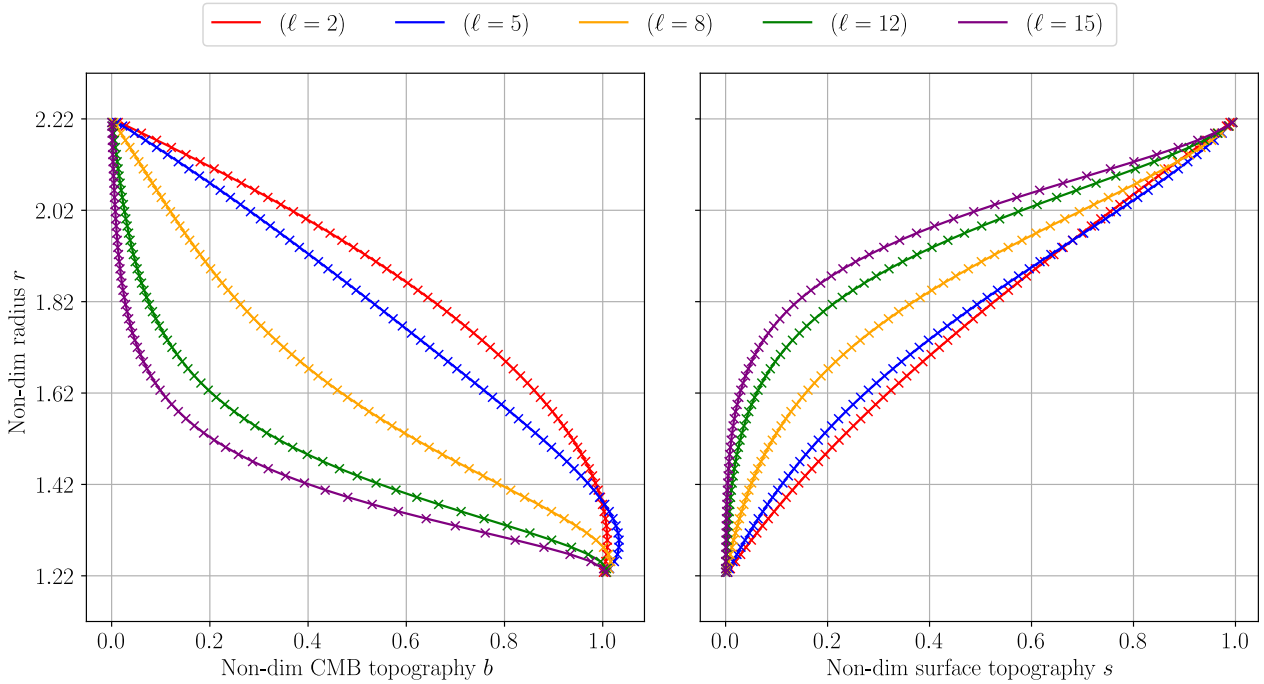

Figure 2. Kernels for the response of the surface and CMB topography to a δ-function forcing of the specified spherical harmonic degree (ℓ) placed at various depths in the mantle; solid lines denote the analytical solution and the cross marks (x) denote the FE computed solution.

4.1.2 Rotational Mode Penalty

In this part we focus on the convergence of the penalty approach described in section 3.4 for the case of free-slip boundary conditions imposed on both the inner and outer surface. For this we prescribe a smooth forcing function eq. (26) for the Stokes system. We use the same domain configuration as in the previous section and also keep the multigrid preconditioned FGMRES solver. Again we study convergence going from $\mathcal{T}_2$ up to $\mathcal{T}_4$, while using $\mathcal{T}_2$ as coarsest level for multigrid. The number of FGMRES iterations is kept the same at every refinement level $\mathcal{T}_l$.

In fig. 1 (right) we give errors with respect to an analytic solution without rigid body modes in the velocity field. The results demonstrate that choosing appropriate values for the penalty parameter maintains the expected $O(h^3)$ and $O(h^2)$ convergence for velocity and pressure, while taking care of the rigid body modes.

A theoretical study of the choice of penalty value is out of the scope of this contribution. Thus, the values for c_{rot} were tuned manually and set roughly to behave as $O(h)$. Also to get this convergence behaviour, a different mesh configuration was chosen for $\mathcal{T}_0$. The precise values for both can be found in the Python notebooks, see Ilangovan (2025).

4.2 Geophysical Benchmarks

In this section, we will consider numerical convection experiments that are commonly used in the geophysics community to benchmark mantle convection frameworks. We start in section 4.2.1 with a compressible convection simulation using the TALA approximation on a unit square. We test the approximation techniques proposed in section 2.3.1 and section 2.3.2 and verify the resulting Nusselt number against other codes. Then in section 4.2.2, we perform the benchmark from Tosi et al. (2015) which employs a viscoplastic rheology. Finally, in section 4.2.3, we consider the benchmark from Ratcliff et al. (1996); Zhong et al. (2008) on the spherical shell and compare the radially averaged temperatures and Rayleigh number vs Nusselt number trends reported in other works.

4.2.1 Compressible case

We consider the compressible Stokes equations eq. (5) and eq. (6) coupled with the energy equation eq. (8) on the unit square. The goal is to perform an experiment using our methodologies with the setup being identical to that in King et al. (2010) and compare the Nusselt numbers that we obtain on the top boundary. In the benchmark free-slip boundary conditions are imposed on all sides. In addition to diffusion and adiabatic heating/cooling in the energy equation eq. (8), we also include the shear heating term. The experiment is started from an initial sinusoidal perturbation,

$$T(\mathbf{x}, t=0) = (1-y) + A_0 \cos \pi x \sin \pi y, \tag{29}$$

with $A_0 = 0.1$, which then induces a single convection cell in the square. An isoviscous, compressible case with $\bar{\rho} = \exp[\mathrm{Di}(1-y)]$ is considered, where Di is the dissipation number taken as $\mathrm{Di} = 0.5$. The unit square mesh contains 4 subdivisions in the x and y direction generating a total of 32 triangles at the coarsest level $\mathcal{T}_0$. The triangles are refined to reach the operating mesh at level n denoted as $\mathcal{T}_n$. To solve the linear system of equations the FGMRES solver with multigrid as preconditioner is used. The finest level for the multigrid is $\mathcal{T}_n$ and the coarsest level $\mathcal{T}_0$. Time-steps are calculated based on the CFL condition. At each time-step, 10 iterations are performed with the FGMRES solver, with two coupling Picard iterations between the Stokes and energy equation. The system eq. (5), eq. (6) and eq. (8) is non-dimensionalised following King et al. (2010), which introduces the Rayleigh and Dissipation number. The simulation is performed at Rayleigh number $\mathrm{Ra} = 10^5$ and run until the temperature reaches a steady state as seen in fig. 3. After this, the Nusselt number is computed on the top boundary and checked for convergence. The Nusselt number describes the ratio between heat transported by convection to heat transported purely by diffusion. Taking the pure diffusion solution as reference, T_{ref}, the Nusselt number can, thus, be computed for the final temperature solution T via

$$\mathrm{Nu} = \frac{\int_\Gamma \nabla T \cdot \hat{\mathbf{n}} \, d\Gamma}{\int_\Gamma \nabla T_{\mathrm{ref}} \cdot \hat{\mathbf{n}} \, d\Gamma} \ . \tag{30}$$

