A pointwise conservative method for thermochemical convection under the compressible anelastic liquid approximation

Nathan Sime[1], Cian R. Wilson[1], and Peter van Keken[1]

[1]Carnegie Institution for Science

November 26, 2022

Abstract

In prior work we found that precise approximation of the continuity constraint is crucial for accurate propagation of tracer data when advected through a background incompressible velocity field (Sime et al., 2021). Here we extend this investigation to compressible flows using the anelastic liquid approximation (ALA) and address four related issues: 1. exact conservation of tracer discretized fields through a background compressible velocity; 2. exact mass conservation; 3. addition and removal of tracers without affecting (exact) conservation to preserve a consistent number of tracers per cell; and 4. the diffusion of tracer data, for example, as induced by thermal or chemical effects. In this process we also present an abstract formulation of the interior penalty hybrid discontinuous Galerkin (HDG) finite element formulation for diffusion problems, and apply it to the advection-diffusion and compressible Stokes systems. Finally we present numerical experiments exhibiting the HDG compressible Stokes momentum formulation's superconvergent compressibility approximation and reproduce community numerical benchmarks from the literature for the ALA.

A pointwise conservative method for thermochemical convection under the compressible anelastic liquid approximation

Nathan Sime[1], **Cian R. Wilson**[1], **and Peter E. van Keken**[1]

[1]Earth and Planets Laboratory, Carnegie Institution for Science, Washington, D.C., U.S.A.

Key Points:

- Pointwise divergence free momentum fields offer an accurate approximation of tracer advection in compressible velocity fields.
- PDE-constrained l_2 projection exactly conserves mass of tracer data in compressible flows.
- Operator-splitting methods open up the use of tracer-based methods in advection-diffusion problems
- Tracers that discretize fields may be added and removed with no impact on conservation.

Corresponding author: Nathan Sime, `nsime@carnegiescience.edu`

Abstract

In prior work we found that precise approximation of the continuity constraint is crucial for accurate propagation of tracer data when advected through a background incompressible velocity field (Sime et al., 2021). Here we extend this investigation to compressible flows using the anelastic liquid approximation (ALA) and address four related issues: 1. exact conservation of tracer discretized fields through a background compressible velocity; 2. exact mass conservation; 3. addition and removal of tracers without affecting (exact) conservation to preserve a consistent number of tracers per cell; and 4. the diffusion of tracer data, for example, as induced by thermal or chemical effects. In this process we also present an abstract formulation of the interior penalty hybrid discontinuous Galerkin (HDG) finite element formulation for diffusion problems, and apply it to the advection-diffusion and compressible Stokes systems. Finally we present numerical experiments exhibiting the HDG compressible Stokes momentum formulation's superconvergent compressibility approximation and reproduce community numerical benchmarks from the literature for the ALA.

Keywords: Finite element analysis, tracer methods, compressible flow, mantle convection, geodynamics

1 Introduction

Plate tectonics causes continuous differentiation of the Earth's mantle. At mid-oceanic ridges decompression melting forms a basaltic crust on top of a depleted peridotite. Melting at subduction zones helps generate the continental crust and provides further chemical modifications to the subducting crust and mantle. Subduction introduces the newly formed heterogeneity to the deep mantle where it remixes with older subducted materials as well as the likely remnants of early Earth formation, differentiation, and magma ocean solidification (Elkins-Tanton, 2008; Labrosse et al., 2007). Large scale mantle convection associated with plate tectonics also remixes past and newly generated heterogeneities leading to a "marble cake" mantle (Allègre & Turcotte, 1986; van Keken et al., 2014) with its complicated geochemical history expressed in the broad heterogeneity of mid-oceanic ridge and ocean island basalts (Hofmann, 2014; van Keken et al., 2002; Zindler & Hart, 1986), with indications of preservation of heterogeneity caused by very early differentiation of the Earth's mantle (Boyet & Carlson, 2005) even in modern basalts (Horan et al., 2018). The long-term recycling of oceanic crust mixed in with the depleted harzburgite-derived component and ambient mantle is an attractive explanation (Christensen & Hofmann, 1994; Jones et al., 2020) for the bulk of the seismologically observed Large Low Shear Velocity Provinces (LLSVPs: Garnero & McNamara, 2008) at the base of the mantle. However, it is as yet not clear how significant a contribution any primordial heterogeneity has to these LLSVPs (Ballmer et al., 2016; Li et al., 2014) and whether other, potentially more exotic, processes are at play to maintain chemical heterogeneity (Ballmer et al., 2017; Kellogg et al., 1999).

To aid in the understanding of the thermal and chemical evolution of the Earth's mantle we can test hypotheses using predictive models of mantle convection driven by thermal and chemical buoyancy forces. This requires reliable thermochemical methods that are based on the numerical solution of the governing equations following from conservation of momentum, mass, energy, and chemical species. Significant work has relied on the incompressible Boussinesq approximation (e.g., Brandenburg et al., 2008; Christensen & Hofmann, 1994; Li & McNamara, 2018; McNamara & Zhong, 2004; van Summeren et al., 2009; Zhong, 2006). A more realistic approach, that also allows for better comparison with seismological observations and for use of realistic equations of state as determined from mineral physics (Stixrude & Lithgow-Bertelloni, 2011), is to use (weakly) compressible convection that takes into account the compression of the Earth's mantle under its own weight. The anelastic liquid approximation (ALA) (Jarvis & McKenzie, 1980) has been used increasingly to take into account the effects of density increasing into the Earth's interior, along with associated latent heat and buoyancy effects of solid-solid phase transitions, and the interplay of viscous

dissipation and adiabatic heating and cooling caused by work against gravity (Schubert et al., 2001). A growing body of work reflects the community's interest in solving a more complete set of compressible equations, including ALA, instead of the simplified Boussinesq equations (Glatzmaier, 1986; Tackley, 1996; Tan & Gurnis, 2007; Nakagawa et al., 2010; Bossman & van Keken, 2013; Gassmöller et al., 2020).

Aside from the solution of the Stokes and mass conservation equations we generally need to solve time-dependent advection-diffusion equations for thermal and compositional fields that, in their simplest form, can be written

$$\frac{\partial \phi}{\partial t} + \mathbf{u} \cdot \nabla \phi - \nabla \cdot \kappa_\phi \nabla \phi \;=\; 0. \tag{1}$$

Here ϕ is the physical quantity being advected and diffused, κ_ϕ is the thermal or chemical diffusivity, and $\mathbf{u}$ is the velocity field. A measure of the importance of the advective over diffusive terms is the Péclet number $\mathrm{Pe}_\phi = LU/\kappa_\phi$ where L and U are a representative length and velocity scales respectively. In typical models of whole mantle convection $\mathrm{Pe}_\phi \backsim \mathcal{O}(10^4)$ for thermal advection-diffusion but becomes practically infinite for chemical species due to their very low effective diffusivity.

Various methods have been used in geodynamics to solve (1) including: 1. field-based methods where temperature or composition are discretized on the mesh (e.g., Hansen & Yuen, 2000; He et al., 2017; Kellogg & King, 1993); 2. interface-tracking methods where boundaries between distinct thermal or compositional regions are advected through the mesh using, for example, the marker chain (Christensen & Yuen, 1984; Lin & van Keken, 2006; Schmeling, 1987; van Keken et al., 1997), volume or moment of fluid (Pilliod & Puckett, 2004; Robey & Puckett, 2019; Zalesak, 1979), or level-set (Hillebrand et al., 2014; Samuel & Evonuk, 2010; Suckale et al., 2010) methods; 3. tracer methods where individual tracers carry thermal or chemical information as they are advected pointwise (Brandenburg et al., 2008; Christensen & Hofmann, 1994; Gerya & Yuen, 2003; O'Neill et al., 2006; Tackley & King, 2003).

While field-based methods are generally robust when the Péclet number is moderate, they tend to become unstable or unsuitable in the case of high Pe_ϕ and may suffer from numerical oscillations or numerical over-diffusion. Interface-tracking and tracer-based methods on the other hand are appropriate in zero-diffusivity scenarios but are more difficult to use when Pe_ϕ is finite. In the case of negligible diffusion, tracer methods are particularly well suited for the solution of (1) since the positions of tracers can be traced over time by solving ordinary differential equations after interpolation of the velocity field $\mathbf{u}$ at the tracers' positions. However, even when $\mathbf{u}$ is incompressible, its numerical approximation may not be. This incompressibility approximation error may cause tracers to drift apart, creating gaps in the data and requiring the addition or deletion of particles (e.g., Moresi et al., 2003). Unless treated carefully this can introduce issues with conservation and artificial diffusion or mixing.

Tracer drift can be avoided in incompressible flows if the numerical velocity approximation can be guaranteed to be pointwise divergence-free (Jenny et al., 2001; McDermott & Pope, 2008; Wang et al., 2015) and the mesh spacing is regular. These requirements ensure that the number of tracers per cell remains roughly constant. In Sime et al. (2021) we demonstrated this using a hybrid discontinuous Galerkin (HDG) finite element (FE) method (Cockburn et al., 2010; Labeur & Wells, 2012; Maljaars et al., 2018; Rhebergen & Wells, 2018) to discretize the incompressible Stokes equation. The first major goal of this paper is to extend our presentation of the HDG FE discretization of the Stokes equations to a more general and compressible mass-conservation equation and investigate its properties.

Regardless of the discretization used, the physical dilation and contraction introduced by compressible flows means that tracer drift is unavoidable. It is also unlikely that the requirement of regular mesh spacing will be generally satisfied. It therefore becomes necessary to add and remove tracers to maintain data coverage across the domain. Appropriate

methods for doing this while maintaining other important properties, such as conservation, depend on how the compositional or thermal tracer data is transferred to the mesh for use in setting the buoyancy or other material parameters.

Many transfer methods have been developed principally for application to compositional data and chemical buoyancy. In the Dirac delta source, or 'Stokeslet' method (e.g., Christensen & Hofmann, 1994), each particle is normally associated with a mass that is used to derive the buoyancy source. When setting up the initial distribution or adding and subtracting tracers during a simulation, care must be taken to (re-)distribute this mass in a manner consistent with the background density field. Errors introduced during this process will affect both conservation and the flow field. The so-called tracer ratio or absolute methods (Tackley & King, 2003), where tracers or their ratios are counted per cell, still require careful consideration of the positioning of new tracers to prevent artificial diffusion. The addition or removal of tracers will also generally adversely affect the conservation properties of the absolute method. The partial differential equation (PDE)-constrained l_2 projection method (Maljaars et al., 2019, 2020) projects tracer data to the mesh under the constraint that it satisfies a PDE. When advecting compositional data we showed (Sime et al., 2021) that this method will guarantee conservation. The second major goal of this paper is then to demonstrate that it also allows particles to be added or removed without affecting these important conservation properties.

We further note that operator-splitting (Lanser & Verwer, 1999) offers a mechanism to extend tracer methods beyond the zero-diffusion limit. These solve for the advective part of the problem on the tracers while diffusing the data on the mesh after projection. Care must be taken when discretizing each stage to ensure that accuracy is maintained but these methods are essential in bridging the gap between the negligible and small diffusion regimes, for example when chemical diffusion is present or in high Péclet number thermally-driven flows. They also offer the tantalizing possibility of exceeding the mesh-dependent Courant limit on time step size. The third major goal of this paper is then to demonstrate the use of the PDE-constrained l_2 projection method to solve (1) from κ_ϕ=0 (as in Sime et al., 2021) to finite κ_ϕ. This also allows us to use tracer methods on the temperature field.

Several approaches are available to estimate the accuracy and applicability of computational methods for mantle convection studies. These include community benchmarks (Blankenbach et al., 1989; Busse et al., 1994; King et al., 2010; van Keken et al., 1997), or internal convergence tests and comparisons with physical or manufactured analytical solutions (e.g., Curbelo et al., 2019; Kronbichler et al., 2012; Zhong et al., 2008). The comparison with analytical solutions is attractive as it allows for direct determination of the error in approximation, but can be limited to the investigation of individual aspects of the governing equations rather than the full system. The fourth major goal of this paper is to present new manufactured solutions that satisfy the full set of coupled compressible convection equations to allow us to determine the absolute error of the numerical solution to the full coupled system, as well as the spatial and temporal orders of convergence. We hope these new manufactured solutions will be used to test existing and newly developed computational techniques for mantle convection modeling.

1.1 Tracers in compressible flows with diffusion

In our prior work (Sime et al., 2021) we highlighted the necessity for exact satisfaction of the continuity constraint in the *incompressible* Stokes system for precise advection of tracer data. When combined with a PDE-constrained l_2 projection method (Maljaars et al., 2019) we are also able to exactly conserve mass when advecting chemical data without diffusion. Our aim here is to extend this work to the weakly compressible Stokes system coupled to the heat equation (such as described in the ALA benchmark of King et al., 2010). This requires us to achieve the following goals:

1. Satisfy the compressible continuity equation exactly (pointwise) to conserve mass.
2. Achieve exact conservation of the field discretized by tracers in a compressible flow regime.
3. Retain an approximately constant number of tracers per cell in a mesh for computational efficiency and data resolution.
4. Facilitate diffusion of the tracer discretized field.

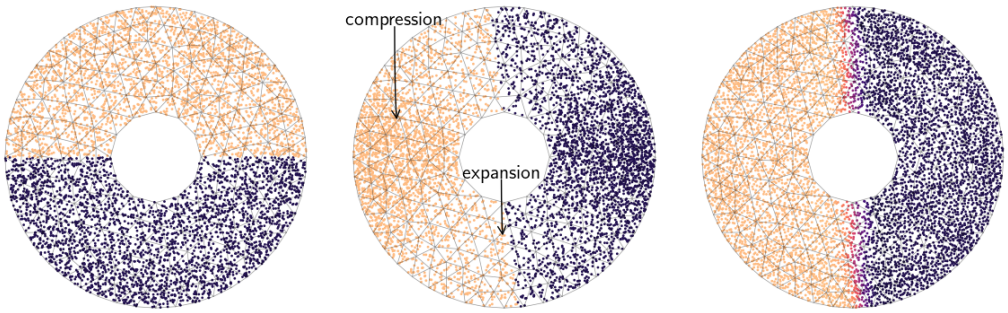
Point 1. is offered by the HDG discretization of the compressible Stokes system, which provides a momentum approximation that exactly satisfies the continuity equation (shown in section 5). Assuming point 1. is achieved, point 2. follows in a straightforward manner from the PDE-constrained l_2 projection method employed in Sime et al. (2021) and introduced by Maljaars et al. (2019). Maintaining exact conservation while achieving point 3. is addressed in section 2.3 where we note that by discretizing a field by tracers, we may add or remove tracers without affecting mass conservation. The PDE-constrained l_2 projection of the tracers to the underlying field ensures mass is exactly conserved. Point 4. is addressed in sections 4 and 6.3 for second order advection-diffusion problems. In this setting the advection and diffusion components are considered individually by operator splitting. The advection component is treated as in Sime et al. (2021) whilst we discretize the diffusion component employing the method introduced in Maljaars et al. (2018).

