Responses to reviewer comments on the manuscript: “Slate: extending Firedrake’s domain-specific abstraction to hybridized solvers for geoscience and beyond.”

Thomas H. Gibson†
†Email: t.gibson15@imperial.ac.uk

November 20, 2019

1 Anonymous Referee # 1

The authors would like to sincerely thank the referee for the careful review and constructive comments for improving the manuscript.

1) How is the .inv() command translated to generated code? Are you using LU/Cholesky or directly inverting the matrix? It is clear the former for the A.solve() command as the decomposition is passed, but not for .inv().

This is a good question. As mentioned on page 6, Slate expressions will get transformed into C++ code using Eigen as the main linear algebra interface. By default, A.inv() gets translated to the corresponding Eigen call: .inverse(), which uses LU with partial pivoting for general matrices. This is also the default behavior of A.solve(B) when a factorization strategy is not provided. A complete list of factorization strategies which can be used with Slate are provided here: https://eigen.tuxfamily.org/dox/group__TopicLinearAlgebraDecompositions.html. This has been clarified in the revised manuscript (after Section 2.1.2).

2) How would you deal in practice with non-linear problems solved using, e.g. a Newton-Krylov method. Is it possible to compose SLEPc and the Preconditioned Krylov Solvers? How would one setup SLEPc to call the code to recover the internal/local variables between Newton iterations?

For global non-linear problems, one can use the technology provided in Section 4 of the paper together with PETSc or SLEPc non-linear solvers. Standard Newton, for example, requires the solution of the linearized Jacobian system for the linear updates. Each Newton iteration therefore requires a linear solve. In Firedrake, the Jacobian is constructed by differentiating the non-linear PDE (expressed in UFL), which generates yet another UFL expression for the Jacobian. The result is then used as the operator for the linear solver (KSP). Hybridization or static condensation can then be used by specifying the correct solver options. In practice, a single Newton iteration would look very similar to the Picard method described in Section 5.2. By design, all preconditioners presented in Section 4 are entirely composable with the PETSc library and can be applied in the usual way, even when nested inside of an outer non-linear method (e.g. snes_type newtonls, ksp_type preonly, pc_type python, pc_python_type firedrake.SCPC). This follows from the framework developed in [1].

3) What does the generated Eigen code look like?

Slate’s linear algebra compiler generates templated C++ code conforming to PyOP2’s application program interface, described in [2]. A typical generated kernel will take an output tensor, any coefficients, and possibly external data as arguments. The body of the kernel will make appropriate function calls to local assembly kernels, generated by Firedrake’s form compiler TSFC [3]. Local data structures (element matrices/vectors) are derived types of Eigen’s Matrix class (https://eigen.tuxfamily.org/dox/group__TutorialMatrixClass.html), and populated by the output of the local assembly kernels. Once all local data structures are populated, the dense linear algebra operations are performed and populated into the output tensor. The result is then passed onto Firedrake’s global assembler (PyOP2).

We have discussed the possibility of providing a code listing in the paper, which would show some generated C++. We came to the conclusion that it would be unhelpful at best and, at worst, distracting away from the central message of the paper, which is the Slate abstraction itself.
4) It is not totally clear how the output from TSFC is fed into the linear algebra compiler. Do you call the TSFC kernel, get the complete cell tensor, split it and then perform the dense linear algebra operations? Or is everything interleaved into one single cell tensor kernel by the linear algebra compiler? Or do you call multiple TSFC kernels, one for each sub-block and then perform the dense linear algebra operations?

TSFC splits an assembly kernel for a mixed operator into separate kernels for each sub-block of the local tensor. Then the single kernel responsible for performing the dense linear algebra will first need to make multiple function calls to gather each local contribution. This made more clear in Section 2.

5) The high-level problem setup in sections 1 and 2 is pretty terse. I don’t think someone who has some knowledge of FEM but is not a real subject expert could get through this section. Some of the wording is quite heavy on jargon too, e.g. global data structure sparse matrix? I appreciate you don’t have time to do a deep-dive into FE, but a bit more text and some pointers to more detailed explanations (other Firedrake papers?) would be useful.

This is a very fair point. We have updated both Sections 1 and 2 to include some further explanation on finite element assembly and what the global data structures are.

Small notational comments: * You do not mention the bold symbol = vector function convention. * Your convention of non-bold capitals being (local? - not sure) discrete linear operators (matrices and vectors) is also not mentioned, although K (the finite element cell) would break this convention. I’d also note that K is used twice for different objects (the finite element cell and a matrix in eq.76). It would probably aid readability if it was possible to distinguish between matrices and vectors, global and local, especially when the operations become more complex in the later sections. There also seem to be (global? - not sure) discrete linear operators in bold.

Notation throughout the entire paper has been updated and made more consistent.