We compare three different approaches for the compressible setting here. Using the full asymmetric Stokes system, employing the frozen velocity approach (from section 2.3.1) and additionally also the frozen divergence approach (from section 2.3.2). The corresponding Nusselt numbers are reported in fig. 3 (right). From the figure we can see that all three approaches behave

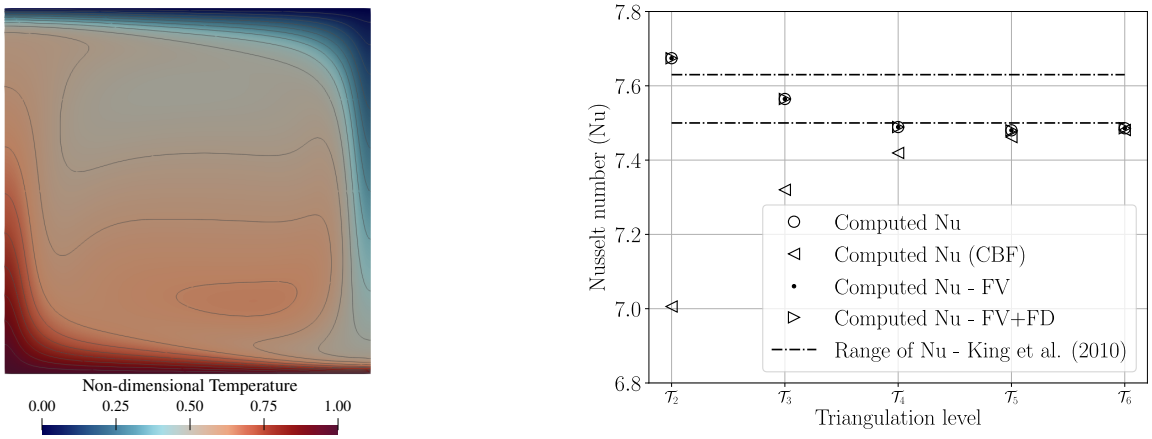

Figure 3. (Left) Final Temperature contours of the compressible benchmark on an unit square of the TALA case from King et al. (2010) presented in section 4.2.1 with Ra $= 10^5$ and Di $= 0.5$; (Right) Comparison of Nusselt number values between the Frozen velocity approach (FV, see section 2.3.1) and including Frozen divergence approach (FD, see section 2.3.2) under the setup from section 4.2.1 identical to the Truncated Anelastic Liquid Approximation (TALA) case from King et al. (2010) at Ra$= 10^5$ and Di$= 0.5$.

very similarly and converge to the same Nusselt number. Note that this final value is on the lower end of the Nusselt numbers reported in King et al. (2010) (Nu $= 7.50$). Specifically this is the value for the CZ code, which also uses a semi-Lagrangian approach to treat advection. Hence, as our approach is also similar to this, the Nusselt numbers computed here are expected.
455 Note that the temparature gradient in eq. (30) appears in the integrand. It can, thus, be evaluated and used locally, without re-sorting to the methods discussed in section 3.5, as we have done here. For comparison, we also computed Nu for the asymmetric approach using the CBF method. As can be seen it approaches the same final value albeit from below.

4.2.2 Nonlinear Rheology

Although the rheology of the mantle is not completely known, it is clear that it exhibits a nonlinearity (Boioli et al., 2018).
460 A simple example would be plate boundaries, where large stresses (which depend on the velocity $\mathbf{u}$) make the plates exhibit plastic behaviour. This decreases viscosities giving rise to weak zones at these locations. In this section, we conduct a plume rising experiment with a sinusoidal temperature perturbation (eq. (29)) on the unit square with a nonlinear rheology. The setup is identical to case 4 from Tosi et al. (2015), where the viscosity is defined as

$$\mu = \left(\frac{1}{\mu_{\text{linear}}} + \frac{1}{\mu_{\text{nonlinear}}} \right)^{-1} \tag{31}$$

465 $$\mu_{\text{linear}} = \exp(-\gamma_T T + \gamma_z z) \tag{32}$$

$$\mu_{\text{nonlinear}} = \mu^* + \frac{\sigma_y}{\sqrt{\dot{\boldsymbol{\epsilon}} : \dot{\boldsymbol{\epsilon}}}} \ . \tag{33}$$

Here γ_T and γ_z are positive constants, and the constant μ^* models the effective viscosity in the nonlinear regime, which is active when the stresses exceed the yield stress σ_y. We treat the viscosities as a P_0 FE function, i.e. constant per cell,

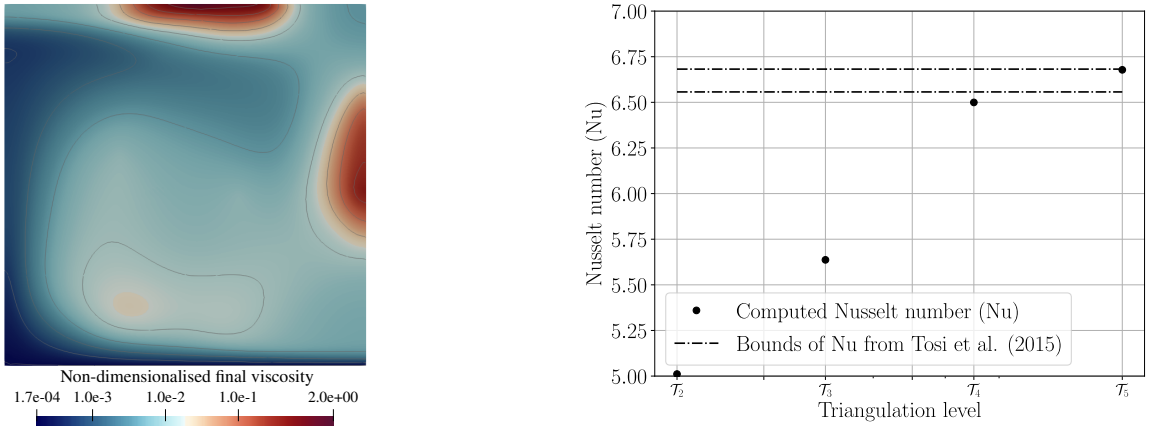

Figure 4. Results for the setup described in section 4.2.2 after reaching steady state: (Left) Final viscosity contours showing formed weak zones; (Right) Convergence of the Nusselt numbers computed on the top boundary with mesh refinement.

and average them on coarser grids. The FGMRES solver with a multigrid preconditioner applied to the viscous A-Block is used with 10 FGMRES iterations per Stokes solve. We run the simulation until a single convection cell has stabilised in the domain. The dependence of viscosity on velocity makes the Stokes system nonlinear, which we handle by performing five Picard type iterations at every time-step. This number could be reduced, though, once the system starts to stabilise. The final viscosity contour in log scale can be seen from fig. 4 (left). If viscosity was only temperature-dependent (μ_{linear}), then it would consistently be large along the top and right boundary of the domain, where temperature is coldest. However, as we consider a pseudoplastic rheology, the viscosity is considerably smaller where stress is large. This is the case in the upper left and right boundary, where rising material hits the top or starts sinking down, but also at the lower part of the right boundary. Figure 4 (right) shows the resulting Nusselt number at the top boundary and how it behaves when we refine the mesh. For a high enough resolution the value falls inside the region reported for other codes in Tosi et al. (2015) and Davies et al. (2022).