A sketch of the issues we address is shown in figure 1. Here we have manufactured an anticlockwise rotational compressible velocity field FE approximation in a cylindrical annulus geometry with no flow through the internal or external boundaries. The system is evolved for approximately one quarter rotation. We show the impact of the compressible velocity field on the tracer distribution with and without tracer addition/removal and diffusion in figure 1(b) and figure 1(c), respectively. In figure 1(b) we note that tracers congregate and disperse in compressive and expansive flow regions, respectively. When using a PDE-constrained l_2 projection, the congregation and dispersion of tracers has no bearing on the mass localized to those regions (cf. Maljaars et al., 2018; Sime et al., 2021). However, the congregation of tracers leads to computational inefficiency by over prescription of data per cell, whilst dispersion of tracers yields insufficient data to resolve the underlying discretized field. Figure 1(c) depicts the benefits of the methods shown in this work by including diffusion effects and addition/removal of tracers to preserve approximately uniform tracer coverage in each cell of the mesh.

1.2 HDG methods

To accommodate the necessity for exact mass conservation we require a spatial discretization scheme of the compressible Stokes system which exactly satisfies the continuity constraint at all points in the domain. Many FE methods exist with this property (Cockburn et al., 2007; Evans & Hughes, 2013; Guzmán & Neilan, 2014; Morgan & Scott, 2018; Scott & Vogelius, 1985). However we employ the HDG FE method (as developed in Rhebergen & Wells, 2020), where the projection of the continuity constraint onto the discontinuous pressure space matches with the formulation of the continuity constraint itself, yielding exact pointwise satisfaction (see Rhebergen & Wells, 2018).

The HDG scheme has been applied to the discretization of the compressible Navier–Stokes system (La Spina et al., 2020; Peraire et al., 2010; Woopen et al., 2014) and the weakly compressible Navier–Stokes system (Vila-Pérez et al., 2021). However, we are concerned with the weakly compressible Stokes system. To this end we will present an abstract interior penalty (IP) HDG formulation of second order diffusion operators which will be applied to the weakly compressible Stokes system such as that found in the ALA model. This abstract formulation will further serve us by providing a means to discretize the diffusive component of the advection-diffusion equation. This advection-diffusion approximation is then used in updating tracer data.



(a) Initial state

(b) Neither diffusion nor tracer addition/removal

(c) $\mathrm{Pe}_\phi \approx 10^3$ diffusion and tracer addition/removal

Figure 1: Here we sketch the challenges we address regarding tracer discretizations in compressible velocity fields in this work. We manufacture a compressible system with a rotational flow acting anti-clockwise in the cylindrical annulus geometry with no flow through the inner and outer radii, 0.4292 and 1.4292 respectively. (a) The initial state of a field discretized by an even coverage of tracers where color is added as a visual aid only. (b) After approximately one quarter revolution that with no addition or removal the tracers congregate in regions of compression and disperse in regions of expansion leading to inefficient tracer coverage. (c) Even coverage is obtained by addition and removal of tracers and diffusion across tracers as required by, for example, thermal or chemical diffusivity. The diffusion of the field discretized by tracers can be seen in the transition zone from the orange to purple regions.

We found in Sime et al. (2021) that precise approximation of the incompressibility, $\nabla \cdot \mathbf{u}_h = 0$, of the velocity field approximation $\mathbf{u}_h$, is crucial for precise advection of tracer data. In the compressible scheme we therefore pay close attention to the compressibility approximation $\nabla \cdot \mathbf{u}_h \neq 0$ error. In section 6.2 we show that the compressibility approximation error is superconvergent.

1.3 Outline

Our investigation proceeds as follows: In section 2 we will outline the compressible Stokes system underlying the numerical models we wish to discretize. Furthermore we will briefly cover the paradigm of temporal discretization of the advection-diffusion equation by operator splitting, and spatial discretization by subdivision of a finite domain into a mesh for computation of FE approximations. In section 2.3 we will reintroduce the concept of discretizing a field with tracers as previously investigated in, for example, Maljaars et al. (2018) and Sime et al. (2021). This includes the method by which we add and remove tracers during a simulation. Section 3 covers the detail of PDE-constrained l_2 projection of tracer data to a field using the HDG FE method in the context of compressible flows.

The abstract IP HDG formulation for the discretization of second order diffusion operators is presented in section 2.4. This abstraction is then applied to the diffusion component of the advection-diffusion equation in section 4 and the compressible Stokes system in section 5. We complete the technical detail of the numerical scheme in section 6.3 by reproducing the result in Maljaars et al. (2019), in the context of compressible flow, which shows careful selection of intermediate states for the tracer projection component of the operator splitting is necessary to ensure a mathematically consistent formulation.

Penultimately in section 7 we demonstrate numerical experiments. We show optimal convergence of our method by comparing computations of approximations of a compressible Stokes system with a manufactured solution. Additionally we replicate the ALA model

community benchmark from King et al. (2010) and compare our results. Finally in section 8 we provide a summary of the numerical method's highlights and some concluding remarks.

While we focus in our development on the use of equation (1) for thermal fields, we note, and will demonstrate, that our new method is equally suited for solution of (1) under the limit $\kappa_\phi \to 0$. This provides therefore a new and highly precise approach to solve the general thermochemical convection equations for mantle convection.

2 Preliminaries

This section is dedicated to an overview of the compressible system and the means by which we handle spatial and temporal discretization. Specifically we will cover the HDG FE scheme in space and an operator splitting in conjunction with a θ differencing scheme in time. The remainder of this work discusses the technical detail and implementation of these numerical methods. Our complete method to solve the coupled Stokes, mass conservation, and heat advection-diffusion equations is outlined in algorithm 1.

Algorithm 1: Summary of solution procedure

```
n ← 0
t ← 0
Initialize T_h(x, t = 0)
Solve for initial momentum (ρu)_h(x, t = 0) (section 5)
while t ≤ t_F (in time loop) do
    Compute Δt_n (section 6.1)
    Balance number of tracers per cell (section 2.6)
    foreach Runge–Kutta sub step do
        Compute u_h and advect tracers to new sub step position
        Project advective component yielding T_{a,h}^{n+1} (section 3)
        Solve diffusive correction yielding T_{d,h}^{n+1} (section 4)
        Solve for updated momentum (ρu)_h (section 5)
    Update tracer data T_p^{n+1}, p = 1, ..., N_p (section 6.3)
    t ← t + Δt_n
    n ← n + 1
```

2.1 Approximate equations for thermal convection in a weakly compressible fluid

The approximate, non-dimensionalized equations which govern velocity $\mathbf{u}$, temperature T, and pressure p are

$$-\nabla \cdot \underline{\underline{\sigma}} = \mathbf{f}, \tag{2}$$

$$\nabla \cdot (\rho \mathbf{u}) = 0, \tag{3}$$

$$\rho \frac{\mathrm{D}T}{\mathrm{D}t} - \nabla \cdot k \nabla T = H, \tag{4}$$

where we have assumed a constant non-dimensional heat capacity of one. Here, the stress and rate of strain tensors are

$$\underline{\underline{\sigma}}(\mathbf{u}, p) = 2\eta \left(\underline{\underline{\dot{\varepsilon}}}(\mathbf{u}) - \frac{1}{3} (\nabla \cdot \mathbf{u}) \underline{\underline{\mathrm{I}}} \right) - p \underline{\underline{\mathrm{I}}}, \tag{5}$$

$$\underline{\underline{\dot{\varepsilon}}}(\mathbf{u}) = \frac{1}{2} \left(\nabla \mathbf{u} + \nabla \mathbf{u}^\top \right), \tag{6}$$

respectively, $\rho(\mathbf{x})$ is the density, $\eta(\mathbf{x},t)$ is the viscosity, $k(\mathbf{x},t)$ is the thermal conductivity, $\mathbf{f}(\mathbf{x},t)$ is the momentum source and $H(\mathbf{x},t)$ is the heat source. The convective derivative is defined by

$$\frac{\mathrm{D}T}{\mathrm{D}t} = \frac{\partial T}{\partial t} + \mathbf{u}\cdot\nabla T. \tag{7}$$

2.2 Advection-diffusion on tracers by operator splitting

Our goal is to discretize advection-diffusion problems such as the heat equation (4) using tracers. The case of pure advection ($k = 0$) has been specifically covered in Sime et al. (2021) where the l_2 (least squares) and PDE-constrained l_2 tracer projection methods were demonstrated. The latter method (introduced in Maljaars et al., 2019) was shown to exactly conserve mass (to machine precision). The case of advection-diffusion ($k>0$) requires some additional work. We will exploit the exact conservation of the PDE-constrained l_2 projection method and augment our numerical scheme with the capability of diffusing thermal data between tracers (as introduced in Maljaars et al., 2018).

The premise of the discretization scheme is to split the advective and diffusive components of the heat equation for individual consideration. For the advection terms we exploit the tracer projection method shown in Sime et al. (2021). We then compute the diffusive correction of the projected field followed by updating the tracer data. In sequence this involves:

1. Advect tracer data in the background velocity approximation.
2. Project advective component of the advection-diffusion equation (the material derivative).
3. Diffuse the intermediate advection field yielding the next time step's field approximation.
4. Update the tracer data with the diffused field.

As part of this procedure we split equation (4) into two parts and seek the advective and diffusive field components, T_a and T_d respectively, by solving

$$\rho\frac{\partial T_a}{\partial t} + \rho\mathbf{u}\cdot\nabla T_a = 0, \tag{8}$$

$$\rho\frac{\partial T_d}{\partial t} - \nabla\cdot k\nabla T_d = H, \tag{9}$$

which will be coupled through temporal discretization.

2.2.1 Coupling by temporal discretization

Let t be the current model time with initial time t_0=0 and final time t=t_F. These limits define the temporal interval of the simulation $t \in \mathcal{I}_t := [t_0, t_F]$. The simulation time interval $\mathcal{I}_t$ is discretized into steps

$$\mathcal{I}_{\Delta t} := \{t_0, t_1, \ldots, t_F\}, \quad t_0 < t_1 < \ldots < t_F. \tag{10}$$

Employing the operator splitting method in the time interval $\mathcal{I}_{\Delta t}$ we define the states $T_a^n = T_a(\mathbf{x}, t_n)$ and $T_d^n = T_d(\mathbf{x}, t_n)$ to be used in a θ time differencing scheme where

$$T_a^{n+\theta} = \theta T_a^{n+1} + (1-\theta)T_a^n \quad \text{and} \quad T_d^{n+\theta} = \theta T_d^{n+1} + (1-\theta)T_d^n, \tag{11}$$

and $\theta \in [0, 1]$ is a constant. We approximate the time derivatives by

$$\frac{\partial T_a}{\partial t} \approx \frac{T_a^{n+1} - T_a^{*,n}}{\Delta t_n} \quad \text{and} \quad \frac{\partial T_d}{\partial t} \approx \frac{T_d^{n+1} - T_d^{*,n}}{\Delta t_n}. \tag{12}$$

Here $T_a^{*,n}$ and $T_d^{*,n}$ are intermediate states and $\Delta t_n = t_{n+1} - t_n$. A simple first choice for the intermediate states would appear to be $T_a^{*,n} = T_d^n$ and $T_d^{*,n} = T_a^{n+1}$ with the initial

state $T_a^{*,0} = T(\mathbf{x}, t_0)$ (commensurate with Lie splitting), however, we will present a better, mathematically consistent choice in section 6.3.

The time discretization seeks advective and diffusive temperature states, T_a^{n+1} and T_d^{n+1}, respectively, such that

$$\rho \frac{T_a^{n+1} - T_a^{*,n}}{\Delta t_n} + \rho \mathbf{u} \cdot \nabla T_a^{n+\theta} = 0, \tag{13}$$

$$\rho \frac{T_d^{n+1} - T_d^{*,n}}{\Delta t_n} - \nabla \cdot k \nabla T_d^{n+\theta} = H^{n+\theta}. \tag{14}$$

2.3 Tracer discretization of a field

Our starting point is advecting a distribution of tracers which discretize the field T as introduced in Maljaars et al. (2018) and adopted in the context of purely advective geodynamic flows in Sime et al. (2021).

Consider N_p tracers with position and temperature data

$$\mathrm{X}(t) := \left\{ \mathbf{x}_p(t) \right\}_{p=1}^{N_p}, \tag{15}$$

$$\mathrm{T}(t) := \left\{ T_p(t) \right\}_{p=1}^{N_p}, \tag{16}$$

respectively. The initial values of these tracer data are interpolated from the initial temperature field

$$\mathrm{T}(t=0) = \left\{ T(\mathbf{x}_p, t=0) \right\}_{p=1}^{N_p}. \tag{17}$$

It is important to note that the tracer data do not carry a notion of mass or volume. They are simply pointwise discretizations of a continuum field. It is the conservation of these fields which is important when approximating underlying physical models.

2.3.1 Tracer data advection

Following Sime et al. (2021) we discretize the advection equation (13) by updating tracers' positions using a Runge–Kutta (RK) method of degree $\ell \in \mathbb{N}$ to numerically integrate the total derivatives in $\mathcal{I}_{\Delta t}$

$$\frac{\mathrm{d}\mathbf{x}_p}{\mathrm{d}t} = \mathbf{u}(\mathbf{x}_p, t), \quad p = 1, \ldots, N_p. \tag{18}$$

We also note the total derivative of the temperature data

$$\begin{aligned} \rho(\mathbf{x}_p) \frac{\mathrm{d}T_p}{\mathrm{d}t} &= \rho(\mathbf{x}_p) \left(\frac{\partial T(\mathbf{x}_p, t)}{\partial t} + \frac{\mathrm{d}\mathbf{x}_p}{\mathrm{d}t} \cdot \nabla T(\mathbf{x}_p, t) \right), \\ &= \rho(\mathbf{x}_p) \left(\frac{\partial T(\mathbf{x}_p, t)}{\partial t} + \mathbf{u}(\mathbf{x}_p, t) \cdot \nabla T(\mathbf{x}_p, t) \right), \\ &= 0 \qquad p = 1, \ldots, N_p, \end{aligned} \tag{19}$$

by equation (8), where $T = T_a$.

2.4 Abstract HDG FE formulation for diffusion problems

Throughout this work we employ the HDG method for spatial discretization for its benefits of exact (global, local, and pointwise) conservation, in addition to lending itself well to global problem size reduction by static condensation[1]. However, the verbosity and

[1] The process by which the local degrees of freedom defined in mesh cells may be eliminated in favor of global degrees of freedom defined on the mesh facets. Also known as Guyan reduction (Guyan, 1965).

complexity of the formulations invite human error in their computational implementation. In an effort to alleviate these concerns we present here the abstract HDG FE formulation for diffusion problems. Employing this abstraction in computational symbolic algebra (Alnæs et al., 2014) allows for automatic formulation of HDG terms (using the techniques in Houston & Sime, 2018). The abstract formulation follows from the techniques developed in the discontinuous Galerkin (DG) literature (Hartmann & Houston, 2006) and the work on the HDG formulation of the diffusion operator in Labeur and Wells (2012) and Rhebergen and Wells (2017).

2.4.1 Domain discretization

Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be the domain of interest with boundary $\partial\Omega$ which has unit outward point unit vector $\mathbf{n}$. To impose boundary data we subdivide $\partial\Omega$ into Dirichlet and Neumann components, $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$, respectively, such that the Dirichlet component is non-empty and the Dirichlet and Neumann components do not overlap $\partial\Omega_D \cap \partial\Omega_N = \emptyset$.

The domain Ω is subdivided into conforming simplices (d=2 triangles or d=3 tetrahedra). Each simplex is called a cell or element κ in the mesh $\mathcal{T}_h$, such that $\mathcal{T}_h := \{\kappa\}$. The outward pointing normal unit vector on the boundary of each cell $\partial\kappa$ is $\mathbf{n}_\kappa$. Each cell has size h_κ measured by the diameter of the smallest circle or sphere in which the cell may be contained. We write the quantity h to be the maximum cell size of all cells in a mesh and the subscript h denotes a quantity which has been discretized conforming to an underlying mesh. Interior facets are those shared by two adjacent cells. Exterior facets are those cell boundaries for which $\partial\kappa$ overlaps $\partial\Omega$.