2 Anonymous Referee # 2

The authors would like to sincerely thank the referee for carefully reviewing the manuscript and providing helpful comments to improve the quality of the discussion.

2.1. There are a few inaccuracies and omissions in the historical references for hybridization of mixed finite element methods, particularly in the introduction.

2.1.1. The bibliographic reference to Brezzi and Fortin’s book (from 1991, not 2012) is incorrect. The correct reference is:


The authors may also wish to cite the following, which is essentially an updated version of the Brezzi and Fortin book:


Thank you for pointing this out. We have corrected the reference.

2.1.2. Hybridization and static condensation of mixed methods was apparently first introduced in the following reference:


The referee is absolutely correct. We have added an appropriate citation in the introduction.

2.1.3. Local post-processing appeared in Arnold and Brezzi (1985) for the hybridized RT method and in Brezzi, Douglas, and Marini (1985) for the hybridized BDM method. The authors frequently cite the former but not the latter when mentioning hybridization and post-processing of mixed methods. Also, the 1991 paper of Stenberg on post-processing is cited elsewhere in the paper but not in the introduction. I believe that these three papers should be cited in the relevant part of the introduction (p. 2, l. 18).

This has been fixed in the introduction.
2.2. Section 2.1 is somewhat confusing and could be improved.

2.2.1. The notation \( a(c; v) \) originally made me think that \( a(\cdot; \cdot) \) was a bilinear form. It took a few times through before I made sense of the notation and realized how linear, bilinear, etc., forms are specified. It might be helpful to include one or two concrete examples before introducing a "general form." This is done nicely in Section 2.1 of Alnaes et al. (2014), which I suggest that the authors emulate.

This is a very good point. We have updated Section 2 to include more detailed constructions and definitions, following the convention of Alnaes et al. (2014).

2.2.2. The notation for the \( I_c \) and \( I_f \) integrals is also confusing. Presumably \( c \) stands for "cell" and \( f \) for "facet," but at first I thought \( c \) was the coefficient function \( c \) and \( f \) was some source function, as in equation (10). Since the authors have been using \( T_h \) for cells and \( E_h \) for facets, perhaps a clearer notation would be to call these \( I^T_h \), \( I^{E,c}_h \), and \( I^{E,\partial}_h \). A short sentence mentioning that the three \( I \)s correspond to the contributions from the cells, internal facets, and boundary facets, respectively, would make this easier for the reader to understand.

Another good point and an excellent suggestion for improving our notation. We have incorporated this style in Section 2.

3 Revised Manuscript

All changes suggested by the reviewers have been incorporated. We have also made some clarifying edits to the discussion throughout Section 5. The numerical results themselves have not changed. A complete diff-summary of the paper is provided at the end of this document.

References


Slate: extending Firedrake’s domain-specific abstraction to hybridized solvers for geoscience and beyond

Thomas H. Gibson¹, Lawrence Mitchell², David A. Ham¹, and Colin J. Cotter¹

¹Department of Mathematics, Imperial College London, London, SW7 2AZ, UK
²Department of Computer Science, Durham University, Durham, DH1 3LE, UK

Correspondence: Thomas H. Gibson (t.gibson15@imperial.ac.uk)

Abstract. Within the finite element community, discontinuous Galerkin (DG) and mixed finite element methods have become increasingly popular in simulating geophysical flows. However, robust and efficient solvers for the resulting saddle-point and elliptic systems arising from these discretizations continue to be an on-going challenge. One possible approach for addressing this issue is to employ a method known as hybridization, where the discrete equations are transformed such that classic static condensation and local post-processing methods can be employed. However, it is challenging to implement hybridization as performant parallel code within complex models, whilst maintaining separation of concerns between applications scientists and software experts. In this paper, we introduce a domain-specific abstraction within the Firedrake finite element library that permits the rapid execution of these hybridization techniques within a code-generating framework. The resulting framework composes naturally with Firedrake’s solver environment, allowing for the implementation of hybridization and static condensation as runtime-configurable preconditioners via the Python interface to PETSc, petsc4py. We provide examples derived from second order elliptic problems and geophysical fluid dynamics. In addition, we demonstrate that hybridization shows great promise for improving the performance of solvers for mixed finite element discretizations of equations related to large-scale geophysical flows.