4.2.3 Spherical Shell

This section involves experiments on a thick spherical shell where an incompressible Stokes system with variable viscosity coupled to an energy equation is considered (Ratcliff et al., 1996; Zhong et al., 2008; Davies et al., 2022; Euen et al., 2023). The boundary conditions prescribed are of free-slip type on both inner and outer surface. The initial condition for the temperature field is a radially linear function with an added symmetrical perturbation according to a spherical harmonic function $\mathcal{Y}_{lm}$. The benchmark consists in checking, if the spherical harmonic components of the resulting temperature field largely maintain the prescribed order and degree. The initial condition for the temperature field is given by

$$T(t=0) = T_{\text{ref}} + \epsilon T_{\text{perturb}}, \tag{34}$$

Case	Rayleigh number (Ra)	r_μ	$\mathrm{Nu}_{\mathrm{top}}$ - HYTEG	$\mathrm{Nu}_{\mathrm{top}}$ - ASPECT/CitcomS
A3	7×10^3	20	3.139	3.14 - 3.19
C1	1×10^5	1	7.758	7.37 - 7.81
C3	1×10^5	30	6.756	6.50 - 6.79

Table 1. Test cases considered for comparison of Nusselt numbers (Nu) and radially averaged temperatures and velocity RMS values with published data from other codes (Euen et al., 2023)

with $\epsilon > 0$. T_{ref} is a reference profile which satisfies the steady diffusion equation. We consider two types pertubations $T_{\mathrm{perturb}} = \mathcal{Y}_{\ell m}$: A choice of degree $\ell = 3$ and order $m = 2$ results in a tetrahedral symmetry of plumes and downwellings, while a cubic symmetry is obtained for $T_{\mathrm{perturb}} = \mathcal{Y}_{40} + \frac{5}{7}\mathcal{Y}_{44}$. The prescribed viscosity variation is dependent on the evolving temperatures and is given by the following expression,

$$\mu = 10^{r_\mu(1/2-T)} \quad , \tag{35}$$

where $r_\mu > 0$ and T is the non-dimensional temperature whose values lie between 0 and 1. By increasing the value of r_μ, the viscosity contrast between the plumes and downwellings can be increased. By taking a pure diffusion solution as the reference, $T_{\mathrm{ref}} = \frac{r_{\mathrm{min}} r_{\mathrm{max}}}{r} - r_{\mathrm{min}}$, the Nusselt number can be computed with eq. (30). Experiments are performed on a thick spherical shell with $r_{\mathrm{min}} = 1.22$ and $r_{\mathrm{max}} = 2.22$, which has an aspect ratio of $\frac{r_{\mathrm{min}}}{r_{\mathrm{max}}} \simeq 0.55$, which is similar to that of Earth's mantle, while keeping non-dimensionalised thickness of 1, (Ratcliff et al., 1996; Davies et al., 2022).

The Stokes system is solved with the geometric multigrid preconditioned FGMRES solver described in section 3.3. Simulations are started with 25 FGMRES steps for the first 25 time-steps. Afterwards the number of FGMRES iterations is reduced to between 5 and 15 depending on the magnitude of viscosity contrasts (15 when $r_\mu > 10^2$). On each multigrid level three pre- and postsmoothing steps with a 3rd order Chebyshev smoother are applied. MINRES is used as the coarse grid solver. The coarsest level $(\mathcal{T}_2)$ of the multigrid solver corresponds to a resolution of $(n_{\mathrm{tan}}, n_{\mathrm{rad}}) = (9, 5)$ and the finest level $(\mathcal{T}_4)$ to $(33, 17)$. For comparison to other published results, we consider the scenarios A3, C1, and C3 from Zhong et al. (2008), also presented in Euen et al. (2023), for which the respective values used can be seen from table 1. We then also simulate the A7 case from Zhong et al. (2008) to assess the regime switching point. In addition, we perform multiple simulations and plot the variation of the Nusselt number with respect to the Rayleigh number as was done in Ratcliff et al. (1996).

All simulations are run until the system reaches a statistical steady state i.e. the radially averaged temperatures and the Nusselt numbers converge to a certain value. A CFL condition threshold (see eq. (13)) of $C_{\mathrm{CFL}} = 1.2$ is used, which we gradually increased to $C_{\mathrm{CFL}} = 1.5$ when we approach the steady state. The resulting temperature isosurfaces at $T = 0.5$ coloured by the magnitude of velocity are given in fig. 5 for the A3 and C3 case. The radially averaged temperatures and velocity RMS values from the final time-step are plotted for our simulations in combination with the data presented in Euen et al. (2023) for the A3 case in fig. 6 and the C1 and C3 case in fig. 7. The corresponding Nusselt numbers are presented in table 1. The radius values from Euen et al. (2023) are scaled $(\times \frac{1}{0.45})$ for comparison with the results computed here as the mantle thickness is different

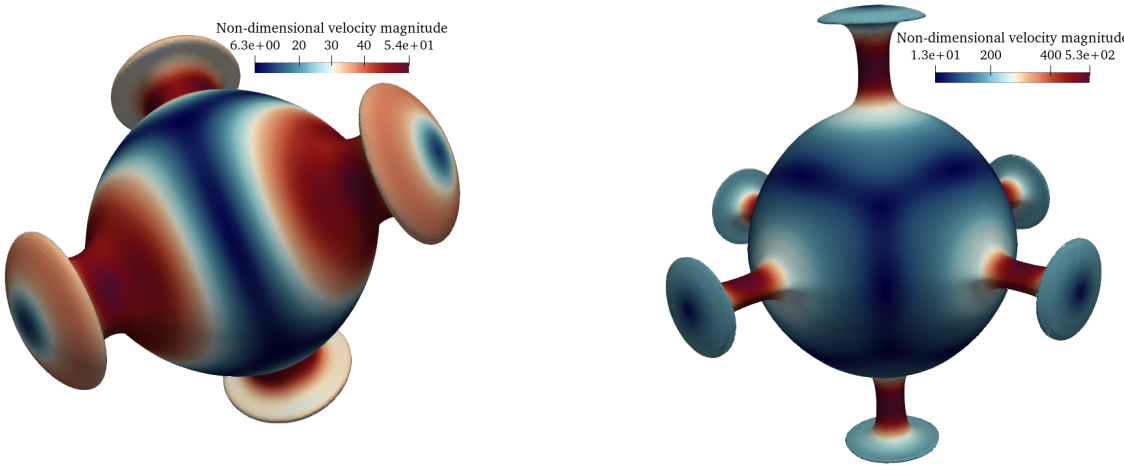

Figure 5. (Left) Temperature isosurfaces at $T = 0.5$ for the A3 case coloured with non-dimensional velocity magnitudes; (Right) Temperature isosurfaces at $T = 0.5$ for the C3 case coloured with non-dimensional velocity magnitudes;

in both setups, although they are comparable since the aspect ratio of the domain is the same in both setups. As can be seen from the figures, the results obtained with our framework agree well with those presented in Euen et al. (2023) for ASPECT and CitComS.