2.4.2 HDG and EDG FE function spaces

Here we define the FE function spaces we will use later for the spatial discretization of fields conforming to the mesh $\mathcal{T}_h$. Considering we are concerned with both the discretization of the heat equation and the compressible Stokes system, we present these spaces in an abstract setting. This abstraction will then be specialized for the individual cases in the later sections.

Let $\mathbf{w}$ be the $m \in \mathbb{N}$ dimensional abstract solution of an appropriate PDE. The HDG formulation seeks the approximate solution composed of a hybrid of functions defined on the cells and facets, $\mathbf{w}_h$ and $\overline{\mathbf{w}}_h$, respectively. The choice of the FE function spaces in which these solutions are sought impact the numerical properties of the approximation. Consider the following FE function spaces in which we will seek our HDG approximations:

1. $\left[W_{\mathrm{DG}}^{h,\mathsf{k}}\right]^m := \{m-$dimensional piecewise polynomials of degree k defined and continuous in each $\kappa \in \mathcal{T}_h$, and *discontinuous* between elements$\}$,
2. $\left[\overline{W}_{\mathrm{DG}}^{h,\mathsf{k}}\right]^m := \{m-$dimensional piecewise polynomials of degree k defined and continuous on each facet in $\mathcal{T}_h$, and *discontinuous* between facets$\}$,
3. $\left[\overline{W}_{\mathrm{CG}}^{h,\mathsf{k}}\right]^m := \{m-$dimensional piecewise polynomials of degree k defined and continuous on each facet in $\mathcal{T}_h$, and *continuous* between facets$\}$.

The combinations $(\mathbf{w}_h, \overline{\mathbf{w}}_h) \in \left[W_{\mathrm{DG}}^{h,\mathsf{k}}\right]^m \times \left[\overline{W}_{\mathrm{DG}}^{h,\mathsf{k}}\right]^m$ and $(\mathbf{w}_h, \overline{\mathbf{w}}_h) \in \left[W_{\mathrm{DG}}^{h,\mathsf{k}}\right]^m \times \left[\overline{W}_{\mathrm{CG}}^{h,\mathsf{k}}\right]^m$ we call the hybrid discontinuous Galerkin (HDG) and embedded discontinuous Galerkin (EDG) methods, respectively.

Throughout this work we exploit the EDG approximation for the reduced global problem size owing to fewer degrees of freedom arising on the facets. However, as we shall specify later in section 5.1 we employ a combined EDG-HDG method for the Stokes system momentum and pressure, respectively. This element pairing yields a pointwise divergence free momentum approximation.

2.4.3 Abstract diffusion problem

Consider the following homogeneous second order PDE where we seek the m dimensional solution $\mathbf{w}$ subject to appropriate Dirichlet and Neumann boundary data, $\mathbf{w}_D$ and $\mathbf{w}_N$, respectively,

$$-\nabla \cdot \underline{\underline{\tau}}(\mathbf{w}, \nabla\mathbf{w}) = \mathbf{0} \qquad \text{in } \Omega, \tag{20}$$

$$\underline{\underline{\tau}}(\mathbf{w}, \nabla\mathbf{w}) \cdot \mathbf{n} = \mathbf{w}_N \quad \text{on } \partial\Omega_N, \tag{21}$$

$$\mathbf{w} = \mathbf{w}_D \quad \text{on } \partial\Omega_D. \tag{22}$$

Here, $\underline{\underline{\tau}}(\mathbf{w}, \nabla\mathbf{w})$ is an $m \times d$ tensor which is linear or nonlinear in $\mathbf{w}$ and linear in $\nabla\mathbf{w}$. For example, in the context of thermal diffusion, $m = 1$, $\mathbf{w} = T$ and $\underline{\underline{\tau}} = k\nabla T$. In the context of the compressible Stokes system momentum equation $m = d$, $\mathbf{w} = \mathbf{u}$ and $\underline{\underline{\tau}} = \underline{\underline{\sigma}}$.

2.4.4 Abstract IP HDG FE formulation

The construction of the abstract HDG FE formulation demands we homogenize equation (20). To this end we construct the homogeneity tensor $G(\mathbf{w})$ such that

$$(G_{ij})_{kl} = \frac{\partial \underline{\underline{\tau}}_{ij}}{\partial (\nabla\mathbf{w})_{kl}}, \quad i,k = 1,\ldots,m, \quad j,l = 1,\ldots,d, \tag{23}$$

and the product and transpose product operations (using Einstein summation notation) read

$$(G\nabla\mathbf{w})_{ij} = (G_{ij})_{kl}(\nabla\mathbf{w})_{kl} = G_{ij} : \nabla\mathbf{w}, \tag{24}$$

$$(G^\top\nabla\mathbf{w})_{ij} = (G_{ij})_{kl}(\nabla\mathbf{w})_{ij}. \tag{25}$$

Employing the homogeneity tensor we seek the abstract solution variable approximation $(\mathbf{w}_h, \overline{\mathbf{w}}_h) \in \left[W_{\mathrm{DG}}^{h,\mathsf{k}}\right]^m \times \left[\overline{W}_{\mathrm{DG}}^{h,\mathsf{k}}\right]^m$, such that

$$\begin{aligned} F := &\sum_{\kappa\in\mathcal{T}_h} \int_\kappa \underline{\underline{\tau}} : \nabla\mathbf{z}_h \,\mathrm{d}\mathbf{x} + \sum_{\kappa\in\mathcal{T}_h} \int_{\partial\kappa} (\widehat{\mathbf{w}} - \mathbf{w}_h) \otimes \mathbf{n}_\kappa : G^\top\nabla\mathbf{z}_h \,\mathrm{d}\mathbf{s} \\ &- \sum_{\kappa\in\mathcal{T}_h} \int_{\partial\kappa} \hat{\underline{\underline{\tau}}} : \mathbf{z}_h \otimes \mathbf{n}_\kappa \,\mathrm{d}\mathbf{s} + \sum_{\kappa\in\mathcal{T}_h} \int_{\partial\kappa} \hat{\underline{\underline{\tau}}} : \overline{\mathbf{z}}_h \otimes \mathbf{n}_\kappa \,\mathrm{d}\mathbf{s} \equiv 0 \end{aligned} \tag{26}$$

for all $(\mathbf{z}_h, \overline{\mathbf{z}}_h) \in \left[W_{\mathrm{DG}}^{h,\mathsf{k}}\right]^m \times \left[\overline{W}_{\mathrm{DG}}^{h,\mathsf{k}}\right]^m$. Here the numerical flux functions are given by

$$\widehat{\mathbf{w}} = \overline{\mathbf{w}}_h, \tag{27}$$

$$\hat{\underline{\underline{\tau}}} = \underline{\underline{\tau}} + \frac{C_{\mathrm{IP}}}{h_\kappa} G\left((\overline{\mathbf{w}}_h - \mathbf{w}_h) \otimes \mathbf{n}_\kappa\right), \tag{28}$$

where $C_{\mathrm{IP}} \geq C_{\mathrm{IP}_0} > 0$ is an IP parameter independent of the mesh where the (typically unknown) minimum value C_{IP_0} ensures stability of the numerical method. Specific choices of C_{IP} will be stated in the following sections which are chosen based on observed stability thresholds of small computational experiments.

2.4.5 Implementation

We note here that construction of the homogeneity tensor and the residual formulation may appear daunting. However with the unified form language (UFL) (Alnæs et al., 2014) and the techniques described in Houston and Sime (2018) the specification of such formulations is straightforward. Furthermore, the symbolic computation of the Gâteaux derivative for use with Newton's iterative method follows using automatic differentiation (available in the UFL). Specific examples applied to a range of problems for use with the components of the FEniCS project (Alnæs et al., 2015) are available in the public repositories Houston and Sime (2018) and Sime (2021).

2.5 Momentum as a solution variable

The standard FE approximation of the compressible Stokes system equations (2) and (3) would seek velocity and pressure approximations $\mathbf{u}_h$ and p_h, respectively. However, here we instead compute momentum and pressure FE approximations $(\rho\mathbf{u})_h$ and p_h, respectively. This is motivated by the HDG method momentum approximation of the compressible Stokes system exactly satisfying the continuity constraint of equation (3) pointwise. By this we mean that

$$\nabla \cdot (\rho\mathbf{u})_h(\mathbf{x}, t) = 0 \quad \forall \mathbf{x} \in \mathcal{T}^h, \tag{29}$$

where $(\rho\mathbf{u})_h$ is the momentum solution variable computed from the HDG formulation. Furthermore this property is a key component in our tracer projection method being exactly conservative. More detail is provided in section 5.

2.6 Balancing the number of tracers per mesh cell

By the nature of compressible flow, the distribution of tracers will congregate and disperse in compressive and expansive regions, respectively. In order to preserve an accurate approximation of a field discretized by tracers we strive to keep the number of tracers per cell roughly constant throughout a simulation.

We remind ourselves that the field which our tracer method discretizes is the host of physical continuum data. Therefore we may add tracers with data interpolating their corresponding field and likewise delete tracers (maintaining a minimum resolution defined later) at will. We emphasize that this addition and removal of tracers will *not* impact conservation of the underlying field.

To this end we define the following tracer addition and removal method. Let N_{p_κ} be the number of tracers in cell $\kappa \in \mathcal{T}_h$. Furthermore let the upper and lower limit of the number of tracers per cell be $N_{p_\kappa}^{\max}$ and $N_{p_\kappa}^{\min}$, respectively, where $N_{p_\kappa}^{\max} \geq N_{p_\kappa}^{\min} > 0$. In each cell $\kappa \in \mathcal{T}_h$:

1. If $N_{p_\kappa} > N_{p_\kappa}^{\max}$ then $N_{p_\kappa} - N_{p_\kappa}^{\max}$ tracers with the shortest distances to a neighboring tracer are removed.
2. If $N_{p_\kappa} < N_{p_\kappa}^{\min}$ then $N_{p_\kappa}^{\min} - N_{p_\kappa}$ tracers are added with positions generated from a uniform random distribution defined on the cell's geometry, and their data set to the interpolant of the underlying field, $T_h(\mathbf{x}_{p_{\text{new}}}, t)$, for each new tracer at position $\mathbf{x}_{p_{\text{new}}}$.

3 Advection of tracer data and their projection to a field

First we consider the advection problem equation (13). We employ the methods which have been extensively detailed in Sime et al. (2021) regarding tracer advection and projection to a field. We will summarize this technique in this section by stating the PDE-constrained l_2 projection scheme in the context of compressible flow. Two key additions are the advection through the background momentum field $(\rho\mathbf{u})_h$ and the choice of the intermediate state approximation $T_{a,h}^{*,n}$ (following Maljaars et al., 2019).

PDE-constrained l_2 projection minimizes the square distance between the field and the tracers whilst constrained by the advection PDE (8). The full problem reads: find $T_{a,h}^{n+1} \in \left[W_{\text{DG}}^{h,\mathsf{s}}\right]^1$ such that

$$\min_{T_{a,h}^{n+1} \in W_{\text{DG}}^{h,\mathsf{s}}} \mathcal{J}(T_{a,h}^{n+1}) := \sum_p^{N_p} \frac{1}{2} \left(T_{a,h}^{n+1}(\mathbf{x}_p^{n+1}) - T_p^n \right)^2 \tag{30a}$$

$$\text{subject to:} \quad \begin{aligned} \rho \frac{\partial T_{a,h}}{\partial t} + (\rho\mathbf{u})_h \cdot \nabla T_{a,h}^{n+\theta} &= 0 && \text{in } \Omega, \\ T_{a,h}^{n+1} &= T_D(\mathbf{x}, t_{n+1}) && \text{on } \partial\Omega_{\text{in}}, \end{aligned} \tag{30b}$$

where $\mathsf{s} \in \mathbb{N}$ is the polynomial degree of temperature FE approximation, $\partial\Omega_{\text{in}} = \{\mathbf{x} \in \partial\Omega : \mathbf{u}_h \cdot \mathbf{n} < 0\}$ is the inlet boundary and we define $\partial\Omega_{\text{out}} = \partial\Omega \setminus \partial\Omega_{\text{in}}$ as the outlet boundary.

In order to distinguish the FE function spaces in which we seek the advective and diffusive components of the temperature approximation, we define

$$S_{\text{DG}}^{h,\mathsf{s}} := \left[W_{\text{DG}}^{h,\mathsf{s}}\right]^1, \quad \overline{S}_{\text{DG}}^{h,\mathsf{s}} := \left[\overline{W}_{\text{DG}}^{h,\mathsf{s}}\right]^1, \quad \overline{S}_{\text{CG}}^{h,\mathsf{s}} := \left[\overline{W}_{\text{CG}}^{h,\mathsf{s}}\right]^1, \tag{31}$$

and the spaces enforcing Dirichlet temperature data $T_D(\mathbf{x}, t)$ and homogenized boundary data

$$\overline{S}_{\text{BC}}^{h,\mathsf{s}} := \{\overline{s} \in \overline{S}^{h,\mathsf{s}} : \overline{s}|_{\partial\Omega_D} = T_D\} \quad \text{and} \quad \overline{S}_{\text{BC}_0}^{h,\mathsf{s}} := \{\overline{s} \in \overline{S}^{h,\mathsf{s}} : \overline{s}|_{\partial\Omega_D} = 0\}, \tag{32}$$

respectively. Adhering to the method and notation in Sime et al. (2021) we forego stating the semi-discrete Lagrangian functional to be minimized and instead state the corresponding discretization of the equivalent linear system by the HDG method. Employing the θ scheme and time discretization discussed in section 2.2.1 and introducing a Lagrange multiplier λ_h, the EDG discretization of equations (30a) and (30b) reads: find $(T_{a,h}^{n+1}, \overline{T}_{a,h}^{n+1}, \lambda_h) \in S_{\text{DG}}^{h,\mathsf{s}} \times \overline{S}_{\text{CG}}^{h,\mathsf{s}} \times W_{\text{DG}}^{h,0}$ such that

$$\begin{aligned}
&\sum_{p=1}^{N_p} \left(T_{a,h}^{n+1}(\mathbf{x}_p^{n+1}) - T_p^n\right) s_h(\mathbf{x}_p^{n+1}) - \sum_{\kappa\in\mathcal{T}_h} \int_{\partial\kappa} \beta \left(\overline{T}_{a,h}^{n+1} - T_{a,h}^{n+1}\right) s_h \,\mathrm{d}\mathbf{s} \\
&+ \int_\Omega \rho \frac{s_h}{\Delta t_n} \lambda_h \,\mathrm{d}\mathbf{x} - \theta \sum_{\kappa\in\mathcal{T}_h} \int_\kappa (\rho\mathbf{u})_h^n \cdot \nabla\lambda_h s_h \,\mathrm{d}\mathbf{x} + \theta \int_{\partial\Omega_{\text{out}}} (\rho\mathbf{u})_h^n \cdot \mathbf{n} \lambda_h s_h \,\mathrm{d}\mathbf{s} = 0, \qquad (33)\\
&\int_\Omega \rho \frac{T_{a,h}^{n+1} - T_{a,h}^{*,n}}{\Delta t_n} \delta\lambda_h \,\mathrm{d}\mathbf{x} \\
&- \sum_{\kappa\in\mathcal{T}_h} \int_\kappa (\rho\mathbf{u})_h^n T_{a,h}^{n+\theta} \cdot \nabla\delta\lambda_h \,\mathrm{d}\mathbf{x} + \sum_{\kappa\in\mathcal{T}_h} \int_{\partial\kappa\setminus\partial\Omega} (\rho\mathbf{u})_h^n \cdot \mathbf{n} \overline{T}_{a,h}^{n+1} \delta\lambda_h \,\mathrm{d}\mathbf{s} \\
&+ \int_{\partial\Omega_{\text{out}}} (\rho\mathbf{u})_h^n \cdot \mathbf{n} T_{a,h}^{n+\theta} \delta\lambda_h \,\mathrm{d}\mathbf{s} + \int_{\partial\Omega_{\text{in}}} (\rho\mathbf{u})_h^n \cdot \mathbf{n} T_D^{n+\theta} \delta\lambda_h \,\mathrm{d}\mathbf{s} = 0, \qquad (34)\\
&\sum_{\kappa\in\mathcal{T}_h} \int_{\partial\kappa\setminus\partial\Omega_{\text{in}}} (\rho\mathbf{u})_h^n \cdot \mathbf{n}_\kappa \lambda_h \overline{s}_h \,\mathrm{d}\mathbf{s} + \sum_{\kappa\in\mathcal{T}_h} \int_{\partial\kappa} \beta \left(\overline{T}_{a,h}^{n+1} - T_{a,h}^{n+1}\right) \overline{s}_h \,\mathrm{d}\mathbf{s} = 0, \qquad (35)
\end{aligned}$$

for all $(s_h, \overline{s}_h, \delta\lambda_h) \in S_{\text{DG}}^{h,\mathsf{s}} \times \overline{S}_{\text{CG}}^{h,\mathsf{s}} \times W_{\text{DG}}^{h,0}$ (see Maljaars et al. (2019) and Sime et al. (2021) for derivation details).