1 Introduction

The development of simulation software is an increasingly important aspect of modern scientific computing, in the geosciences in particular. Such software requires a vast range of knowledge spanning several disciplines, ranging from applications expertise to mathematical analysis to high-performance computing and low-level code optimization. Software projects developing automatic code generation systems have become quite popular in recent years, as such systems help create a separation of concerns which focuses on a particular complexity independent from the rest. This allows for agile collaboration between computer scientists with hardware and software expertise, computational scientists with numerical algorithm expertise, and domain scientists such as meteorologists, oceanographers and climate scientists. Examples of such projects in the domain of finite element methods include FreeFEM++ (Hecht, 2012), Sundance (Long et al., 2010), the FEniCS Project (Logg et al., 2012a), Feel++ (Prud’Homme et al., 2012), and Firedrake (Rathgeber et al., 2016).
The finite element method (FEM) is a mathematically robust framework for computing numerical solutions of partial differential equations (PDEs) that has become increasingly popular in fluids and solids models across the geosciences, with a formulation that is highly amenable to code-generation techniques. A description of the weak formulation of the PDEs, together with appropriate discrete function spaces, is enough to characterize the finite element problem. Both the FEniCS and Firedrake projects employ the *Unified Form Language* (UFL) (Alnæs et al., 2014) to specify the finite element integral forms and discrete spaces necessary to properly define the finite element problem. UFL is a highly expressive domain-specific language (DSL) embedded in Python, which provides the necessary abstractions for code generation systems.

There are classes of finite element discretizations resulting in discrete systems that can be solved more efficiently by directly manipulating local tensors. For example, the static condensation technique for the reduction of global finite element systems (Guyan, 1965; Irons, 1965) produces smaller globally-coupled linear systems by eliminating interior unknowns to arrive at an equation for the degrees of freedom defined on cell-interfaces only. This procedure is analogous to the point-wise elimination of variables used in staggered finite difference codes, such as the ENDGame dynamical core (Melvin et al., 2010; Wood et al., 2014) of the UK Meteorological Office (Met Office), but requires the local inversion of finite element systems. For finite element discretizations of coupled equations relevant to geophysical flows PDEs, the hybridization technique (Arnold and Brezzi, 1985; Brezzi and Fortin, 1991; Cockburn et al., 2009a) provides a mechanism for enabling the static condensation of more complex linear systems. First introduced by Fraeijs de Veubeke (1965) and analyzed further by Brezzi and Fortin (1991); Cockburn et al. (2009a); Boffi et al. (2013), the hybridization method introduces Lagrange multipliers enforcing certain continuity constraints. Local static condensation can then be applied to the augmented system to produce a reduced equation for the multipliers. Methods of this type are often accompanied by local post-processing techniques that produce superconvergent approximations, which exploit the approximation properties of the Lagrange multipliers. This enables the manufacturing of fields exhibiting superconvergent phenomena, or enhanced conservation properties (Bramble and Xu, 1989; Cockburn et al., 2010b). These procedures require invasive manual intervention during the equation assembly process in intricate numerical code.

In this paper, we provide a simple yet effective high-level abstraction for localized dense linear algebra on systems derived from finite element problems. Using embedded DSL technology, we provide a means to enable the rapid development of hybridization and static condensation techniques within an automatic code-generation framework. In other words, the main contribution of this paper is in solving the problem of automatically translating from the mathematics of static condensation and hybridization to compiled code. This automated translation facilitates the separation of concerns between applications scientists and computational computer scientists, and facilitates the automated optimization of compiled code. This framework provides an environment for the development and testing of numerics relevant to the Gung-Ho Project, an initiative by the UK Met Office in designing the next-generation atmospheric dynamical core using mixed finite element methods (Melvin et al., 2019). Our work is implemented in the Firedrake finite element library and the PETSc (Balay et al., 1997; 2019) solver library, accessed via the Python interface petsc4py (Dalcin et al., 2011).

The rest of the paper is organized as follows. We introduce common notation used throughout the paper in Section 1.1. The embedded DSL, called “Slate”, is introduced in Section 2, which allows concise expression of localized linear algebra...
operations on finite element tensors. We provide some contextual examples for static condensation and hybridization in Section 3, including a discussion on post-processing. We then outline in Section 4 how, by interpreting static condensation techniques as a preconditioner, we can go further, and automate many of the symbolic manipulations necessary for hybridization and static condensation. We first demonstrate our implementation on a manufactured problem derived from a second-order elliptic equation, starting in Section 5. The first example compares a hybridizable discontinuous Galerkin (HDG) method with an optimized continuous Galerkin method. Section 5.2 illustrates the composability and relative performance of hybridization for compatible mixed methods applied to a semi-implicit discretization of the nonlinear rotating shallow water equations. Our final example in Section 5.3 demonstrates time-step robustness of a hybridizable solver for a compatible finite element discretization of a rotating linear Boussinesq model. Conclusions follow in Section 6.