Results for scenario A7 are given in the left of fig. 8. Here one keeps the Rayleigh number at $\mathrm{Ra} = 7 \times 10^3$, but increases r_μ to 10^5. This takes us past the transition point from a steady state regime to a stagnant lid regime. The viscosity contrast is now so high that viscosity increases drastically towards the outer surface preventing the development of downwellings there. Also, plumes are restricted to the lower region near the inner surface.

Next we consider multiple cases with different Rayleigh numbers and r_μ values to analyse the variation of the achieved Nusselt number with respect to Rayleigh numbers. For a given r_μ value and a given degree ℓ of spherical harmonic perturbation $T_{\text{perturb}} = \mathcal{Y}_{\ell m}$, there exists a specific value $\mathrm{Ra}_{\text{crit}}$, that indicates the onset of convection, see the stability analysis in Zebib (1993). Specifically we consider all combinations of Rayleigh numbers $\mathrm{Ra} \in \{7 \times 10^3, 2 \times 10^4, 4 \times 10^4, 6 \times 10^4\}$ and $r_\mu \in \{3, 10, 20, 40, 100, 500, 700, 1000\}$ for the case of the perturbation with cubic symmetry (giving us degree $\ell = 4$). The choice of values is motivated by Ratcliff et al. (1996). For each combination we perform one simulation run with a CFL threshold of 1.2 and compute the Nusselt number at the end of the simulation. These are then plotted in fig. 8 (right) versus $\frac{\mathrm{Ra}}{\mathrm{Ra}_{\text{crit}}}$, where the critical Rayleigh number was taken from Zebib (1993). As can be seen, a power law exponent of 0.31 fits the Nusselt number data well, which is in accordance with earlier studies (Jarvis, 1993).

Although these experiments are a challenging test to the framework, we see that the benchmarks only compare the steady state result reached at the end of the simulation, while there is nothing that can be said about how the Stokes and Energy solver gets there. As the result is only verified at the end of the time evolution, the temperature and velocity patterns through which the solvers reach the solution can be different.

5 Scalability of the Framework

Development of the HyTeG framework is performed with a focus on HPC applications at the extreme scale. Thus, scalability studies have already been presented, e.g. in Kohl and Rüde (2022) and Böhm et al. (2025). These were, however, either limited to simple (isoviscous) Stokes problems or other operators such as curl-curl.

In order to highlight the capabilities of our state-of-the-art framework when running high-resolution global-scale mantle convection models, we consider in this section its scaling behaviour for a realistic model with varying viscosity and mixed boundary conditions posed on a thick spherical shell. Since the major chunk of the computational effort goes into solving the Stokes system, eq. (5) and eq. (6), we focus on the latter for our experiments.

As before we again use block preconditioned FGMRES with geometric multigrid for the A-block as solver, see section 3.3. Experiments were performed on the Hawk supercomputer at HLRS (66th on the Top500 list as of Nov 2024). We demonstrate scalability with up to approximately 10^{11} DoFs on $15,360$ cores. The timings reported are for 5 FGMRES steps with a restart at every 3 steps and 1 pre- and post-smoothing iteration each per level in the multigrid solver. For a single weak scaling test, we fix the number of coarse mesh elements to 1 per MPI process and refine it several times as outlined in section 2.1. Keeping the maximal refinement level fixed, while increasing the number of processes, then maintains the number of DoFs per process. We perform multiple weak scaling tests at different refinement levels $\mathcal{T}_k$.

We first present in fig. 9 (left) a comparison between the runtime of the application compiled once with and once without vectorization. As can be seen from the plot, we achieve nearly a 4x improvement with the HOG-generated kernels supporting AVX extensions, (Böhm et al., 2025) and the following results also use this optimization.

In the runtimes reported in fig. 9 (right), we consider multiple weak scaling cases. Each line indicates a single weak scaling test and the $\mathcal{T}_k$ next to the line denotes the highest level (k) of refinement the corresponding test was performed at. From the plots, we see that the framework scales extremely well for up to 61,440 MPI processes. The slightly reduced runtime for the $\mathcal{T}_6$ case in the right figure stems from a technical change with respect to the implementation of the treatment of free-slip boundary

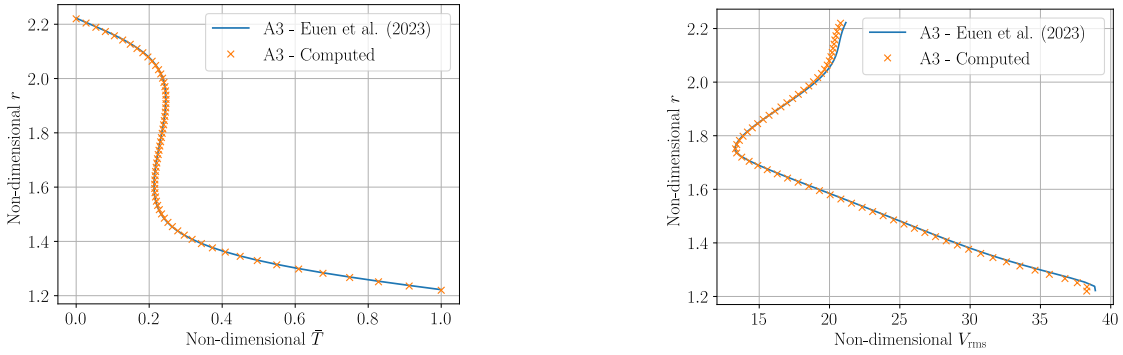

Figure 6. Radially averaged temperature values $\bar{T}$ (left) and radially averaged velocity RMS values V_{rms} (right) for the A3 case from Zhong et al. (2008); we compare the results published in Euen et al. (2023) (solid line) with results computed in our setup

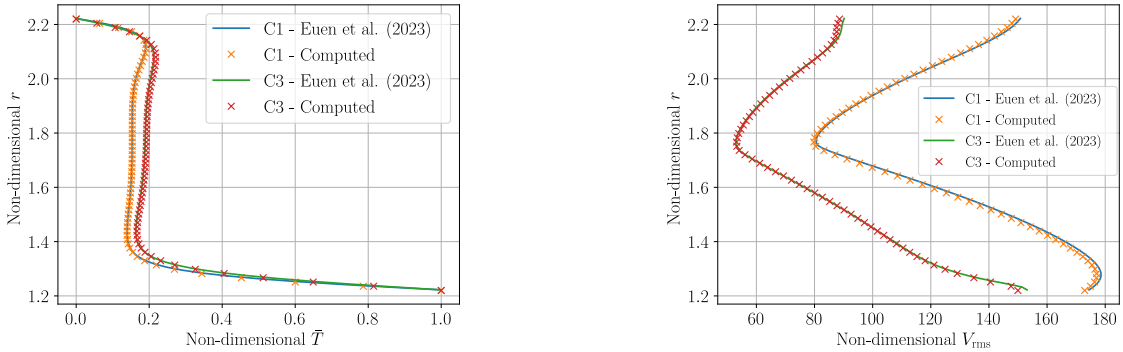

Figure 7. Radially averaged temperature values $\bar{T}$ (left) and radially averaged velocity RMS values V_{rms} (right) for the C1 and C3 cases from Zhong et al. (2008); we compare the results published in Euen et al. (2023) (solid lines) with results computed in our setup

555 conditions, see section 2.1. This change applies to all our weak scaling runs. These also used a representation of viscosity as P_1 function. This, however, had no significant influence on the runtime behaviour.