4 Heat equation discrete formulation

In this section we define the backward Euler ($\theta = 1$) HDG formulation of the heat equation specified in equation (14) using the time discretization scheme discussed in section 2.2.1. We show later in section 6.3 that employing the backward Euler method in the split scheme with careful choice of the intermediate state $T_{a,h}^{*,n}$ can be shown to yield second order accuracy in time (Maljaars et al., 2019).

Reflecting on the abstract formulation in section 2.4 we note that in the case of the isotropic heat equation $\mathbf{w}_h = T_h$ and $\underline{\underline{\tau}} = k\nabla T_h$. We further note that in this simple case $G^\top \nabla T_h = G\nabla T_h = k\nabla T_h$. The HDG FE residual formulation reads: find $(T_{d,h}^{n+1}, \overline{T}_{d,h}^{n+1}) \in$

$S_{\mathrm{DG}}^{h,\mathsf{s}} \times \overline{S}_{\mathrm{BC}}^{h,\mathsf{s}}$ such that

$$\begin{aligned} F_T := & \int_\Omega \rho \frac{T_{d,h}^{n+1} - T_{d,h}^{*,n}}{\Delta t_n} s_h \,\mathrm{d}\mathbf{x} \\ & + \sum_{\kappa \in \mathcal{T}_h} \int_\kappa k \nabla T_{d,h}^{n+1} : \nabla s_h \,\mathrm{d}\mathbf{x} + \sum_{\kappa \in \mathcal{T}_h} \int_{\partial\kappa} (\widehat{T_d} - T_{d,h}^{n+1}) k \nabla s_h \cdot \mathbf{n}_\kappa \,\mathrm{d}\mathbf{s} \\ & - \sum_{\kappa \in \mathcal{T}_h} \int_{\partial\kappa} \widehat{k \nabla T_d} \cdot \mathbf{n}_\kappa s_h \,\mathrm{d}\mathbf{s} + \sum_{\kappa \in \mathcal{T}_h} \int_{\partial\kappa} \widehat{k \nabla T_d} \cdot \mathbf{n}_\kappa \overline{s}_h \,\mathrm{d}\mathbf{s} \\ & - \sum_{\kappa \in \mathcal{T}_h} \int_\kappa H(\mathbf{x}, t_{n+1}) s_h \,\mathrm{d}\mathbf{x} \equiv 0 \end{aligned} \tag{36}$$

for all $(s_h, \overline{s}_h) \in S_{\mathrm{DG}}^{h,\mathsf{s}} \times \overline{S}_{\mathrm{BC}_0}^{h,\mathsf{s}}$. The numerical flux functions are given by

$$\widehat{T_d} = \overline{T}_{d,h}^{n+1} \quad \text{and} \quad \widehat{k \nabla T_d} = k \nabla T_{d,h}^{n+1} + \frac{C_{\mathrm{IP}_T}}{h_\kappa} k \left(\overline{T}_{d,h}^{n+1} - T_{d,h}^{n+1} \right) \mathbf{n}_\kappa, \tag{37}$$

where C_{IP_T} is the penalty parameter selected specifically for the heat equation problem, chosen in this work (based on numerical experiment) to be $C_{\mathrm{IP}_T} = 24\mathsf{s}^2$.

5 Conserved momentum Stokes HDG formulation

Careful consideration of the weakly compressible formulation of the Stokes system in equations (2) and (3) is required in the context of tracer advection. Previous work in the context of *incompressible* flows demonstrated the necessity for precise approximation of the continuity constraint (Sime et al., 2021). Furthermore in Jones et al. (2021) it was found that imprecise approximation of the incompressibility constraint would lead to spurious model results.

5.1 Compressible Stokes EDG-HDG function spaces

The EDG-HDG compressible Stokes formulation seeks momentum and pressure approximations defined in the cells $((\rho\mathbf{u})_h, p_h)$ and on the facets $(\overline{(\rho\mathbf{u})}_h, \overline{p}_h)$, respectively. The choice of the FE function spaces in which these solutions are sought impact the numerical properties of the approximation (see Rhebergen & Wells, 2020).

As in Sime et al. (2021), we exploit the EDG-HDG formulation. In this setting we define the following function spaces:

$$\mathbf{V}_{\mathrm{DG}}^{h,\mathsf{p}} := \left[W_{\mathrm{DG}}^{h,\mathsf{p}} \right]^d, \qquad \overline{\mathbf{V}}_{\mathrm{CG}}^{h,\mathsf{p}} := \left[\overline{W}_{\mathrm{CG}}^{h,\mathsf{p}} \right]^d, \tag{38}$$

$$Q_{\mathrm{DG}}^{h,\mathsf{p}-1} := \left[W_{\mathrm{DG}}^{h,\mathsf{p}-1} \right]^1, \quad \overline{Q}_{\mathrm{DG}}^{h,\mathsf{p}} := \left[\overline{W}_{\mathrm{DG}}^{h,\mathsf{p}} \right]^1, \tag{39}$$

where $\mathsf{p} \in \mathbb{N}$ is the polynomial degree of the momentum FE approximation. Additionally the following modified spaces which satisfy Dirichlet boundary data $(\rho\mathbf{u})_D$ and homogeneous boundary data are defined

$$\overline{\mathbf{V}}_{\mathrm{BC}}^{h,\mathsf{p}} := \{ \overline{\mathbf{v}} \in \overline{\mathbf{V}}^{h,\mathsf{p}} : \overline{\mathbf{v}}|_{\partial\Omega_D} = (\rho\mathbf{u})_D \} \quad \text{and} \quad \overline{\mathbf{V}}_{\mathrm{BC}_0}^{h,\mathsf{p}} := \{ \overline{\mathbf{v}} \in \overline{\mathbf{V}}_{\mathrm{CG}}^{h,\mathsf{p}} : \overline{\mathbf{v}}|_{\partial\Omega_D} = \mathbf{0} \}, \tag{40}$$

respectively. On the Neumann boundary we enforce the homogeneous condition $\underline{\underline{\sigma}} \cdot \mathbf{n}|_{\partial\Omega_N} = \mathbf{0}$.

5.2 Compressible Stokes EDG-HDG FE formulation

The compressible Stokes system is composed of the conservation of momentum and mass, equations (2) and (3), respectively. The EDG-HDG formulation of the conservation

of momentum follows from the abstract formulation stated in section 2.4. The EDG-HDG formulation of the continuity equation is also presented (see e.g., Rhebergen & Wells, 2017).

We remark that (by the quotient rule)

$$\mathbf{u}_h = \frac{(\rho\mathbf{u})_h}{\rho} \quad \text{and} \quad \nabla\mathbf{u}_h = \frac{\rho\nabla(\rho\mathbf{u})_h - (\rho\mathbf{u})_h \otimes \nabla\rho}{\rho^2}, \tag{41}$$

which may be used in construction of the stress tensor $\underline{\underline{\sigma}}$ in equation (5). Using the formulation in section 2.4 we see that in the case of isotropic viscosity where $d = 2$, $\mathbf{w}_h = (\rho\mathbf{u})_h$, $\underline{\underline{\tau}} = \underline{\underline{\sigma}}$ and

$$G = \begin{pmatrix} \begin{pmatrix} \frac{2}{\rho}(\eta - \frac{1}{3}) & 0 \\ 0 & -\frac{2}{3\rho} \end{pmatrix} & \begin{pmatrix} 0 & \frac{\eta}{\rho} \\ \frac{\eta}{\rho} & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & \frac{\eta}{\rho} \\ \frac{\eta}{\rho} & 0 \end{pmatrix} & \begin{pmatrix} -\frac{2}{3\rho} & 0 \\ 0 & \frac{2}{\rho}(\eta - \frac{1}{3}) \end{pmatrix} \end{pmatrix}. \tag{42}$$

The residual EDG-HDG formulation of the Stokes system reads: find the momentum and pressure approximations $\left((\rho\mathbf{u})_h, \overline{(\rho\mathbf{u})}_h, p_h, \overline{p}_h\right) \in V_{\mathrm{DG}}^{h,\mathsf{p}} \times \overline{V}_{\mathrm{BC}}^{h,\mathsf{p}} \times Q_{\mathrm{DG}}^{h,\mathsf{p}-1} \times \overline{Q}_{\mathrm{DG}}^{h,\mathsf{p}}$, such that

$$\begin{aligned} F_{\rho\mathbf{u}} :=& \sum_{\kappa\in\mathcal{T}_h} \int_\kappa \underline{\underline{\sigma}} : \nabla\mathbf{v}_h \,\mathrm{d}\mathbf{x} + \sum_{\kappa\in\mathcal{T}_h} \int_{\partial\kappa} (\widehat{\rho\mathbf{u}} - (\rho\mathbf{u})_h) \otimes \mathbf{n}_\kappa : G^\top \nabla\mathbf{v}_h \,\mathrm{d}\mathbf{s} \\ &- \sum_{\kappa\in\mathcal{T}_h} \int_{\partial\kappa} \hat{\underline{\underline{\sigma}}} : \mathbf{v}_h \otimes \mathbf{n}_\kappa \,\mathrm{d}\mathbf{s} + \sum_{\kappa\in\mathcal{T}_h} \int_{\partial\kappa} \hat{\underline{\underline{\sigma}}} : \overline{\mathbf{v}}_h \otimes \mathbf{n}_\kappa \,\mathrm{d}\mathbf{s} \\ &- \sum_{\kappa\in\mathcal{T}_h} \int_\kappa \mathbf{f}(\mathbf{x}, t^{n+1}) \cdot \mathbf{v}_h \,\mathrm{d}\mathbf{x} \equiv 0, \end{aligned} \tag{43}$$

$$F_{\nabla\cdot\rho\mathbf{u}} := \sum_{\kappa\in\mathcal{T}_h} \int_\kappa \nabla \cdot (\rho\mathbf{u})_h q_h \,\mathrm{d}\mathbf{x} + \sum_{\kappa\in\mathcal{T}_h} \int_{\partial\kappa} ((\rho\mathbf{u})_h - \overline{(\rho\mathbf{u})}_h) \cdot \mathbf{n}_\kappa \overline{q}_h \,\mathrm{d}\mathbf{s} \equiv 0, \tag{44}$$

for all $(\mathbf{v}_h, \overline{\mathbf{v}}_h, q_h, \overline{q}_h) \in V_{\mathrm{DG}}^{h,\mathsf{p}} \times \overline{V}_{\mathrm{BC}_0}^{h,\mathsf{p}} \times Q_{\mathrm{DG}}^{h,\mathsf{p}-1} \times \overline{Q}_{\mathrm{DG}}^{h,\mathsf{p}}$. Here the HDG numerical flux functions are given by

$$\widehat{\rho\mathbf{u}} = \overline{(\rho\mathbf{u})}_h, \tag{45}$$

$$\hat{\underline{\underline{\sigma}}} = \underline{\underline{\sigma}}(\mathbf{u}_h, \overline{p}_h) + \frac{C_{\mathrm{IP}_{\rho\mathbf{u}}}}{h_\kappa} G \left(\left(\overline{(\rho\mathbf{u})}_h - (\rho\mathbf{u})_h \right) \otimes \mathbf{n}_\kappa \right), \tag{46}$$

where $C_{\mathrm{IP}_{\rho\mathbf{u}}}$ is an IP parameter selected through experimentation to be $C_{\mathrm{IP}_{\rho\mathbf{u}}} = 12\mathsf{p}^2$.

6 Discussion

With the discretization schemes outlined in algorithm 1 now established, we provide pertinent topics for consideration in this section.

6.1 Time step size and operator splitting error

In our numerical experiments

$$h_{\min} \max_{\mathbf{x}\in\Omega} \left|\mathbf{u}_h(\mathbf{x}, t_n)\right| > k.$$

Therefore we select our time step according to the hyperbolic Courant–Friedrichs–Lewy (CFL) criterion

$$\Delta t_n = C_{\mathrm{CFL}} \frac{h_{\min}}{\max_{\mathbf{x}\in\Omega} \left|\mathbf{u}_h(\mathbf{x}, t_n)\right|}. \tag{47}$$

where C_{CFL} is the desired maximum Courant number bound.

We also consider the numerical error incurred by operator splitting by writing equation (4)

$$\rho \frac{\partial T}{\partial t} + A(T) + B(T) = 0, \tag{48}$$

$$A(T) = \rho \mathbf{u} \cdot \nabla T, \tag{49}$$

$$B(T) = -\nabla \cdot k \nabla T - H \tag{50}$$

and considering the discretization in equations (13) and (14). Analysis (LeVeque, 1999) shows us we incur an error

$$T_d^{n+1} - T^{n+1} = \frac{(\Delta t_n)^2}{2}(A \circ B - B \circ A)(T^n) + \mathcal{O}({\Delta t_n}^3). \tag{51}$$

In essence, if the operators $A(\cdot)$ and $B(\cdot)$ commute, then the splitting scheme is exact. It can be shown (following Lanser & Verwer, 1999) that $A(\cdot)$ and $B(\cdot)$ commute only in the case $\rho\mathbf{u}$ and k are independent of $\mathbf{x}$.

Although the splitting scheme we use is globally first order accurate, in our numerical experiments in section 7 we observe that the splitting scheme error is a concern only when Δt_n is larger than that prescribed by the CFL criterion. The splitting scheme in this work may be improved by employing higher order methods, such as globally second order accurate Strang splitting. However, higher order splitting methods will incur greater computational expense due to the increased number of global FE systems to be solved each time step.