1.1 Notation

We begin by establishing notation used throughout this paper. Let $T_h$ denote a tessellation of $\Omega \subset \mathbb{R}^n$, the computational domain, consisting of polygonal elements $K$ associated with a mesh size parameter $h$, and $\partial T_h = \{e \in \partial K : K \in T_h\}$ the set of facets of $T_h$. The set of facets interior to the domain $\Omega$ is denoted by $E^\circ_h = \partial T_h \setminus \partial \Omega$; $E^\circ_h$ denotes the set of facets exterior to the domain $\Omega$. Similarly, we denote the set of exterior facets as $E^\partial_h = \partial T_h \cap \partial \Omega$. For brevity, we denote the finite element integral forms over $T_h$ and any facet set $\Gamma \subset \partial T_h$ by

$$
\left(\frac{u, v}{K}\right) = \int_K u \cdot v \, dx,
$$

$$
\langle u, v \rangle_e = \int_e u \cdot v \, ds,
$$

$$
\left(\frac{u, v}{T_h}\right) = \sum_{K \in T_h} \left(\frac{u, v}{K}\right),
$$

$$
\langle u, v \rangle_\Gamma = \sum_{e \in \Gamma} \langle u, v \rangle_e,
$$

where $dx$ and $ds$ denote appropriate integration measures. The operation $\cdot$ should be interpreted as standard multiplication for scalar functions or a dot product for vector functions.

For any double-valued vector field $w$ on a facet $e \in \partial T_h$, we define the jump of its normal component across $e$ by

$$
[w]_e = \begin{cases} w|_e^+ \cdot n^e_+ + w|_e^- \cdot n^e_-, & e \in E^\circ_h \\ w|_e \cdot n_0, & e \in E^\partial_h \end{cases}
$$

where $+ \text{ and } -$ denote arbitrarily but globally defined sides of the facet. Here, $n^e_+$ and $n^e_-$ are the unit normal vectors with respect to the positive and negative sides of the facet $e$. Whenever the facet domain is clear by the context, we omit the subscripts for brevity and simply write $\left[ \frac{w}{e} \right]$.

2 A system for localized algebra on finite element tensors

We present an expressive language for dense linear algebra on the elemental matrix systems arising from finite element problems. The language, which we call Slate, provides typical mathematical operations performed on matrices and vectors, hence the input syntax is comparable to high-level linear algebra software such as MATLAB. The Slate language provides basic
abstract building blocks which can be used by a specialized compiler for linear algebra to generate low-level code implementations.

Slate is heavily influenced by the Unified Form Language (UFL) (Alnæs et al., 2014; Logg et al., 2012a), a DSL embedded in Python which provides symbolic representations of finite element forms. The expressions can be compiled by a form compiler, which translates UFL into low level code for the local assembly of a form over the cells and facets of a mesh. In a similar manner, Slate expressions are compiled to low level code that performs the requested linear algebra element-wise on a mesh.

2.1 An overview of Slate

To clarify conventions and the scope of Slate, we start by considering a general form. Suppose we have a establishing our notation for a general finite element form following the convention of Alnæs et al. (2014). We define a real-valued multi-linear form as an operator which maps a list of arguments \( v = (v_0, \ldots, v_{\alpha-1}) \in V_0 \times \cdots \times V_{\alpha-1} \) into \( \mathbb{R} \):

\[
a : V_0 \times \cdots \times V_{\alpha-1} \to \mathbb{R}, \quad a \mapsto a(v_0, \ldots, v_{\alpha-1}) = a(v),
\]

(4)

where \( a \) is linear in each argument \( v_k \). The arity of a form is \( \alpha \), an integer denoting the total number of form arguments. In traditional finite element nomenclature (for \( \alpha \leq 2 \)), \( V_0 \) is referred to as the space of test functions and \( V_1 \) as the space of trial functions. Each \( V_k \) are referred to as argument spaces. Forms with arity \( \alpha = 0, 1 \) or 2 are best interpreted as the more familiar mathematical objects: scalars (0-forms), linear forms or functionals (1-forms), and bilinear forms (2-forms) respectively.

If a given form \( a \) is parameterized by one or more coefficients, say \( e = (c_0, \ldots, c_q) \in C_0 \times \cdots \times C_q \) where \( \{C_k\}_{k=0}^q \) are coefficient spaces, then we write:

\[
a : C_0 \times \cdots \times C_q \times V_0 \times \cdots \times V_{\alpha-1} \to \mathbb{R}, \quad a \mapsto a(c_0, \ldots, c_q; v_0, \ldots, v_{\alpha-1}) = a(e; v).
\]

(5)

From here on, we shall work exclusively with forms that are linear in \( v \) and possibly nonlinear in the coefficients \( e \). This is reasonable since nonlinear methods based on Newton iterations produce linear problems via Gâteaux differentiation of a nonlinear form corresponding to a PDE (also known as the form Jacobian). We refer the interested reader to Alnæs et al. (2014, Section 2.1.2) for more details. For clarity, we present examples of multi-linear forms of arity \( \alpha = 0, 1 \) and 2 that frequently appear in finite element form:

\[
a(c; v) = \sum_{K \in \mathcal{T}_h} \int_K I^c(c; v)dx + \sum_{e \in \mathcal{E}^h_\partial} \int_{T_e} I_e^f(c; v)ds + \sum_{e \in \mathcal{E}^h_\partial} \int_{T_e} I_e^{f,\partial}(c; v)ds,
\]
\[ a(\kappa; v, u) := (\nabla v, \kappa \nabla u)_{\mathcal{T}_h} \equiv \sum_{K \in \mathcal{T}_h} \int_K \nabla v \cdot (\kappa \nabla u) \, dx, \quad \kappa \in C_0, \ u \in V_1, \ v \in V_0, \ \alpha = 2, \ q = 1, \] 

\[ a(f; v) := (v, f)_{\mathcal{T}_h} \equiv \sum_{K \in \mathcal{T}_h} \int_K vf \, dx, \quad f \in C_0, \ v \in V_0, \ \alpha = 1, \ q = 1, \] 

\[ a(f, g) := (f - g, f - g)_{\mathcal{T}_h} \equiv \sum_{K \in \mathcal{T}_h} \int_K |f - g|^2 \, dx, \quad g \in C_1, \ f \in C_0, \ \alpha = 0, \ q = 2, \] 

\[ a(\gamma; \sigma) := \langle \gamma, [[\sigma]] \rangle_{\partial \mathcal{T}_h} = \sum_{e \in \mathcal{E}_h^\partial} \int_e \gamma \, ds + \sum_{e \in \mathcal{E}_h^\partial} \int_e \gamma \sigma \cdot n \, ds, \quad \sigma \in V_1, \ \gamma \in V_0, \ \alpha = 2, \ q = 0. \] 

In general, a finite element form will consist of integrals over various geometric domains: integration over cells \( \mathcal{T}_h \), interior facets \( \mathcal{E}_h^\circ \), and exterior facets \( \mathcal{E}_h^\partial \). Therefore, we express a general multi-linear form in terms of integrals over each set of geometric entities:

\[ a(c; v) = \sum_{K \in \mathcal{T}_h} \int_K \mathcal{I}_K^c(c; v) \, dx + \sum_{e \in \mathcal{E}_h^\circ} \int_e \mathcal{I}_e^\circ(c; v) \, ds + \sum_{e \in \mathcal{E}_h^\partial} \int_e \mathcal{I}_e^\partial(c; v) \, ds, \] 

where \( dx \) and \( ds \) denote appropriate integration measures. The integral form in is uniquely determined by its lists (possibly of 0 length) of arbitrary coefficient functions \( c = (c_0, \ldots, c_p) \) in the associated finite element spaces, arguments \( v = (v_0, \ldots, v_q) \) describing any test or trial functions, and its integrand expressions for each integral type: \( \mathcal{I}_K^c, \mathcal{I}_e^\circ, \mathcal{I}_e^\partial, \mathcal{I}_K^T \) denotes a cell integrand on \( K \in \mathcal{T}_h \), \( \mathcal{I}_e^\circ \) is an integrand on the interior facet \( e \in \mathcal{E}_h^\circ \), and \( \mathcal{I}_e^\partial \) is an integrand defined on the exterior facet \( e \in \mathcal{E}_h^\partial \). The form \( a(c; v) \) describes a finite element form globally over the entire problem domain.

Here, we will consider the case where the integrand \( \mathcal{I}_e^\partial(c; v) \) is an integrand on the interior facet \( e \in \mathcal{E}_h^\circ \) and \( \mathcal{I}_e^\partial(c; v) \) can be decomposed into two independent parts for each on each interior facet \( e \) one for the positive restriction (+) and the negative restriction (−).

The integrand \( \mathcal{I}_e^\circ(c; v) \) is the sum of integrals:

\[ \int_{e \subset \partial K^+ \cup \partial K^-} \mathcal{I}_e^\circ(c; v) \, ds = \int_{e \subset \partial K^+} \mathcal{I}_e^\circ(c; v) \, ds + \int_{e \subset \partial K^-} \mathcal{I}_e^\circ(c; v) \, ds. \]

The local contribution of (10) in each cell \( K \) of the mesh \( \mathcal{T}_h \) is simply along with its associated facets \( e \subset \partial K \), is then

\[ a_K(c; v) = \int_K \mathcal{I}_K(c; v) \, dx + \sum_{e \subset \partial K \cap \partial K^+} \int_e \mathcal{I}_e^\circ(c; v) \, ds + \sum_{e \subset \partial K \cap \partial K^-} \int_e \mathcal{I}_e^\circ(c; v) \, ds. \]

We call (12) the cell-local contribution of \( a(c; v) \), with

\[ a(c; v) = \sum_{K \in \mathcal{T}_h} a_K(c; v)|_K. \]
Equation produces an element tensor which is mapped.