To further test the capabilities of TerraNeo we increased the number of DoFs to about 10^{11} and simulated a variable viscosity Stokes system on 15,360 cores. Due to memory limitations we needed to adjust the solver setup and restarted FMGRES already after 2 steps. The results are marked by *-R2* in fig. 9 and demonstrate that the code can also handle a system of this size.

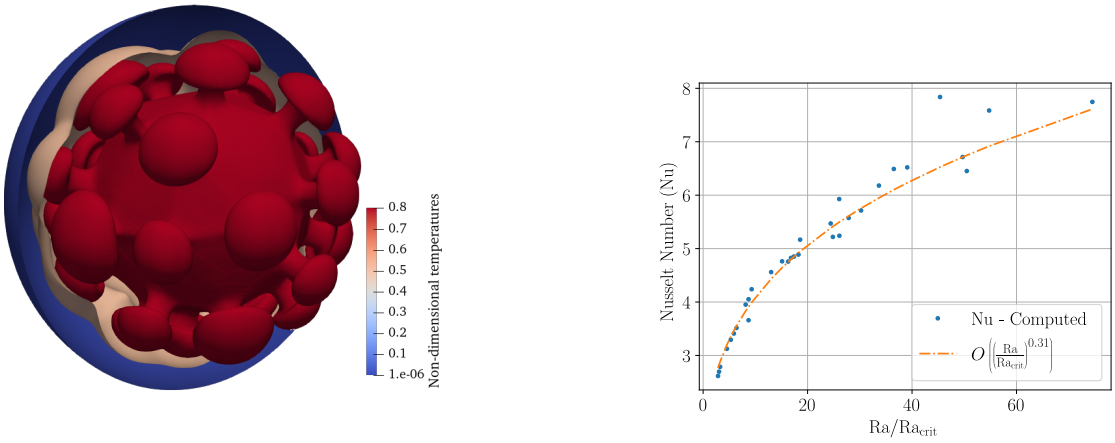

Figure 8. (Left) Temperature isosurfaces at $T = 0.8$ (red), $T = 0.5$ (light red) and $T = 10^{-6}$ (blue) for the A7 case from Zhong et al. (2008); (Right) Variation of Nu vs $\frac{\text{Ra}}{\text{Ra}_{\text{crit}}}$ for the cubic symmetry case; each data point denotes a simulation run where Ra_{crit} is computed depending on the r_μ value for a given spherical harmonic perturbation, with the best fit exponent plotted side by side $\left(\frac{\text{Ra}}{\text{Ra}_{\text{crit}}}\right)^{0.31}$

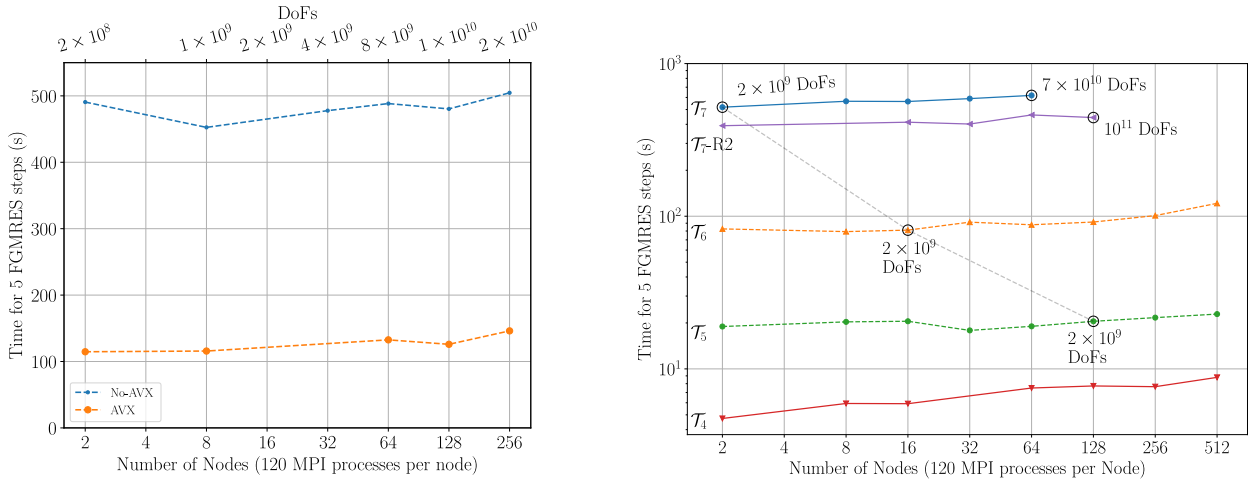

Figure 9. (Left) Weak scaling at $\mathcal{T}_6$ for an identical setup with and without vectorization using the code generated from the HOG generator, (Right) Weak scaling behaviour on the Hawk supercomputer at HLRS (66th on Top500 as of Nov 2024) up to 512 nodes; $\mathcal{T}_4$, $\mathcal{T}_5$, $\mathcal{T}_6$ or $\mathcal{T}_7$ gives the highest refinement level when simulating the variable viscosity Stokes system; the postfix -R2 indicates a weak scaling test with peak number of unknowns being 10^{11} (but a restart of FGMRES after already two steps due to memory limitations)

With respect to the strong scaling behaviour of the framework, we highlight in fig. 9 (right) three runs using the same problem size, i.e. 2×10^9 DoFs, but different number of MPI processes and, thus, also different refinement levels. The system when solved at $\mathcal{T}_7$ on 2 nodes takes about 500 seconds, while it can be brought down to $\simeq 20$ seconds on 128 nodes at refinement level $\mathcal{T}_5$. These results clearly show that the framework is scalable and suitable for handling globally resolved mantle convection models.

An important point to note is that, these tests only assess the Stokes solver for the scalability of its implementation in HyTeG and not on its solvability to achieve a specific residual reduction.

6 Conclusions

In this paper, we presented a comprehensive evaluation of the HyTeG Finite Element framework when applied to problems from computational geodynamics. We assessed the key components required for large-scale mantle convection simulations, including the multigrid-based linear solver for the Stokes system and a semi-Lagrangian advection scheme, through a series of targeted tests and geophysical benchmarks.