6.2 Superconvergence of the HDG compressibility approximation

Superconvergent properties of the HDG method have been documented in, for example, Giacomini et al. (2018) and Nguyen et al. (2011). Here we examine the superconvergence of the HDG compressibility approximation. Consider the momentum variable $\rho\mathbf{u}$ and its HDG approximation $(\rho\mathbf{u})_h$. We define the $L_2(\mathcal{T}_h)$ norm

$$\|\mathbf{v}\|_{L_2(\mathcal{T}_h)} = \sqrt{\sum_{\kappa \in \mathcal{T}_h} \int_\kappa \mathbf{v}^2 \, \mathrm{d}\mathbf{x}}, \tag{52}$$

such that we have the property

$$\|\nabla \cdot (\rho\mathbf{u})\|_{L_2(\mathcal{T}_h)} = \|\nabla \cdot (\rho\mathbf{u})_h\|_{L_2(\mathcal{T}_h)} = 0. \tag{53}$$

The precise advection of tracers through a velocity field is crucial for stable and accurate simulation. So we examine the accuracy of our approximation of both $\mathbf{u}_h$ and $\nabla \cdot \mathbf{u}_h$.

Following from Rhebergen and Wells (2020, Theorem 3) we have a known, pressure robust, bound on the error of the momentum approximation

$$\|\rho\mathbf{u} - (\rho\mathbf{u})_h\|_{L_2(\mathcal{T}_h)} \leq C_{\rho\mathbf{u}} h^{\mathsf{p}+1} \|\rho\mathbf{u}\|_{H^{\mathsf{p}+1}(\Omega)}, \tag{54}$$

where $C_{\rho\mathbf{u}}$ is a constant independent of the mesh and $\|\rho\mathbf{u}\|_{H^{\mathsf{p}+1}(\Omega)}$ is the $H^{\mathsf{p}+1}$ norm of the momentum. We now consider the compressibility approximation error $\|\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\|_{L_2(\mathcal{T}_h)}$.

First we note by the quotient rule and equation (53)

$$\nabla \cdot \mathbf{u} = \nabla \cdot \left(\frac{\rho\mathbf{u}}{\rho} \right) = \frac{\rho \nabla \cdot (\rho\mathbf{u}) - \rho\mathbf{u} \cdot \nabla\rho}{\rho^2} = -\rho\mathbf{u} \cdot \frac{\nabla\rho}{\rho^2}, \tag{55}$$

where it follows that in terms of the HDG FE approximation

$$\nabla \cdot \mathbf{u}_h = \nabla \cdot \left(\frac{(\rho\mathbf{u})_h}{\rho} \right) = -(\rho\mathbf{u})_h \cdot \frac{\nabla\rho}{\rho^2}. \tag{56}$$

We rewrite the HDG FE compressibility error

$$\begin{aligned}\|\nabla\cdot\mathbf{u}-\nabla\cdot\mathbf{u}_h\|_{L_2(\mathcal{T}_h)} &= \left\|\nabla\cdot\frac{\rho\mathbf{u}}{\rho}-\nabla\cdot\frac{(\rho\mathbf{u})_h}{\rho}\right\|_{L_2(\mathcal{T}_h)},\\ &= \left\|((\rho\mathbf{u})_h-\rho\mathbf{u})\cdot\frac{\nabla\rho}{\rho^2}\right\|_{L_2(\mathcal{T}_h)},\\ &\leq \left\|(\rho\mathbf{u})_h-\rho\mathbf{u}\right\|_{L_2(\mathcal{T}_h)}\left\|\frac{\nabla\rho}{\rho^2}\right\|_{L_2(\mathcal{T}_h)}. \end{aligned} \tag{57}$$

In essence we have that the HDG FE compressibility approximation converges at the same rate as the momentum approximation, $\mathcal{O}(h^{\mathsf{p}+1})$, as in equation (54). The term $\|\nabla\rho/\rho^2\|_{L_2(\mathcal{T}_h)}$ indicates that as the density field's smoothness tends towards a constant we recover an exactly incompressible approximation. Evidently the compressibility approximation error may suffer from highly irregular density fields.

This superconvergent error bound is enabled by the numerical property of equation (53). Standard FE methods for which $\|\nabla\cdot(\rho\mathbf{u})_h\|_{L_2(\mathcal{T}_h)}\neq 0$ or $\|\nabla\cdot(\rho\mathbf{u}_h)\|_{L_2(\mathcal{T}_h)}\neq 0$ will not exhibit this phenomenon.

We demonstrate the HDG method's superconvergent compressibility approximation where $\mathsf{p}=2$ in figure 2. Here we compute HDG momentum formulation and Taylor–Hood (TH) (as defined in Sime et al., 2021) momentum and velocity formulations of the compressible Stokes system with a manufactured solution for comparison. The system is solved in the domain $\Omega=(0.5,1.5)^2$ where the prescribed velocity and density fields are

$$\mathbf{u}(\mathbf{x},t)=\left(\frac{y^2}{x^2+y^2}+1\right)\begin{pmatrix}-y\\ x\end{pmatrix}\quad\text{and}\quad\rho=\frac{\rho_0+\sqrt{x^2+y^2}}{3-\frac{x^2-y^2}{x^2+y^2}}, \tag{58}$$

respectively. We may examine the influence of the $\|\nabla\rho/\rho^2\|_{L_2(\mathcal{T}_h)}$ term by adjusting the constant value $\rho_0\geq 0$. We measure the error by computing the following functionals which we expect to converge at the stated rates in the case of the EDG-HDG discretization scheme:

1. $\|\mathbf{u}-\mathbf{u}_h\|_{L_2(\mathcal{T}_h)}\backsim\mathcal{O}(h^{\mathsf{p}+1})$: velocity error,
2. $\|\rho\mathbf{u}-(\rho\mathbf{u})_h\|_{L_2(\mathcal{T}_h)}\backsim\mathcal{O}(h^{\mathsf{p}+1})$: momentum error,
3. $\|\nabla\cdot\mathbf{u}-\nabla\cdot\mathbf{u}_h\|_{L_2(\mathcal{T}_h)}\backsim\mathcal{O}(h^{\mathsf{p}+1})$: *superconvergent* compressibility error,
4. $\|\nabla\cdot(\rho\mathbf{u})_h\|_{L_2(\mathcal{T}_h)}=0$: mass continuity error,
5. $\|\nabla\mathbf{u}-\nabla\mathbf{u}_h\|_{L_2(\mathcal{T}_h)}\backsim\mathcal{O}(h^{\mathsf{p}})$: velocity gradient error,
6. $\|\nabla(\rho\mathbf{u})-\nabla(\rho\mathbf{u})_h\|_{L_2(\mathcal{T}_h)}\backsim\mathcal{O}(h^{\mathsf{p}})$: momentum gradient error.

6.3 Choice of intermediate states

In this section we specify how we choose the intermediate states of the advective and diffusive temperature updates, $T^{*,n}_{a,h}$ and $T^{*,n}_{d,h}$, respectively. The diffusive state is simply chosen as the most recent advective update $T^{*,n}_{d,h}=T^{n+1}_{a,h}$. However, more care must be taken regarding the choice of the advective update to maintain mathematical consistency. The following derivation was originally developed in Maljaars et al. (2019) for advection-diffusion in incompressible flows, which we now frame in the context of the advection-diffusion splitting of equation (4).

6.3.1 Tracer updates

Consider the abstract HDG formulation equation (26) in the context of equation (4). In section 4 we stated the HDG FE formulation employing a fully implicit backward Euler method. Here we write the formulation employing a full θ scheme, and show why we choose the backward Euler method and retain second order accuracy.

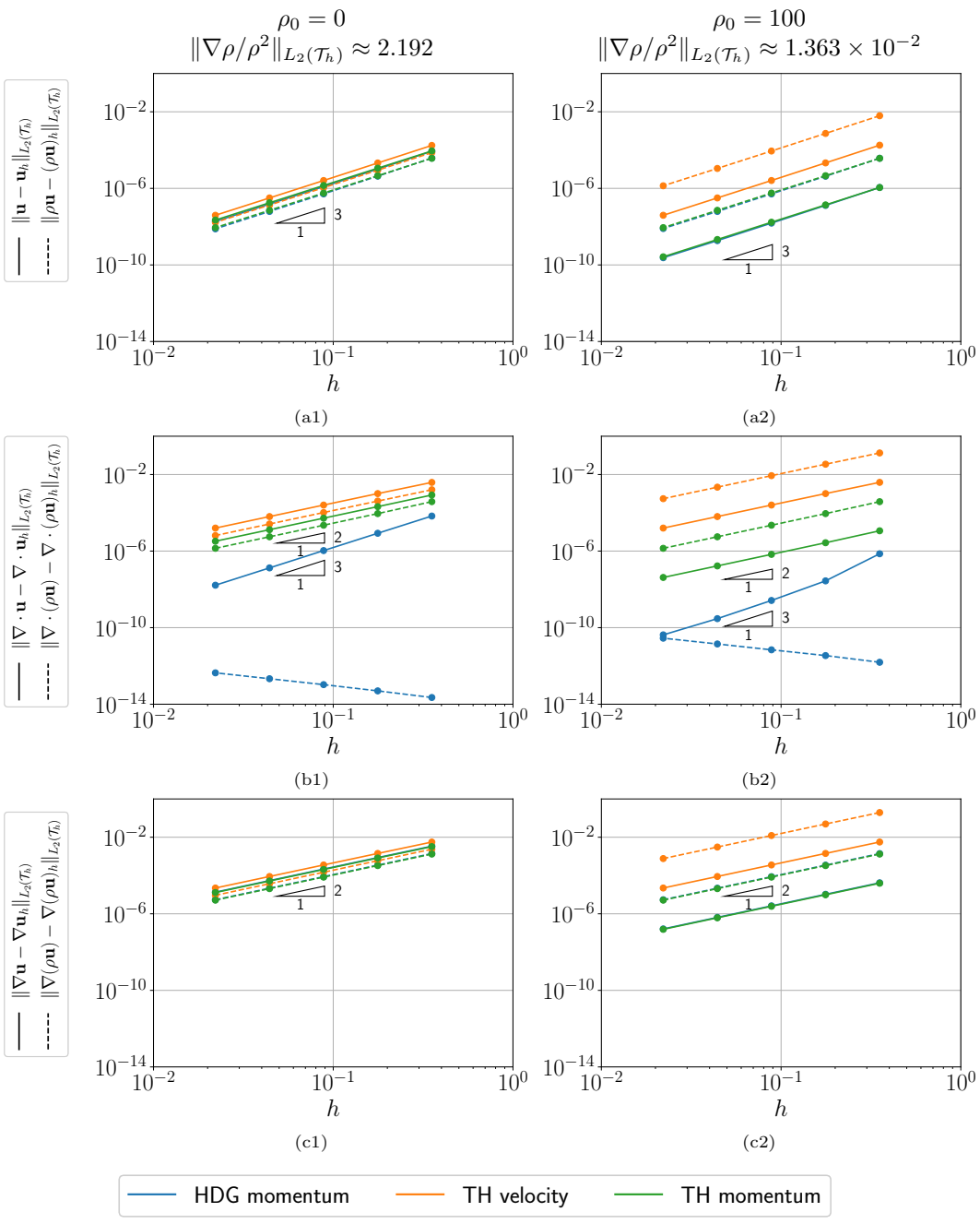


Figure 2: Comparison of the p=2 HDG momentum formulation with the p=2 TH momentum and velocity formulations using the manufactured solution in equation (58). The left and right columns show the 'weakly' ρ_0=0 and 'strongly' ρ_0=100 compressible cases, respectively (see section 6.2). In (a1) and (a2) we observe optimal convergence of both the momentum and velocity formulations. In (b1) and (b2) we observe the HDG method's superconvergent rate of the compressibility approximation error due to pointwise satisfaction of the continuity constraint (see section 6.2). The growth in the error of the HDG method's satisfaction of the continuity constraint is due to accumulation of floating point error. In (c1) and (c2) we observe the standard optimal rates of convergence of the velocity and momentum approximations' gradients. The difference between the HDG and TH momentum formulation error in plots (a1), (a2), (c1) and (c2) is indistinguishable at this scale.

Let us design the θ scheme with $\theta_L \in [0,1]$ which is separate and possibly different from θ in equation (11) such that

$$T_{d,h}^{n+1} = \theta_L T_{I,h}^{n+1} + (1-\theta_L) T_{E,h}^{n+1}. \tag{59}$$

Here $T_{I,h}^{n+1}$ and $T_{E,h}^{n+1}$ are the implicit and explicit components, respectively. We write the abstract FE formulation of the diffusion problem

$$\int_\Omega \frac{T_{d,h}^{n+1} - T_{d,h}^{*,n}}{\Delta t_n} s_h \,\mathrm{d}\mathbf{x} + \theta_L F^{n+1} + (1-\theta_L) F^n = 0, \tag{60}$$

where $F^{n+1}{=}F(T_{I,h}^{n+1}, s_h)$ and $F^n{=}F(T_{E,h}^{n+1}, s_h)$ are the HDG FE discretizations shown in section 2.4 and equation (26). Therefore

$$\begin{aligned}
\int_\Omega \frac{T_{d,h}^{n+1} - T_{d,h}^{*,n}}{\Delta t_n} s_h \,\mathrm{d}\mathbf{x} &= \theta_L \int_\Omega \frac{T_{I,h}^{n+1} - T_{d,h}^{*,n}}{\Delta t_n} s_h \,\mathrm{d}\mathbf{x} + (1-\theta_L) \int_\Omega \frac{T_{E,h}^{n+1} - T_{d,h}^{*,n}}{\Delta t_n} s_h \,\mathrm{d}\mathbf{x}, \\
&= \theta_L \int_\Omega \dot{T}_{I,h}^{n+1} s_h \,\mathrm{d}\mathbf{x} + (1-\theta_L) \int_\Omega \dot{T}_{E,h}^{n+1} s_h \,\mathrm{d}\mathbf{x}, \\
&= -(\theta_L F^{n+1} + (1-\theta_L) F^n),
\end{aligned} \tag{61}$$

where we have employed the notation $\dot{T}_{\cdot,h}^{n+1}{=}(T_{\cdot,h}^{n+1} - T_{d,h}^{*,n})/\Delta t_n$. We can see that

$$\int_\Omega \dot{T}_{I,h}^{n+1} s_h \,\mathrm{d}\mathbf{x} = -F^{n+1} \quad \text{and} \quad \int_\Omega \dot{T}_{E,h}^{n+1} s_h \,\mathrm{d}\mathbf{x} = -F^n, \tag{62}$$

which implies that $\dot{T}_{E,h}^{n+1}{=}\dot{T}_{I,h}^{n}$.