To make matters concrete, let us suppose \(a(c; v)\) is a bilinear form with arguments \(v = (v_0, v_1) \in V_0 \times V_1\). Now let \(\{\Phi_i\}_{i=1}^N\) and \(\{\Psi_j\}_{j=1}^M\) denote bases for \(V_0\) and \(V_1\) respectively. Then the global \(N \times M\) matrix \(A\) corresponding to \(a(c; v_0, v_1)\) has its entries defined via

\[
A_{ij} = a(c; \Phi_i, \Psi_j) = \sum_{K \in \mathcal{T}_h} A_{K,ij}, \quad A_{K,ij} = a_K(c; \Phi_i, \Psi_j).
\]  

(14)

By construction, \(A_{K,ij} \neq 0\) if and only if \(\Phi_i\) and \(\Psi_j\) take non-zero values in \(K\). Now we introduce the cell-node map \(i = e(K, i)\) as the mapping from the local node number \(i\) in \(K\) to the global node number \(i\). Suppose there are \(n\) and \(m\) nodes defining the degrees of freedom for \(V_0\) and \(V_1\), respectively, in \(K\). Then all non-zero entries of \(A_{K,ij}\) arise from integrals involving basis functions with local indices corresponding to the global indices \(i, j\):

\[
A^K_{ij} := a_K\left(c; \Phi_{e(K,i)}, \Psi_{e(K,j)}\right), \quad i \in \{1, \ldots, n\}, \quad j \in \{1, \ldots, m\}.
\]  

(15)

These local contributions are collected in the \(n \times m\) dense matrix \(A^K\), which we call the element tensor. The global matrix \(A\) is assembled from the collection of element tensors: \(A = \{A^K\}_{K \in \mathcal{T}_h}\). For details on the general evaluation of finite element basis functions and multi-linear forms, we refer the reader to Kirby (2004); Kirby and Logg (2006); Logg et al. (2012b); Homolya et al. (2018). Further details on the global assembly of finite element operators, with a particular focus on code-generation, are summarized in the work of Logg and Wells (2010); Markall et al. (2013).

In standard finite element software packages, the element tensor is mapped entry-wise into a global data structure. However, before doing so, a sparse array using the cell-node map \(e(K, \cdot)\). Within Firedrake, this operation is handled by PyOP2 (Rathgeber et al., 2012) and serves as the main user-facing abstraction for global finite element assembly. For many applications, one may want to produce a new local tensor-global operator by algebraically manipulating different element tensors. This is precisely the job relatively invasive in numerical code, as it requires bypassing direct operator assembly to produce the new tensor. This is precisely the scope of Slate.

Like UFL, Slate relies on the grammar of the host-language: Python. The entire Slate language is implemented as a Python module which defines its types (classes) and operations on said types. Together, this forms a high-level language for expressing dense linear algebra on element tensors. The Slate language consists of two primary abstractions for linear algebra:

1. terminal element tensors corresponding to multi-linear integral forms (matrices, vectors, and scalars), or assembled data (coefficient vectors for example, coefficient vectors of a finite element function); and

2. expressions consisting of algebraic operations on terminal tensors.

The composition of binary and unary operations on terminal tensors produces a Slate expression. Such expressions can be composed with other Slate objects in arbitrary ways, resulting in concise representations of complex algebraic operations on locally assembled arrays. We summarize all currently supported Slate abstractions here.
2.1.1 Terminal tensors

In Slate, one associates a tensor with data on an element a cell either by using a multi-linear form, or assembled coefficient data:

- $\text{Tensor}(a(c; v))$

associates a form, expressed in UFL, with its local element tensor:

$$A^K A^K \leftarrow a_K(c; v)|_K, \text{ for all } K \in \mathcal{T}_h. \quad (16)$$

The number of arguments $v$ determine-form arity $\alpha$ of $a_K(c; v)$ determines the rank of the corresponding Tensor, i.e. scalars, vectors, and matrices are produced from 0-forms, 1-forms, and 2-forms\textsuperscript{1} respectively scalars, linear forms, and bilinear forms respectively.\textsuperscript{1} The shape of the element tensor is determined by both the number of arguments, and total number of degrees of freedom local to the cell.

- $\text{AssembledVector}(f)$

where $f$ is some finite element function. The result associates a function with its local coefficient vector function $f \in V$ is expressed in terms of the finite element basis of $V$: $f(x) = \sum_{i=1}^{N} f_i \Phi_i(x)$. The result is the local coefficient vector of $f$ on $K$:

$$F^K \leftarrow \{f_{\epsilon(K,i)}\}_{i=1}^{n}, \quad (17)$$

where $\epsilon(K,i)$ is the local node numbering and $n$ is the number of nodes local to the cell $K$.

2.1.2 Symbolic linear algebra

Slate supports typical binary and unary operations in linear algebra, with a high-level syntax close to mathematics. At the time of this paper, these include:

- $A + B$, the addition of two equal shaped tensors: $A^K + B^K$.

- $A * B$, a contraction over the last index of $A$ and the first index of $B$. This is the usual multiplicative operation on matrices, vectors, and scalars: $A^K B^K$.