We verified the expected convergence rates of the Finite Element solutions against analytical results for Stokes flow with Dirac-delta-type forcing, demonstrating the robustness of the implemented multigrid solver. We also proposed and validated a novel methodology to eliminate rigid body modes in the thick spherical shell geometry with tangential stress-free, no-normal-flow boundary conditions.

We conducted benchmark simulations representative of realistic geodynamic scenarios, including compressible flows and nonlinear, temperature-dependent rheologies. These experiments spanned from unit square domains to thick spherical shells and confirmed the physical fidelity and numerical stability of the proposed methods.

Scalability studies on the Hawk supercomputer (ranked 66th on the Top500 list as of November 2024) showed that our implementation is capable of handling up to 10^{11} degrees of freedom in the Stokes system. While the FGMRES solver exhibits robustness under large viscosity contrasts, the convergence is hindered by inadequate coarse-grid approximations in the geometric multigrid preconditioner. Addressing this bottleneck is an ongoing area of research, with continued exploration of techniques for improved coarse-level correction.

Overall, our results demonstrate that HyTeG is a highly scalable and capable framework for finely resolved mantle convection simulations. Future work will focus on further optimizing multigrid performance and exploring advanced solver strategies to fully exploit the potential of matrix-free methods at extreme scales.

Code and data availability. The software used in the experiments is available on Zenodo: 10.5281/zenodo.17341565 (Ilangovan, 2025), while the resulting data will be made accessible through the Leibniz Supercomputing Center (LRZ): 10.25927/zvnf9-cvc38 (Ilangovan et al., 2025)

Author contributions. PI: implementation, execution and analysis of the benchmarks, writing and editing of draft; NK: supervision, reviewing and editing of draft, acquisition of compute time; MM: supervision, reviewing and editing of draft, acquisition of compute time and funding.

Competing interests. The authors declare that they do not have competing interests.

Acknowledgements. This work was supported by the German Federal Ministry of Research, Technology and Space (BMFTR) as part of its initiative "SCALEXA - New Methods and Technologies for Exascale Computing" (project 16ME0649 - CoMPS). Computing resources were provided by the Institute of Geophysics of LMU Munich, Oeser et al. (2006), funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 495931446 and 518204048 and the Höchstleistungrechenzentrum Stuttgart (HLRS). The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SuperMUC-NG at Leibniz Supercomputing Centre (www.lrz.de). Furthermore the authors wish to thank all their previous and current collaborators from the HyTeG team.

References

Bangerth, W., Dannberg, J., Fraters, M., Gassmoeller, R., Glerum, A., Heister, T., Myhill, R., and Naliboff, J.: geodynamics/aspect: ASPECT 2.5.0, https://doi.org/10.5281/ZENODO.8200213, 2023.

Bauer, S., Mohr, M., Rüde, U., Weismüller, J., Wittmann, M., and Wohlmuth, B.: A two-scale approach for efficient on-the-fly operator assembly in massively parallel high performance multigrid codes, Applied Numerical Mathematics, 122, 14–38, https://doi.org/10.1016/j.apnum.2017.07.006, 2017.

Bauer, S., Bunge, H.-P., Drzisga, D., Ghelichkhan, S., Huber, M., Kohl, N., Mohr, M., Rüde, U., Thönnes, D., and Wohlmuth, B.: Terra-Neo—Mantle Convection Beyond a Trillion Degrees of Freedom, p. 569–610, Springer International Publishing, ISBN 9783030479565, https://doi.org/10.1007/978-3-030-47956-5_19, 2020.

Baumgardner, J.: Three-dimensional treatment of convective flow in the earth's mantle, Journal of Statistical Physics, 39, 501–511, https://api.semanticscholar.org/CorpusID:121461665, 1985.

Baumgardner, J. R. and Frederickson, P. O.: Icosahedral discretization of the two-sphere, SIAM Journal on Numerical Analysis, 22, 1107–1115, https://doi.org/10.1137/0722066, 1985.

Blankenbach, B., Busse, F., Christensen, U., Cserepes, L., Gunkel, D., Hansen, U., Harder, H., Jarvis, G., Koch, M., Marquart, G., Moore, D., Olson, P., Schmeling, H., and Schnaubelt, T.: A benchmark comparison for mantle convection codes, Geophysical Journal International, 98, 23–38, https://doi.org/10.1111/j.1365-246X.1989.tb05511.x, 1989.

Blinova, I., Makeev, I., and Popov, I.: Benchmark Solutions for Stokes Flows in Cylindrical and Spherical Geometry, Bull. Transilvania U. Braşov, 9, 11–16, 2016.

Böhm, F., Bauer, D., Kohl, N., Alappat, C. L., Thönnes, D., Mohr, M., Köstler, H., and Rüde, U.: Code Generation and Performance Engineering for Matrix-Free Finite Element Methods on Hybrid Tetrahedral Grids, SIAM Journal on Scientific Computing, 47, B131–B159, https://doi.org/10.1137/24M1653756, 2025.

Boioli, F., Carrez, P., Cordier, P., Goryaeva, A., Gouriet, K., Hirel, P., Kraych, A., Mahendran, S., Mussi, A., NZOGANG, B. C., Reali, R., Ritterbex, S., and Sun, X.-Y.: Multiscale Modeling of the Mantle Rheology: The RheoMan Project, Université de Lille, 2018.

Boussinesq, J.: Theorie analytique de la chaleur vol 2 (paris: Gauthier-villars), Buoyancy Effects in Fluids, 1903.

Brown, H., Colli, L., and Bunge, H.-P.: Asthenospheric flow through the Izanagi-Pacific slab window and its influence on dynamic topography and intraplate volcanism in East Asia, Frontiers in Earth Science, 10, https://doi.org/10.3389/feart.2022.889907, 2022.

Brown, J.: Efficient Nonlinear Solvers for Nodal High-Order Finite Elements in 3D, Journal of Scientific Computing, 45, 48–63, https://doi.org/10.1007/s10915-010-9396-8, 2010.

Bunge, H.-P. and Baumgardner, J. R.: Mantle convection modeling on parallel virtual machines, Computers in Physics, 9, 207–215, https://doi.org/10.1063/1.168525, 1995.

Burkhart, A., Kohl, N., Wohlmuth, B., and Zawallich, J.: A robust matrix-free approach for large-scale non-isothermal high-contrast viscosity Stokes flow on blended domains with applications to geophysics, https://doi.org/10.48550/ARXIV.2506.04157, 2025.

Burstedde, C., Stadler, G., Alisic, L., Wilcox, L. C., Tan, E., Gurnis, M., and Ghattas, O.: Large-scale adaptive mantle convection simulation, Geophysical Journal International, 192, 889–906, https://doi.org/10.1093/gji/ggs070, 2013.

Choblet, G., Čadek, O., Couturier, F., and Dumoulin, C.: ŒDIPUS: a new tool to study the dynamics of planetary interiors, Geophysical Journal International, 170, 9–30, https://doi.org/10.1111/j.1365-246x.2007.03419.x, 2007.