The update of the tracer data requires

$$\begin{aligned}
\dot{T}_p^{n+1} &= \dot{T}_{d,h}^{n+1}(\mathbf{x}_p^{n+1}), \\
&= \theta_L \dot{T}_{I,h}^{n+1}(\mathbf{x}_p^{n+1}) + (1-\theta_L) \dot{T}_{E,h}^{n+1}(\mathbf{x}_p^{n}), \\
&= \theta_L \dot{T}_{I,h}^{n+1}(\mathbf{x}_p^{n+1}) + (1-\theta_L) \dot{T}_{I,h}^{n}(\mathbf{x}_p^{n}), \qquad p = 1, \ldots, N_p,
\end{aligned} \tag{63}$$

which rearranged in terms of the unknown tracer update yields

$$T_p^{n+1} = T_p^n + \Delta t_n \left(\theta_L \dot{T}_{I,h}^{n+1}(\mathbf{x}_p^{n+1}) + (1-\theta_L) \dot{T}_{I,h}^{n}(\mathbf{x}_p^{n}) \right), \quad p = 1, \ldots, N_p. \tag{64}$$

We exploit this relationship by noting $T_{d,h}^n{=}T_{I,h}^n$ in our backward Euler formulation in equation (36). Provided we make the choice $\theta_L{=}\frac{1}{2}$ we obtain a second order accurate time discretization. We then only need to store $\dot{T}_{d,h}^{n+1}$ and $\dot{T}_{d,h}^{n}$ between time steps in order to update tracer data by:

$$T_p^{n+1} = T_p^n + \Delta t_n \left(\theta_L \dot{T}_{d,h}^{n+1}(\mathbf{x}_p^{n+1}) + (1-\theta_L) \dot{T}_{d,h}^{n}(\mathbf{x}_p^{n}) \right), \quad p = 1, \ldots, N_p. \tag{65}$$

6.3.2 Deriving a mathematically consistent advective initial state

We may now define the intermediate state of the advective update, $T_{a,h}^{*,n}$. We do so by enforcing mathematical consistency of the projection operators between the tracers and the mesh. By this we mean the following: provided tracers do not move, the composition of projection from field to tracers and back to field (and vice versa) yields the identity. Therefore we define the field relation consistent with equation (63)

$$\dot{T}_{a,h}^{*,n} = \frac{T_{a,h}^{*,n} - T_{a,h}^{n}}{\Delta t_n} = \theta_L \dot{T}_{d,h}^{n} + (1-\theta_L) \dot{T}_{d,h}^{n-1}, \tag{66}$$

which rearranged in terms of the intermediate state yields

$$T_{a,h}^{*,n} = T_{a,h}^{n} + \Delta t_n \left(\theta_L \dot{T}_{d,h}^{n} + (1-\theta_L) \dot{T}_{d,h}^{n-1} \right). \tag{67}$$

7 Numerical experiments

In this section we present results from numerical experiments designed to verify the implementation of tracer projection and diffusion in the context of compressible geodynamics. These experiments comprise a system with a manufactured solution where we exactly quantify approximation error and reproductions of benchmarks established in the literature.

In all cases the tracer configuration is generated with $15\mathsf{s}$ tracers per cell whose initial positions are drawn from a uniform random distribution defined on the cells' geometries. Furthermore we maintain a minimum and maximum number of tracers per cell, $(N_{p_\kappa}^{\min}, N_{p_\kappa}^{\max}) = (15\mathsf{s}, 20\mathsf{s})$, respectively, using the method described in section 2.6.

The code used to generate all results presented in this section is available in the public repository Sime (2021). In our computational implementation we use the Lagrangian–Eulerian on Particles (LEoPart) library (Maljaars et al., 2021) for RK numerical integration, PDE-constrained l_2 projection and assembly of HDG discretizations with static condensation. We use this in conjunction with the components of the FEniCS project (Alnæs et al., 2015) facilitating the computation of FE solutions. The data structures and direct solver used to manage the underlying linear systems are provided by the Portable Extensible Toolkit for Scientific Computation (PETSc) (Balay et al., 2019b, 2019a) and the Multifrontal Massively Parallel Sparse Direct Solver (MUMPS) (Amestoy et al., 2000), respectively.

7.1 Coupled manufactured solution

Our implementation of advection and projection of tracers in a background velocity field using the RKℓ method and PDE-constrained l_2 projection have been verified in Sime et al. (2021). In this section we take inspiration from the coupled manufactured solution experiment presented therein.

7.1.1 Problem description

Let the computational domain $\Omega := \{(x, y) : r_0 < \sqrt{x^2 + y^2} < r_1\}$ be the annulus of inner and outer radii $r_0 = 0.4292$ and $r_1 = r_0 + 1$, respectively. Furthermore let $\mathcal{I}_t = [0, 0.5]$ be the time domain. We prescribe analytic velocity and temperature fields

$$\mathbf{u}(\mathbf{x}, t) = \left(\frac{y^2}{x^2 + y^2} + 1\right)\left(\cos^2 t + \frac{1}{2}\right)\begin{pmatrix}-y\\ x\end{pmatrix}, \tag{68}$$

$$T(\mathbf{x}, t) = \frac{2}{2 + 4kt}\exp\left(-\frac{x^2 + y^2}{2 + 4kt}\right), \tag{69}$$

and use these to set Dirichlet boundary conditions on the discrete momentum and temperature solutions in addition to an initial condition on temperature. We further prescribe the analytic density and pressure fields

$$\rho(\mathbf{x}) = \frac{\sqrt{x^2 + y^2}}{3 - \frac{x^2 - y^2}{x^2 + y^2}}, \tag{70}$$

$$p = 0. \tag{71}$$

The following residual formulations then complete the system of equations for the manufactured solution

$$H = 8k(1 - \rho)\frac{(2 + 4kt - (x^2 + y^2))}{(2 + 4kt)^3}\exp\left(-\frac{x^2 + y^2}{2 + 4kt}\right), \tag{72}$$

$$\mathbf{f} = \mathsf{r} - T_{d,h}\hat{\mathbf{g}}, \tag{73}$$

where the r is the residual function composed of the true temperature solution

$$\mathsf{r} = \frac{4}{3(x^2 + y^2)^2}\left(\cos^2 t + \frac{1}{2}\right)\begin{pmatrix}5x^2y - 2y^3\\ 5xy^2 - 2x^3\end{pmatrix} + T\hat{\mathbf{g}}. \tag{74}$$

The following functionals are employed to measure the error:

1. $\|\mathbf{u} - \mathbf{u}_h\|_{L_2(\mathcal{T}_h)}$: velocity approximation error,
2. $\|\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\|_{L_2(\mathcal{T}_h)}$: compressibility approximation error,
3. $\|\rho\mathbf{u} - (\rho\mathbf{u})_h\|_{L_2(\mathcal{T}_h)}$: momentum approximation error,
4. $\|\nabla \cdot (\rho\mathbf{u})_h\|_{L_2(\mathcal{T}_h)}$: mass continuity error,
5. $\|T - T_h\|_{L_2(\mathcal{T}_h)}$: temperature approximation error,
6. $\varepsilon\Delta T = \left|\int_\Omega \left(\rho T_h(\mathbf{x}, t) - \rho T_h(\mathbf{x}, 0)\right) \mathrm{d}\mathbf{x}\right| / \int_\Omega \rho T_h(\mathbf{x}, 0)\, \mathrm{d}\mathbf{x}$: conservation.

7.1.2 Results

For three cases where $k \in \{0, 10^{-3}, 10^{-1}\}$ we show the error functionals in the time domain in figure 3. Additionally we show the error measured at time $t = t_F$ demonstrating convergence rates in figure 4. We emphasize the following observations in support of the numerical method development:

1. We expect a second order accurate method in time. This is observed by examining the temperature approximation error in the case p=2, s=2, ℓ=3. In the case k=0 we see third order accuracy in space and time. However increasing to $k \in \{10^{-3}, 10^{-1}\}$ the rate of convergence deteriorates to second order (see figure 4(c1)).
2. The compressibility approximation $\|\nabla \cdot \mathbf{u} - \nabla \cdot \mathbf{u}_h\|_{L_2(\mathcal{T}_h)}$ converges at the super optimal rate of $\mathsf{p} + 1$ (see figure 4(b2)).
3. The HDG compressible Stokes discretization scheme yields an exactly pointwise divergence free momentum approximation to machine precision (see figure 3(a2) and figure 4(a2)).
4. In the case that $k = 0$ we achieve exact conservation of the temperature field (see figure 3(c2) and figure 4(c2)). When k>0 flux through the boundaries prescribed in the analytic solution prevents exact conservation.

7.2 Benchmark reproduction

The previous numerical example demonstrated and verified the implementation of the discretized advection diffusion operator on tracers in the compressible scheme. In this section we draw inspiration from King et al. (2010) by conducting a time-dependent evolution simulation of their ALA cases for two different choices of parameters.

7.2.1 Problem description

Let $\Omega := (0, 1)^2$ be the unit square. We seek the solution of the compressible system where

$$\mathbf{f} = \rho \left(\mathrm{Di}\, p - \mathrm{Ra}\, (T - T_{\mathrm{ref}})\right) \hat{\mathbf{g}}, \tag{75}$$

$$H = -\mathrm{Di}\, \rho w \alpha \left(T + \frac{T_{\mathrm{surf}}}{\Delta T}\right) + \frac{\mathrm{Di}}{\mathrm{Ra}} \underline{\underline{\sigma}} : \nabla \mathbf{u}, \tag{76}$$

where the reference density and temperature are

$$\rho = e^{\mathrm{Di}\, z/\gamma_r} \quad \text{and} \quad T_{\mathrm{ref}} = \frac{T_{\mathrm{surf}}}{\Delta T} e^{\mathrm{Di}\, z}, \tag{77}$$

respectively. Here z=1−y is the depth, $\hat{\mathbf{g}}$ is the non-dimensional gravity (unit vector in the direction of increasing depth), w=−$\mathbf{u}$·$\hat{\mathbf{g}}$ is the vertical velocity component, α=1 is the thermal expansivity, γ_r=1 is the reference Grüneisen parameter, T_{surf}=273 is the surface temperature and ΔT=3000 is the dimensional temperature difference over the domain. Furthermore Ra and Di are the dimensionless Rayleigh number and dissipation number, respectively.

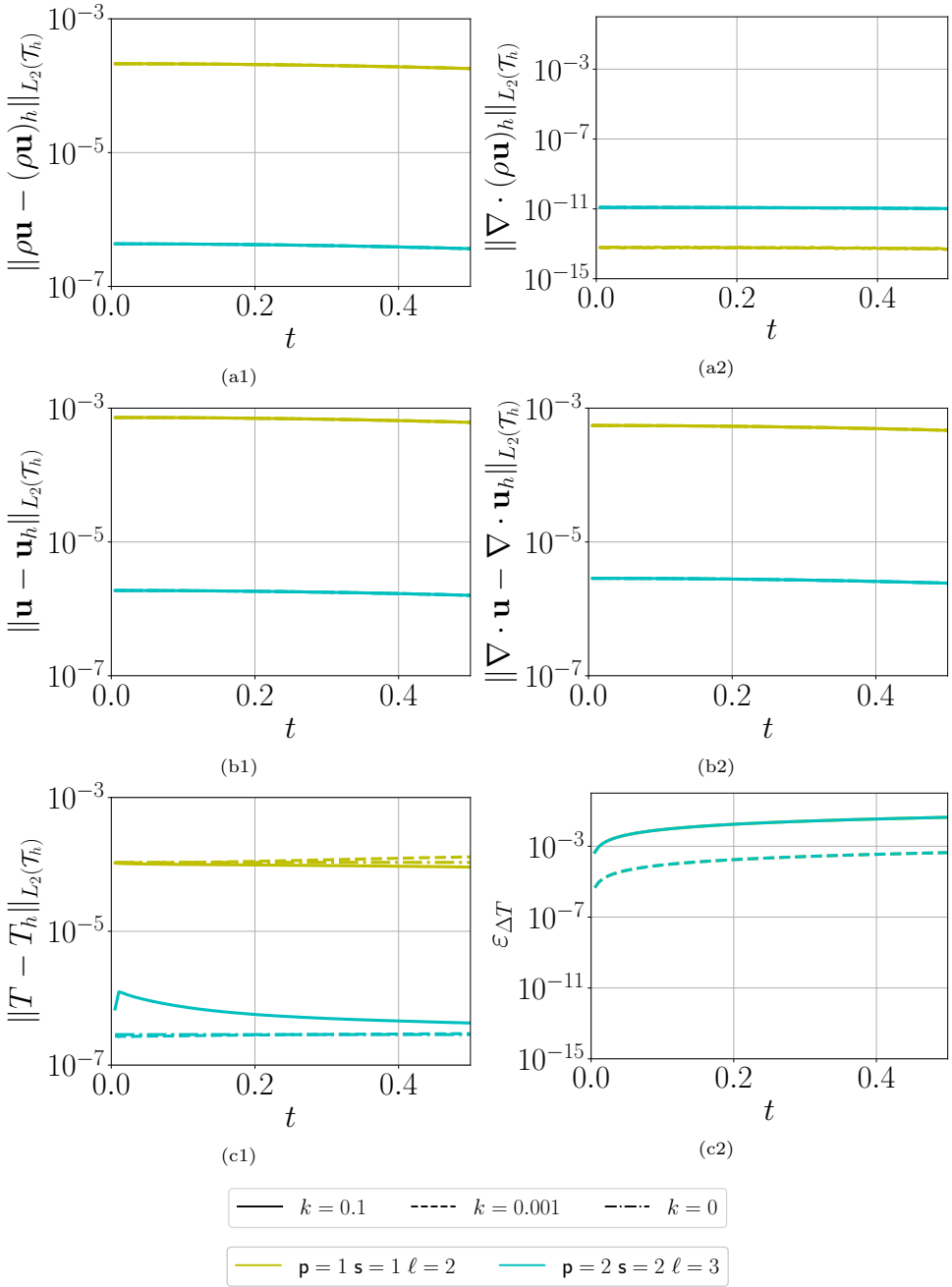


Figure 3: Computed error at time t computed from the coupled manufactured solution experiment described in section 7.1. Here the smooth temperature function solution has very little impact on the momentum and velocity approximation errors. Exact conservation of the temperature field was found in the case of k=0 (so cannot be shown on these graphs). Mass continuity is exactly satisfied to machine precision and the apparent increase in the solenoidal momentum approximation error is due to the accumulation of machine error.