- $-A$, the additive inverse (negation) of a tensor: $-A^K$.

- $A^T$, the transpose of a tensor: $(A^K)^T$.

\textsuperscript{1}As with UFL, Slate is capable of abstractly representing arbitrary rank tensors. However, only rank $\leq 2$ tensors are typically used in most finite element applications and therefore we currently only generate code for those ranks.

\textsuperscript{1}Similarly to UFL, Slate is capable of abstractly representing arbitrary rank tensors. However, only rank $\leq 2$ tensors are typically used in most finite element applications and therefore we currently only generate code for those ranks.
- A.inv, the inverse of a square tensor: \( (A^K)^{-1} \).

- A.solve(B, decomposition="..."), the result \( X^K \), of solving a local linear system \( AX = B \) for \( X A^K X^K = B^K \), optionally specifying a direct factorization strategy.

- A.blocks[indices], where \( A \) is a tensor from a mixed finite element space. This allows for the extraction of subblocks of the tensor, which are indexed by field (slices are allowed). For example, if a matrix \( A \) corresponds to the bilinear form \( a : V \times W \to \mathbb{R} \), where \( V = V_0 \times \cdots \times V_n \) and \( W = W_0 \times \cdots \times W_m \) are product spaces consisting of finite element spaces \( \{V_i\}_{i=0}^n \), \( \{W_i\}_{i=0}^m \), then the cell-local element tensors have the form:

\[
A^K A^K = \begin{bmatrix}
A^{K}_{00} & A^{K}_{01} & \cdots & A^{K}_{0m} \\
A^{K}_{10} & A^{K}_{11} & \cdots & A^{K}_{1m} \\
\vdots & \vdots & \ddots & \vdots \\
A^{K}_{n0} & A^{K}_{n1} & \cdots & A^{K}_{nm}
\end{bmatrix}.
\]

(18)

The associated submatrix of (18) with indices \( i = (p, q) \), \( p = \{p_1, \ldots, p_r\} \), \( q = \{q_1, \ldots, q_e\} \), is

\[
A^K A^K_{pq} = \begin{bmatrix}
A^{K}_{p_1 q_1} & \cdots & A^{K}_{p_1 q_e} \\
\vdots & \ddots & \vdots \\
A^{K}_{p_r q_1} & \cdots & A^{K}_{p_r q_e}
\end{bmatrix} = A^K A^K_{.blocks[p, q]},
\]

(19)

where \( p \subseteq \{1, \ldots, n\} \), \( q \subseteq \{1, \ldots, m\} \) where \( p \subseteq \{0, \ldots, n\} \), \( q \subseteq \{0, \ldots, m\} \).

These building blocks may be arbitrarily composed, giving a symbolic expression for the linear algebra we wish to perform on each cell during assembly. They- Each Tensor object knows all the information about the underlying UFL form that defines it, such as form arguments, coefficients, and the underlying finite element space(s) it operates on. This information is propagated through as unary or binary transformations are applied. The unary and binary operations shown here provide the necessary algebraic framework for a large class of problems, some of which we present in this paper.

In Firedrake, these Slate expressions are transformed into low-level code by a linear algebra compiler. This uses the form compiler, TSFC (Homolya et al., 2018a), to compile kernels for the assembly of terminal tensors and generates a dense linear algebra kernel to be iterated cell-wise. The compiler interprets Slate expressions as a syntax tree, where the tree is visited to identify what local arrays need to be assembled and the sequence of array operations. At the time of this work, our compiler generates C++ code, using the templated library Eigen (Guennebaud et al., 2015) for dense linear algebra. During execution, the local computations in each cell are mapped into global data objects via appropriate indirection mappings using the translation from Slate to C++ is fairly straightforward, as all operations supported by Slate have a representation in Eigen.

The compiler pass will generate a single “macro” kernel, which performs the dense linear algebra operations represented in Slate. The resulting code will also include (often multiple) function calls to local assembly kernels generated by TSFC (Homolya et al., 2018b) to assemble all necessary sub-blocks of an element tensor. All code generated by the linear algebra
Figure 1. The Slate language wraps UFL objects describing the finite element system. The resulting Slate expressions are passed to a specialized linear algebra compiler, which produces a single “macro” kernel assembling the local contributions and executes the dense linear algebra represented in Slate. The kernels are passed to the Firedrake’s PyOP2 interface, which wraps the Slate kernel in a mesh-iteration kernel. Parallel scheduling, code generation, and compilation occurs after the PyOP2 layer.