Clevenger, T. C. and Heister, T.: Comparison between algebraic and matrix-free geometric multigrid for a Stokes problem on adaptive meshes with variable viscosity, Numerical Linear Algebra with Applications, 28, https://doi.org/10.1002/nla.2375, 2021.

Colli, L., Ghelichkhan, S., Bunge, H.-P., and Oeser, J.: Retrodictions of Mid Paleogene mantle flow and dynamic topography in the Atlantic region from compressible high resolution adjoint mantle convection models: Sensitivity to deep mantle viscosity and tomographic input model, Gondwana Research, 53, 252–272, https://doi.org/10.1016/j.gr.2017.04.027, 2018.

Davies, D. R., Davies, J. H., Bollada, P. C., Hassan, O., Morgan, K., and Nithiarasu, P.: A hierarchical mesh refinement technique for global 3-D spherical mantle convection modelling, Geoscientific Model Development, 6, 1095–1107, https://doi.org/10.5194/gmd-6-1095-2013, 2013.

Davies, D. R., Kramer, S. C., Ghelichkhan, S., and Gibson, A.: Towards automatic finite-element methods for geodynamics via Firedrake, Geoscientific Model Development, 15, 5127–5166, https://doi.org/10.5194/gmd-15-5127-2022, 2022.

Davies, G. F.: Dynamic Earth: Plates, Plumes and Mantle Convection, Cambridge University Press, ISBN 9780511605802, https://doi.org/10.1017/cbo9780511605802, 1999.

Engelman, M. S., Sani, R. L., and Gresho, P. M.: The implementation of normal and/or tangential boundary conditions in finite element codes for incompressible fluid flow, International Journal for Numerical Methods in Fluids, 2, 225–238, https://doi.org/10.1002/fld.1650020302, 1982.

Euen, G. T., Liu, S., Gassmöller, R., Heister, T., and King, S. D.: A comparison of 3-D spherical shell thermal convection results at low to moderate Rayleigh number using ASPECT (version 2.2.0) and CitcomS (version 3.3.1), Geoscientific Model Development, 16, 3221–3239, https://doi.org/10.5194/gmd-16-3221-2023, 2023.

Gaspar, F. J., Notay, Y., Oosterlee, C. W., and Rodrigo, C.: A Simple and Efficient Segregated Smoother for the Discrete Stokes Equations, SIAM Journal on Scientific Computing, 36, A1187–A1206, https://doi.org/10.1137/130920630, 2014.

Gassmöller, R., Dannberg, J., Bangerth, W., Heister, T., and Myhill, R.: On formulations of compressible mantle convection, Geophysical Journal International, 221, 1264–1280, https://doi.org/10.1093/gji/ggaa078, 2020.

Gmeiner, B., Huber, M., John, L., Rüde, U., and Wohlmuth, B.: A quantitative performance study for Stokes solvers at the extreme scale, J. Comp. Sci., 17, 509–521, https://doi.org/10.1016/j.jocs.2016.06.006, 2016.

Gough, D. O.: The Anelastic Approximation for Thermal Convection, Journal of Atmospheric Sciences, 26, 448 – 456, https://doi.org/10.1175/1520-0469(1969)026<0448:TAAFTC>2.0.CO;2, 1969.

Gresho, P. M., Lee, R. L., Sani, R. L., Maslanik, M. K., and Eaton, B. E.: The consistent Galerkin FEM for computing derived boundary quantities in thermal and or fluids problems, International Journal for Numerical Methods in Fluids, 7, 371–394, https://doi.org/10.1002/fld.1650070406, 1987.

Ham, D. A., Kelly, P. H. J., Mitchell, L., Cotter, C. J., Kirby, R. C., Sagiyama, K., Bouziani, N., Vorderwuelbecke, S., Gregory, T. J., Betteridge, J., Shapero, D. R., Nixon-Hill, R. W., Ward, C. J., Farrell, P. E., Brubeck, P. D., Marsden, I., Gibson, T. H., Homolya, M., Sun, T., McRae, A. T. T., Luporini, F., Gregory, A., Lange, M., Funke, S. W., Rathgeber, F., Bercea, G.-T., and Markall, G. R.: Firedrake User Manual, Imperial College London and University of Oxford and Baylor University and University of Washington, 1st edn., https://doi.org/10.25561/104839, 2023.

Heister, T., Dannberg, J., Gassmöller, R., and Bangerth, W.: High accuracy mantle convection simulation through modern numerical methods – II: realistic models and problems, Geophysical Journal International, 210, 833–851, https://doi.org/10.1093/gji/ggx195, 2017.

Horbach, A., Mohr, M., and Bunge, H.-P.: A semi-analytic accuracy benchmark for Stokes flow in 3-D spherical mantle convection codes, GEM - International Journal on Geomathematics, 11, https://doi.org/10.1007/s13137-019-0137-3, 2019.

Ilangovan, P.: Software used for experiments in "Highly Scalable Geodynamic Simulations with HyTeG", https://doi.org/10.5281/zenodo.17341565, 2025.

Ilangovan, P., D'Ascoli, E., Kohl, N., and Mohr, M.: Numerical Studies on Coupled Stokes-Transport Systems for Mantle Convection, p. 288–302, Springer Nature Switzerland, ISBN 9783031637599, https://doi.org/10.1007/978-3-031-63759-9_33, 2024.

Ilangovan, P., Kohl, N., and Mohr, M.: Geodynamic Benchmark Simulations with HyTeG: Supplementary Data for Article "Highly Scalable Geodynamic Simulations with HyTeG", https://doi.org/10.25927/zvnf9-cvc38, 2025.

Jarvis, G. T.: Effects of curvature on two-dimensional models of mantle convection: Cylindrical polar coordinates, Journal of Geophysical Research: Solid Earth, 98, 4477–4485, https://doi.org/10.1029/92jb02117, 1993.

King, S. D., Lee, C., van Keken, P. E., Leng, W., Zhong, S., Tan, E., Tosi, N., and Kameyama, M. C.: A community benchmark for 2-D Cartesian compressible convection in the Earth's mantle, Geophysical Journal International, 180, 73–87, https://doi.org/10.1111/j.1365-246x.2009.04413.x, 2010.

Knapek, S.: Matrix-Dependent Multigrid Homogenization for Diffusion Problems, SIAM Journal on Scientific Computing, 20, 515–533, https://doi.org/10.1137/s1064827596304848, 1998.

Kohl, N. and Rüde, U.: Textbook Efficiency: Massively Parallel Matrix-Free Multigrid for the Stokes System, SIAM Journal on Scientific Computing, 44, C124–C155, https://doi.org/10.1137/20M1376005, 2022.

Kohl, N., Thönnes, D., Drzisga, D., Bartuschat, D., and Rüde, U.: The *HyTeG* finite-element software framework for scalable multigrid solvers, International Journal of Parallel, Emergent and Distributed Systems, 34, 477–496, https://doi.org/10.1080/17445760.2018.1506453, 2019.