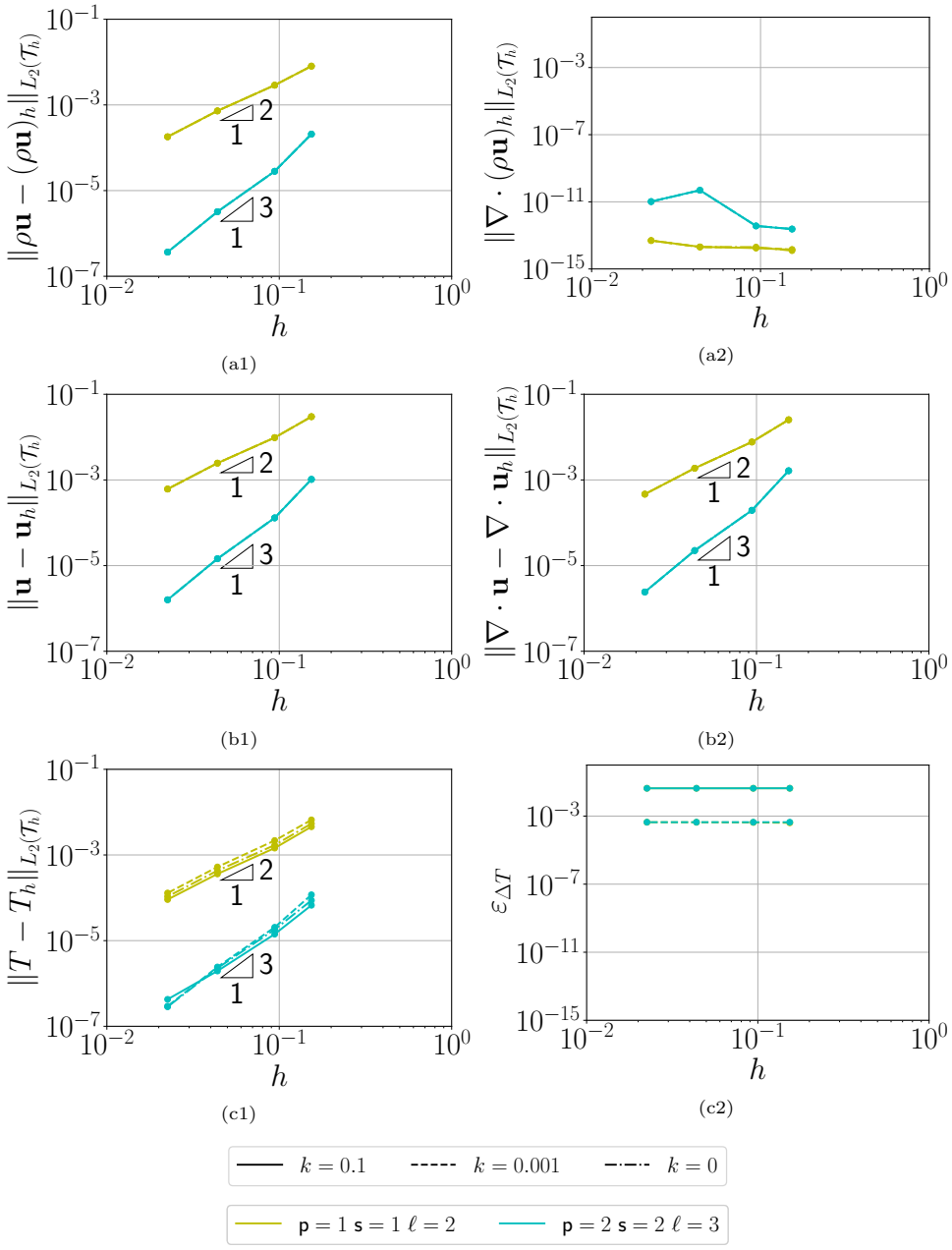


Figure 4: Computed error at time t=t_F from the coupled manufactured solution experiment described in section 7.1. Here we demonstrate optimal convergence rates of the PDE-constrained projection scheme. The difference between the momentum and velocity approximation errors at the specified values of k are indistinguishable at this scale. We highlight that the temperature approximation convergence rate tends towards second order as the value of k grows. Furthermore, in the case of k=0 the method exactly conserves energy (so cannot be shown on these graphs). Additionally we see optimal convergence rates of the velocity field approximation, 'super convergent' rates of the compressibility field approximation and that the mass continuity equation is exactly satisfied to machine precision.

The system is initialized with prescribed temperature field

$$T(\mathbf{x}, t=0) = z + \frac{1}{10}\cos(\pi x)\sin(\pi y) \tag{78}$$

and the boundary conditions are:

1. The temperature $T|_{y=0}$=1 on the bottom and $T|_{y=1}$=0 top boundaries. Furthermore $k\nabla T \cdot \mathbf{n}$=0 on the left and right boundaries (x=0 and x=1, respectively).
2. The momentum is prescribed such that $\mathbf{u}\cdot\mathbf{n}$=0 and $(\sigma\cdot\mathbf{n})\cdot\mathbf{t}$=0 on $\partial\Omega$, where $\mathbf{t}$ is a unit vector lying tangential to the boundary.

The mesh consists of 64×64 quadrilaterals bisected into triangles. We employ local refinement towards the top and bottom of the mesh so as to better resolve the thermal boundary layers. On this locally refined mesh it is crucial we employ the tracer addition and removal scheme described in section 2.6 for computational performance.

7.2.2 Time evolution of functionals

In our first experiment we examine the impact of the choice of C_{CFL} on the accuracy of the numerical approximation. We consider two cases:

Case 1: $\mathrm{Ra} = 10^4$ and $\mathrm{Di} = 0.25$ such that $\|\nabla\rho/\rho^2\|_{L_2(\mathcal{T}_h)} \approx 0.222$,
Case 2: $\mathrm{Ra} = 10^5$ and $\mathrm{Di} = 1.0$ such that $\|\nabla\rho/\rho^2\|_{L_2(\mathcal{T}_h)} \approx 0.658$.

By considering the value of the dissipation number, Di, cases 1 and 2 can be seen as having 'mild' and 'strong' compressibilities, respectively. The intent is to observe the impact of this degree of compressibility in terms of the noncommutativity of the advection and diffusion operators (see section 6.1).

In figure 5 we show the computed values of Nu as compared with the reference values taken from King et al. (2010). We compare the discretization scheme with $(\mathsf{p},\mathsf{s},\ell) \in \{(1,1,2),(2,2,3)\}$. In essence second and third order methods in space and second and *at best* third order methods in time.

We draw our attention first to the mildly compressible Case 1 (Ra=10^4 and Di=0.25) in figure 5. The computed values of Nu at the steady state (t>0.24) show agreement with the University of Michigan (UM) value of King et al. (2010). Furthermore in this case of 'mild' compressibility, by comparing the convergence of low and high order approximations, we can see that we are able to resolve the transient process from the initial condition to the steady state well even when using C_{CFL}>1. Clearly the choice of the RK method order ℓ has significant influence on both the transient and steady state approximation error (see section 6.1).

Now consider the strongly compressible Case 2 (Ra=10^5 and Di=1) in figure 5. Here we see that the choice of C_{CFL} and ℓ is of crucial importance in accurate transient and steady state approximation. Choosing $C_{\mathrm{CFL}} = 8$ is clearly inadequate as the steady state approximation does not agree with the UM value and the transient evolution shows little correspondence with the higher order approximations. In the case C_{CFL}=4, although we achieve a steady state value of Nu which roughly approximates the UM value, the transient behavior is not consistent with the higher order cases. Finally, with the choices C_{CFL}=2 and C_{CFL}=1 we achieve transient behavior which shows convergence and agrees well with the UM value after reaching the steady state.

8 Summary

In this work we have derived and demonstrated the following:

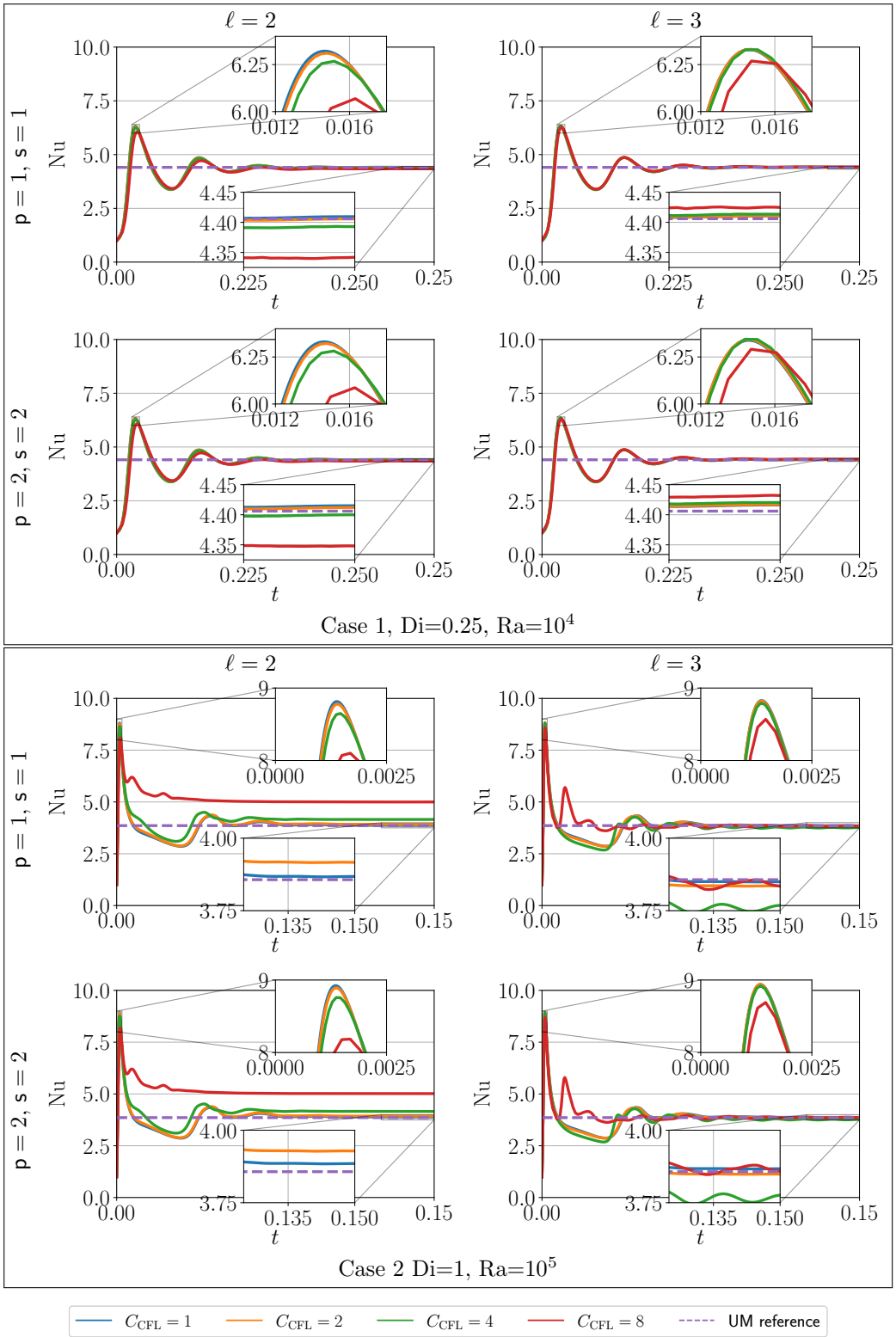


Figure 5: Evaluation of Nu from the numerical benchmark cases 1 and 2 described in section 7.2. The left and right columns correspond to RK2 and RK3 methods, respectively. The top and bottom rows in each case correspond to second (p=1, s=1) and third (p=2, s=2) order HDG spatial discretizations, respectively. Here we emphasize the effect of the 'degree of compressibility' incurred by larger values of Di and Ra. Should one wish to exceed the CFL limit choosing C_{CFL}>1 a corresponding increase in the order of the RK time integration may be necessary.

1. A generalization of IP HDG FE formulations of second order diffusive PDEs (section 2.4).
2. Addition and removal of tracers using the PDE-constrained l_2 projection scheme does not affect mass conservation (section 2.6).
3. An HDG method by which the projection of tracer data to a field exactly conserves mass of a compressible fluid as required by the continuity equation (section 3).
4. An HDG FE formulation of the weakly compressible Stokes system (section 5) which exhibits a superconvergent rate of the approximation of compressibility (section 6.2).
5. The complete tracer-HDG numerical discretization scheme of the coupled compressible Stokes / advection-diffusion system provides approximate solutions which converge at optimal rates in space and second order in time (section 7.1).
6. Tracer methods may allow us to robustly use time step sizes larger than required by the CFL criterion, however, care must be taken in the case of 'strongly compressible' problems (section 7.2).

Acronyms

ALA anelastic liquid approximation
CFL Courant–Friedrichs–Lewy
DG discontinuous Galerkin
EDG embedded discontinuous Galerkin
FE finite element
HDG hybrid discontinuous Galerkin
IP interior penalty
PDE partial differential equation
RK Runge–Kutta
TH Taylor–Hood
UFL unified form language

Acknowledgments

This research was partly supported by NSF-EAR-CSEDI grant 1664642. NS gratefully acknowledges the support of Carnegie Institution for Science President's Fellowship. The authors wish to thank Jakob Maljaars for helpful discussions, in particular regarding the choice of appropriate intermediate state values. No data has been generated or collected as part of this work. The code developed to run all the numerical experiments in this work and a list of its dependencies is available in the public repository Sime (2021).

References

Allègre, C. J., & Turcotte, D. L. (1986). Implications of a two-component marble-cake mantle. *Nature*, *323*, 123–127.

Alnæs, M. S., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., ... Wells, G. N. (2015). The FEniCS Project Version 1.5. *Archive of Numerical Software*, *3*(100), 9–23.

Alnæs, M. S., Logg, A., Ølgaard, K. B., Rognes, M. E., & Wells, G. N. (2014). Unified Form Language: A domain-specific language for weak formulations of partial differential equations. *ACM Transactions on Mathematical Software*, *40*, 9:1–9:37.

Amestoy, P. R., Duff, I. S., & L'Excellent, J.-Y. (2000). Multifrontal parallel distributed symmetric and unsymmetric solvers. *Computer Methods in Applied Mechanics and Engineering*, *184*, 501–520.

Balay, S., Abhyankar, S., Adams, M. F., Brown, J., Brune, P., Buschelman, K., ... Zhang,

H. (2019a). *PETSc users manual* (Tech. Rep. No. ANL-95/11 - Revision 3.12). Argonne National Laboratory. `https://www.mcs.anl.gov/petsc`.

Balay, S., Abhyankar, S., Adams, M. F., Brown, J., Brune, P., Buschelman, K., ... Zhang, H. (2019b). *PETSc Web page.* `https://www.mcs.anl.gov/petsc`.

Ballmer, M. D., Houser, C., Hernlund, J. W., Wentzcovitch, R. M., & Hirose, K. (2017). Persistence of strong silicate-enriched domains in the Earth's lower mantle. *Nature Geoscience*, 236–240.

Ballmer, M. D., Schumacher, L., Lekic, V., Thomas, C., & Ito, G. (2016). Compositional layering within the large low shear-wave velocity provinces in the lower mantle. *Geochemistry, Geophysics, Geosystems*, *17*, 5056–5077.

Blankenbach, B., Busse, F., Christensen, U., Cserepes, L., Gunkel, D., Hansen, U., ... Schnaubelt, T. (1989). A benchmark comparison for mantle convection codes. *Geophysical Journal International*, *98*, 23–38.

Bossman, A., & van Keken, P. E. (2013). Dynamics of plumes in a compressible mantle with phase changes: Implications for phase change boundary topography. *Physics of the Earth and Planetary Interiors*, *224*, 21–31.

Boyet, M., & Carlson, R. W. (2005). ^{142}Nd evidence for early (>4.53 Ga) global differentiation of the silicate Earth. *Science*, *309*, 576–581.

Brandenburg, J. P., Hauri, E. H., van Keken, P. E., & Ballentine, C. J. (2008). A multiple-system study of the geochemical evolution of the mantle with force-balanced plates and thermochemical effects. *Earth and Planetary Science Letters*, *276*, 1–13.

Busse, F. H., Christensen, U., Clever, R., Cserepes, L., Gable, C., Giannandrea, E., ... Travis, B. (1994). 3D convection at infinite Prandtl number in Cartesian geometry – a benchmark comparison. *Geophysical and Astrophysical Fluid Dynamics*, *75*, 39–59.

Christensen, U. R., & Hofmann, A. W. (1994). Segregation of subducted oceanic crust in the convecting mantle. *Journal of Geophysical Research*, *99*, 19867–19884.

Christensen, U. R., & Yuen, D. A. (1984). The interaction of a subducting lithospheric slab with a chemical or phase boundary. *Journal of Geophysical Research*, 4389–4402.

Cockburn, B., Kanschat, G., & Schötzau, D. (2007). A note on discontinuous Galerkin divergence-free solutions of the Navier–Stokes equations. *Journal of Scientific Computing*, *31*, 61–73.

Cockburn, B., Nguyen, N. C., & Peraire, J. (2010). A comparison of HDG methods for Stokes flow. *Journal of Scientific Computing*, *45*(1-3), 215–237.

Curbelo, J., Duarte, L., Alboussière, T., Dubuffet, F., Labrosse, S., & Ricard, Y. (2019). Numerical solutions of compressible convection with an infinite Prandtl number: comparison of the anelastic and anelastic liquid models with the exact equations. *Journal of Fluid Mechanics*, *873*, 646–687.

Elkins-Tanton, L. T. (2008). Linked magma ocean solidification and atmospheric growth for Earth and Mars. *Earth and Planetary Science Letters*, *271*, 181–191.

Evans, J. A., & Hughes, T. J. R. (2013). Isogeometric divergence-conforming B-splines for the steady Navier–Stokes equations. *Mathematical Models and Methods in Applied Sciences*, *23*, 1421–1478.

Garnero, E. J., & McNamara, A. K. (2008). Structure and dynamics of Earth's lower mantle. *Science*, *320*, 626–628.

Gassmöller, R., Dannberg, J., Bangerth, W., Heister, T., & Myhill, R. (2020). On formulations of compressible mantle convection. *Geophysical Journal International*, *221*, 1264–1280.

Gerya, T. V., & Yuen, D. A. (2003). Characteristics-based marker-in-cell method with conservative finite-differences schemes for modeling geological flows with strongly variable transport properties. *Physics of the Earth and Planetary Interiors*, *140*, 293–318.

Giacomini, M., Karkoulias, A., Sevilla, R., & Huerta, A. (2018). A superconvergent HDG method for Stokes flow with strongly enforced symmetry of the stress tensor. *Journal of Scientific Computing*, *77*, 1679–1702.

Glatzmaier, G. A. (1986). Numerical simulation of mantle convection: Time-dependent, three-dimensional, compressible, spherical shell. *Geophysical and Astrophysical Fluid Dynamics*, *43*, 223–264.

Guyan, R. J. (1965). Reduction of stiffness and mass matrices. *AIAA Journal*, *3*(2), 380–380.

Guzmán, J., & Neilan, M. (2014). Conforming and divergence-free Stokes elements on general triangular meshes. *Mathematics of Computation*, *83*, 15–36.

Hansen, U., & Yuen, D. (2000). Extended-Boussinesq thermal-chemical convection with moving heat sources and variable viscosity. *Earth and Planetary Science Letters*, *176*, 400–411.

Hartmann, R., & Houston, P. (2006). Symmetric interior penalty DG methods for the compressible Navier–Stokes equations I: Method formulation. *International Journal of Numerical Analysis and Modeling*, *3*, 1-20.

He, Y., Puckett, E. G., & Billen, M. I. (2017). A discontinuous Galerkin method with a bound preserving limiter for the advection of non-diffusive fields in solid Earth geodynamics. *Physics of the Earth and Planetary Interiors*, *263*, 23–37.

Hillebrand, B., Theuilot, C., Geenen, T., van den Berg, A., & Spakman, W. (2014). Using the level set method in geodynamical modeling of multi-material flows and Earth's free surface. *Solid Earth*, *5*, 1087–1098.

Hofmann, A. W. (2014). Sampling mantle heterogeneity through oceanic basalts: isotopes and trace elements. In H. Holland & K. Turekian (Eds.), *Treatise on Geochemistry (2nd edition)* (Vols. 3 "The Mantle and Core", R. Carlson (Ed.), pp. 67–101). Amsterdam, The Netherlands: Elsevier.

Horan, M. F., Carlson, R. W., Walker, R. J., Jackson, M., Garçon, M., & Norman, M. (2018). Tracking Hadean processes in modern basalts with 142-Neodymium. *Earth and Planetary Science Letters*, *484*, 184–191.

Houston, P., & Sime, N. (2018). Automatic symbolic computation for discontinuous Galerkin finite element methods. *SIAM Journal on Scientific Computing*, *40*, C327-C357.

Jarvis, G. T., & McKenzie, D. P. (1980). Convection in a compressible fluid with infinite Prandtl number. *Journal of Fluid Mechanics*, *96*, 515–583.

Jenny, P., Pope, S., Muradoglu, M., & Caughey, D. (2001). A hybrid algorithm for the joint PDF equation of turbulent reactive flows. *Journal of Computional Physics*, *166*(2), 218–252.

Jones, T. D., Maguire, R. R., van Keken, P. E., Ritsema, J., & Koelemeijer, P. (2020). Subducted oceanic crust as the origin of seismically slow lower-mantle structures. *Progress in Earth and Planetary Science*, *7*, 17.

Jones, T. D., Sime, N., & van Keken, P. E. (2021). Burying Earth's primitive mantle in the slab graveyard. *Geochemistry, Geophysics, Geosystems*, *22*, e2020GC009396.

Kellogg, L. H., Hager, B. H., & van der Hilst, R. D. (1999). Compositional stratification in the deep mantle. *Science*, *283*, 1881–1884.

Kellogg, L. H., & King, S. D. (1993). Effect of mantle plumes on the growth of D" by reaction between the core and mantle. *Geophysical Research Letters*, *20*, 379–382.

King, S. D., Lee, C., van Keken, P. E., Leng, W., Zhong, S., Tan, E., ... Kameyama, M. C. (2010). A community benchmark for 2-D Cartesian compressible convection in the Earth's mantle. *Geophysical Journal International*, *180*, 73-87.

Kronbichler, M., Heister, T., & Bangerth, W. (2012). High accuracy mantle convection simulation through modern numerical methods. *Geophysical Journal International*, 12–29.

Labeur, R. J., & Wells, G. N. (2012). Energy stable and momentum conserving hybrid finite element method for the incompressible Navier–Stokes equations. *SIAM Journal on Scientific Computing*, *34*, A889-A913.

Labrosse, S., Hernlund, J., & Coltice, N. (2007). A crystallizing dense magma ocean at the base of the Earth's mantle. *Nature*, *450*, 866–869.

Lanser, D., & Verwer, J. G. (1999). Analysis of operator splitting for

advection–diffusion–reaction problems from air pollution modelling. *Journal of Computational and Applied Mathematics*, *111*, 201–216.

La Spina, A., Kronbichler, M., Giacomini, M., Wall, W. A., & Huerta, A. (2020). A weakly compressible hybridizable discontinuous Galerkin formulation for fluid–structure interaction problems. *Computer Methods in Applied Mechanics and Engineering*, *372*, 113392.

LeVeque, R. J. (1999). *Numerical methods for conservation laws* (2nd ed.). Boston: Birkhäuser Verlag.

Li, M., McNamara, A., & Garnero, E. (2014). Chemical complexity of hotspots caused by cycling oceanic crust through mantle reservoirs. *Nature Geoscience*, *7*, 366–370.

Li, M., & McNamara, A. K. (2018). The influence of deep mantle compositional heterogeneity on Earth's thermal evolution. *Earth and Planetary Science Letters*, *500*, 86–96.

Lin, S.-C., & van Keken, P. E. (2006). Dynamics of thermochemical plumes: 1. Plume formation and entrainment of a dense layer. *Geochemistry, Geophysics, Geosystems*, *7*, Q02006.

Maljaars, J. M., Labeur, R. J., & Möller, M. (2018). A hybridized discontinuous Galerkin framework for high-order particle–mesh operator splitting of the incompressible Navier–Stokes equations. *Journal of Computional Physics*, *358*, 150–172.

Maljaars, J. M., Labeur, R. J., Trask, N., & Sulsky, D. (2019). Conservative, high-order particle–mesh scheme with applications to advection-dominated flows. *Computer Methods in Applied Mechanics and Engineering*, *348*, 443–465.

Maljaars, J. M., Labeur, R. J., Trask, N., & Sulsky, D. (2020). Optimization based particle-mesh algorithm for high-order and conservative scalar transport. In E. van Brummelen, A. Corsini, S. Perotto, & G. Rozza (Eds.), *Lecture notes in computional science and engineering* (Vol. 132, pp. 265–275). Cham: Springer.

Maljaars, J. M., Richardson, C. N., & Sime, N. (2021). LEoPart: a particle library for FEniCS. *Computers & Mathematics with Applications*, *81*, 289-315.

McDermott, R., & Pope, S. (2008). The parabolic edge reconstruction method (PERM) for Lagrangian particle advection. *Journal of Computional Physics*, *227*(11), 5447–5491.

McNamara, A. K., & Zhong, S. (2004). Thermochemical structures within a spherical mantle: Superplumes or piles? *Journal of Geophysical Research*, *109*, B07402.

Moresi, L., Dufour, F., & Mühlhaus, H.-B. (2003). A Lagrangian integration point finite element method for large deformation modeling of viscoelastic geomaterials. *Journal of Computional Physics*, *184*(2), 476–497.

Morgan, H., & Scott, L. R. (2018). Towards a unified finite element method for the Stokes equations. *SIAM Journal on Scientific Computing*, *40*, A130–A141.

Nakagawa, T., Tackley, P. J., Deschamps, F., & Connolly, J. A. D. (2010). The influence of MORB and harzburgite composition on thermo-chemical mantle convection in a 3-D spherical shell with self-consistently calculated mineral physics. *Earth and Planetary Science Letters*, *296*, 403–412.

Nguyen, N. C., Peraire, J., & Cockburn, B. (2011). Hybridizable discontinuous Galerkin methods. In *Spectral and high order methods for partial differential equations* (pp. 63–84). Berlin, Germany: Springer.

O'Neill, C., Moresi, L., Müller, D., Albert, R., & Dufour, F. (2006). Ellipsis 3d: A particle-in-cell finite element hybrid code for modelling mantle convection and lithospheric deformation. *Computers & Geosciences*, *32*, 1769–1779.

Peraire, J., Nguyen, N., & Cockburn, B. (2010). A hybridizable discontinuous Galerkin method for the compressible Euler and Navier–Stokes equations. In *48th AIAA aerospace sciences meeting and exhibit* (p. 363). doi: 10.2514/6.2010-363

Pilliod, J. E., & Puckett, E. G. (2004). Second-order accurate volume-of-fluid algorithms for tracking material interfaces. *Journal of Computional Physics*, *199*(2), 465–502.

Rhebergen, S., & Wells, G. N. (2017). Analysis of a hybridized/interface stabilized finite element method for the stokes equations. *SIAM Journal on Numerical Analysis*, *55*,

1982–2003.

Rhebergen, S., & Wells, G. N. (2018). A hybridizable discontinuous Galerkin method for the Navier–Stokes equations with pointwise divergence-free velocity field. *Journal of Scientific Computing*, *76*, 1484–1501.

Rhebergen, S., & Wells, G. N. (2020). An embedded–hybridized discontinuous Galerkin finite element method for the Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, *358*, 112619.

Robey, J. M., & Puckett, E. G. (2019). Implementation of a volume-of-fluid method in a finite element code with applications to thermochemical convection in a density stratified fluid in the earth's mantle. *Computers & Fluids*, *190*, 217–253.

Samuel, H., & Evonuk, M. (2010). Modeling advection in geophysical flows with particle level sets. *Geochemistry, Geophysics, Geosystems*, *11*, Q08020.

Schmeling, H. (1987). On the relation between initial conditions and late stages of Rayleigh-Taylor instabilities. *Tectonophysics*, *133*, 65–80.

Schubert, G., Turcotte, D. L., & Olson, P. (2001). *Mantle Convection in the Earth and Planets*. Cambridge, U.K.: Cambridge University Press.

Scott, L. R., & Vogelius, M. (1985). Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials. *ESAIM: Mathematical Modelling and Numerical Analysis*, *19*, 111–143.

Sime, N. (2021). *GeoPart: Particles in geodynamics supporting code.* `https://bitbucket.org/nate-sime/geopart`.

Sime, N., Maljaars, J. M., Wilson, C. R., & van Keken, P. E. (2021). An exactly mass conserving and pointwise divergence free velocity method: application to compositional buoyancy driven flow problems in geodynamics. *Geochemistry, Geophysics, Geosystems*, *22*, e2020GC009349.

Stixrude, L., & Lithgow-Bertelloni, C. (2011). Thermodynamics of mantle minerals – II. Phase equilibria. *Geophysical Journal International*, *184*, 1180–1213.

Suckale, J., Nave, J.-C., & Hager, B. H. (2010). It takes three to tango: 1. Simulating buoyancy-driven flow in the presence of large viscosity contrasts. *Journal of Geophysical Research*, *115*, B07409.

Tackley, P. J. (1996). Effects of strongly variable viscosity on three-dimensional compressible convection in planetary mantles. *Journal of Geophysical Research*, *101*, 3311–3332.

Tackley, P. J., & King, S. D. (2003). Testing the tracer ratio method for modeling active compositional fields in mantle convection simulations. *Geochemistry, Geophysics, Geosystems*, *4*, 8302.

Tan, E., & Gurnis, M. (2007). Compressible thermochemical convection and application to lower mantle structures. *Journal of Geophysical Research*, *112*, B06304.

van Keken, P. E., Ballentine, C., & Hauri, E. H. (2014). Convective mixing in the Earth's mantle. In H. Holland & K. Turekian (Eds.), *Treatise on Geochemistry (2nd edition)* (Vols. 3 "The Mantle and Core", R. Carlson (Ed.), pp. 509–525). Amsterdam, The Netherlands: Elsevier.

van Keken, P. E., Hauri, E. H., & Ballentine, C. J. (2002). Mantle mixing: The generation, preservation, and destruction of chemical heterogeneity. *Annual Review of Earth and Planetary Sciences*, *30*, 493–525.

van Keken, P. E., King, S. D., Schmeling, H., Christensen, U. R., Neumeister, D., & Doin, M.-P. (1997). A comparison of methods for the modeling of thermochemical convection. *Journal of Geophysical Research*, *102*, 22477–22495.

van Summeren, J. R. G., van den Berg, A. P., & van der Hilst, R. D. (2009). Upwellings for a deep mantle reservoir filtered at the 660 km phase transition in thermochemical convection models and implications for intra-plate volcanism. *Physics of the Earth and Planetary Interiors*, *172*, 210–224.

Vila-Pérez, J., Giacomini, M., Sevilla, R., & Huerta, A. (2021). Hybridisable discontinuous Galerkin formulation of compressible flows. *Archives of Computational Methods in Engineering*, *28*, 753–784.

Wang, H., Agrusta, R., & van Hunen, J. (2015). Advantages of a conservative velocity interpolation (cvi) scheme for particle-in-cell methods with application in geodynamic modeling. *Geochemistry, Geophysics, Geosystems*, *16*(6), 2015–2023.

Woopen, M., Balan, A., May, G., & Schütz, J. (2014). A comparison of hybridized and standard DG methods for target-based *hp*-adaptive simulation of compressible flow. *Computers & Fluids*, *98*, 3–16.

Zalesak, S. T. (1979). Fully multidimensional flux-corrected transport algorithms for fluids. *Journal of Computional Physics*, *31*(3), 335–362.

Zhong, S. (2006). Constraints on thermochemical convection of the mantle from plume heat flux, plume excess temperature, and upper mantle temperature. *Journal of Geophysical Research*, *111*, B04409.

Zhong, S., McNamara, A., Tan, E., Moresi, L., & Gurnis, M. (2008). A benchmark study on mantle convection in 3-D spherical shell using CitcomS. *Geochemistry, Geophysics, Geosystems*, *9*, Q10017.

Zindler, A., & Hart, S. (1986). Chemical geodynamics. *Annual Reviews of the Earth and Planetary Sciences*, *14*, 493–571.