Kohl, N., Mohr, M., Eibl, S., and Rüde, U.: A Massively Parallel Eulerian-Lagrangian Method for Advection-Dominated Transport in Viscous Fluids, SIAM Journal on Scientific Computing, 44, C260–C285, https://doi.org/10.1137/21M1402510, 2022.

Kohl, N., Bauer, D., Böhm, F., and Rüde, U.: Fundamental data structures for matrix-free finite elements on hybrid tetrahedral grids, International Journal of Parallel, Emergent and Distributed Systems, 39, 51–74, https://doi.org/10.1080/17445760.2023.2266875, 2024.

Kramer, S. C., Davies, D. R., and Wilson, C. R.: Analytical solutions for mantle flow in cylindrical and spherical shells, Geoscientific Model Development, 14, 1899–1919, https://doi.org/10.5194/gmd-14-1899-2021, 2021.

Kronbichler, M. and Kormann, K.: A generic interface for parallel cell-based finite element operator application, Computers & Fluids, 63, 135–147, https://doi.org/https://doi.org/10.1016/j.compfluid.2012.04.012, 2012.

Liu, S. and King, S. D.: A benchmark study of incompressible Stokes flow in a 3-D spherical shell using ASPECT, Geophysical Journal International, 217, 650–667, https://doi.org/10.1093/gji/ggz036, 2019.

May, D., Brown, J., and Le Pourhiet, L.: A scalable, matrix-free multigrid preconditioner for finite element discretizations of heterogeneous Stokes flow, Computer Methods in Applied Mechanics and Engineering, 290, 496–523, https://doi.org/10.1016/j.cma.2015.03.014, 2015.

McKenzie, D. and Jarvis, G.: The conversion of heat into mechanical work by mantle convection, Journal of Geophysical Research: Solid Earth, 85, 6093–6096, https://doi.org/10.1029/JB085iB11p06093, 1980.

Mizukami, A.: A mixed finite element method for boundary flux computation, Comp. Meth. Appl. Mech. Engg, 57, 239–243, https://doi.org/10.1016/0045-7825(86)90016-2, 1986.

Moresi, L.-N. and Solomatov, V. S.: Numerical investigation of 2D convection with extremely large viscosity variations, Physics of Fluids, 7, 2154–2162, https://doi.org/10.1063/1.868465, 1995.

Oberbeck, A.: Ueber die Wärmeleitung der Flüssigkeiten bei Berücksichtigung der Strömungen infolge von Temperaturdifferenzen, Annalen der Physik, 243, 271–292, https://doi.org/10.1002/andp.18792430606, 1879.

Oeser, J., Bunge, H.-P., and Mohr, M.: Cluster Design in the Earth Sciences: TETHYS, in: High Performance Computing and Communications – Second International Conference, HPCC 2006, Munich, Germany, edited by Gerndt, M. and Kranzlmüller, D., vol. 4208 of *Lecture Notes in Computer Science*, p. 31–40, Springer, https://doi.org/10.1007/11847366_4, 2006.

Ratcliff, J. T., Schubert, G., and Zebib, A.: Steady tetrahedral and cubic patterns of spherical shell convection with temperature-dependent viscosity, Journal of Geophysical Research: Solid Earth, 101, 25 473–25 484, https://doi.org/10.1029/96jb02097, 1996.

Ribe, N.: Analytical Approaches to Mantle Dynamics, p. 167–226, Elsevier, ISBN 9780444527486, https://doi.org/10.1016/b978-044452748-6.00117-6, 2007.

Roache, P. J.: Code Verification by the Method of Manufactured Solutions, Journal of Fluids Engineering, 124, 4–10, https://doi.org/10.1115/1.1436090, 2001.

Rudi, J., Stadler, G., and Ghattas, O.: Weighted BFBT Preconditioner for Stokes Flow Problems with Highly Heterogeneous Viscosity, SIAM Journal on Scientific Computing, 39, S272–S297, https://doi.org/10.1137/16m108450x, 2017.

Schubert, G., Turcotte, D. L., and Olson, P.: Mantle Convection in the Earth and Planets, Cambridge University Press, ISBN 9780511612879, https://doi.org/10.1017/cbo9780511612879, 2001.

Tan, E., Choi, E., Thoutireddy, P., Gurnis, M., and Aivazis, M.: GeoFramework: Coupling multiple models of mantle convection within a computational framework, Geochemistry, Geophysics, Geosystems, 7, https://doi.org/10.1029/2005gc001155, 2006.

Tan, E., Leng, W., Zhong, S., and Gurnis, M.: On the location of plumes and lateral movement of thermochemical structures with high bulk modulus in the 3-D compressible mantle: PLUMES AND CHEMICAL STRUCTURES, Geochemistry, Geophysics, Geosystems, 12, https://doi.org/10.1029/2011gc003665, 2011.

Terrel, A. R., Scott, L. R., Knepley, M. G., Kirby, R. C., and Wells, G. N.: Finite elements for incompressible fluids, pp. 385–397, Springer Berlin Heidelberg, https://doi.org/10.1007/978-3-642-23099-8_20, 2012.

Thieulot, C. and Bangerth, W.: On the choice of finite element for applications in geodynamics, Solid Earth, 13, 229–249, https://doi.org/10.5194/se-13-229-2022, 2022.

Tosi, N., Stein, C., Noack, L., Hüttig, C., Maierová, P., Samuel, H., Davies, D. R., Wilson, C. R., Kramer, S. C., Thieulot, C., Glerum, A., Fraters, M., Spakman, W., Rozel, A., and Tackley, P. J.: A community benchmark for viscoplastic thermal convection in a 2-D square box, Geochemistry, Geophysics, Geosystems, 16, 2175–2196, https://doi.org/10.1002/2015GC005807, 2015.

Trottenberg, U., Oosterlee, C., and Schüller, A.: Multigrid, Academic Press, iSBN 0-12-701070-X, 2001.

Williams, K., Stamps, D. S., Austermann, J., King, S., and Njinju, E.: Effects of using the consistent boundary flux method on dynamic topography estimates, Geophysical Journal International, 238, 1137–1149, https://doi.org/10.1093/gji/ggae203, 2024.

Zebib, A.: Linear and weakly nonlinear variable viscosity convection in spherical shells, Theoretical and Computational Fluid Dynamics, 4, 241–253, https://doi.org/10.1007/bf00417930, 1993.

Zhong, S., Gurnis, M., and Hulbert, G.: Accurate determination of surface normal stress in viscous flow from a consistent boundary flux method, Physics of the Earth and Planetary Interiors, 78, 1–8, https://doi.org/10.1016/0031-9201(93)90078-n, 1993.

Zhong, S., Zuber, M. T., Moresi, L., and Gurnis, M.: Role of temperature-dependent viscosity and surface plates in spherical shell models of mantle convection, Journal of Geophysical Research: Solid Earth, 105, 11 063–11 082, https://doi.org/10.1029/2000jb900003, 2000.

Zhong, S., McNamara, A., Tan, E., Moresi, L., and Gurnis, M.: A benchmark study on mantle convection in a 3-D spherical shell using CitcomS, Geochemistry, Geophysics, Geosystems, 9, https://doi.org/10.1029/2008gc002048, 2008.