Most optimization of the resulting dense linear algebra code is handled directly by Eigen. In the case of unary and binary operations such as \( A.inv \) and \( A.solve(B) \), stable default behaviors are applied by the linear algebra compiler. For example, \( A.solve(B) \) without a specified factorization strategy will default to using an in-place LU factorization with partial pivoting. For local matrices smaller than \( 5 \times 5 \), the inverse is translated directly into Eigen’s \( A.inverse() \) which employs stable analytic formulas. For larger matrices, the linear algebra replaces \( A.inv \) with an LU factorization.\(^2\) Currently, we only support direct matrix factorizations for solving local linear systems. However, it would not be difficult to extend Slate to support more general solution techniques like iterative methods.

3 Examples

We now present a few examples and discuss solution methods which require element-wise manipulations of finite element systems and their specification in Slate. We stress here that Slate is not limited to these model problems; rather these examples were chosen for clarity and to demonstrate key features of the Slate language. In Sections 4 and 5, we discuss more intricate ways Slate is used in custom preconditioners.

\(^2\)For more details on solving linear equations in Eigen, see: https://eigen.tuxfamily.org/dox/group__TutorialLinearAlgebra.html
For the hybridization of mixed and discontinuous Galerkin methods, for our discussion, we use a model elliptic equation defined in a computational domain $\Omega$. Consider the second-order PDE with both Dirichlet and Neumann boundary conditions:

\[
-\nabla \cdot (\kappa \nabla p) + cp = f, \quad \text{in } \Omega, \\
p = p_0, \quad \text{on } \partial \Omega_D, \\
-\kappa \nabla p \cdot n = g, \quad \text{on } \partial \Omega_N,
\]

where $\partial \Omega_D \cup \partial \Omega_N = \partial \Omega$ and $\kappa, c : \Omega \to \mathbb{R}^+$ are positive-valued coefficients. To obtain a mixed formulation of (20)–(22), we introduce the auxiliary velocity variable $u = -\kappa \nabla p$. We then obtain the first-order system, we obtain the mixed problem: system of PDEs:

\[
\mu u + \nabla p = 0, \quad \text{in } \Omega, \\
\nabla \cdot u + cp = f, \quad \text{in } \Omega, \\
p = p_0, \quad \text{on } \partial \Omega_D, \\
u \cdot n = g, \quad \text{on } \partial \Omega_N,
\]

where $\mu = \kappa^{-1}$ and $u = -\kappa \nabla p$ is the velocity variable.

### 3.1 Hybridization of mixed methods

To motivate our discussion in this section, we start by recalling the mixed method for (23)–(26). Methods of this type seek approximations $(u_h, p_h)$ in finite-dimensional subspaces $U_h \times V_h \subset H(\text{div}) \times L^2 U_h \times V_h \subset H(\text{div}; \Omega) \times L^2(\Omega)$, defined by:

\[
U_h = \{ w \in H(\text{div}; \Omega) : w \mid K \in U(K), \forall K \in T_h, w \cdot n = g \text{ on } \partial \Omega_N \}, \\
V_h = \{ \phi \in L^2(\Omega) : \phi \mid K \in V(K), \forall K \in T_h \}.
\]

The space $U_h$ consists of $H(\text{div})$-conforming piecewise vector polynomials, where choices of $U(K)$ typically include the Raviart-Thomas (RT), Brezzi-Douglas-Marini (BDM), or Brezzi-Douglas-Fortin-Marini (BDFM) elements (Brezzi et al., 1987, 1985; Nédélec, 1980; Raviart and Thomas, 1977). The space $V_h$ is the Lagrange family of discontinuous polynomials. These spaces are of particular interest when simulating geophysical flows, since choosing the right pairing results in stable discretizations with desirable conservation properties and avoids spurious computational modes. We refer the reader to Cotter and Shipton (2012); Cotter and Thuburn (2014); Natale et al. (2016); Shipton et al. (2018) for a discussion of mixed methods relevant for geophysical fluid dynamics. Two examples of such a discretization is discretizations are presented in Section 5.2.

The mixed finite element formulation of (23)–(26) is arrived at by multiplying (23)–(24) by test functions and integrating by parts. The resulting finite element problem reads as follows: find $(u_h, p_h) \in U_h \times V_h$ satisfying $(u_h, p_h) \in U_h \times V_h$. satisfying
\[
\begin{align*}
\left( w, \mu u_h \right)_{\Omega_h} - \left( \nabla w, p_h \right)_{\Omega_h} &= - \left( w \cdot n, p_0 \right)_{\partial \Omega_D}, \quad \forall w \in U_{h,0}, \\
\left( \phi, \nabla u_h \right)_{\Omega_h} + \left( \phi, c p_h \right)_{\Omega_h} &= \left( \phi, f \right)_{\Omega_h}, \quad \forall \phi \in V_h,
\end{align*}
\]

where \( U_{h,0} \) is the space of functions in \( U_h \) such that \( \phi \in U_{h,0} \) is the subspace of \( U_h \) with functions whose normal components vanish on \( \partial \Omega_N \). The discrete system is obtained by first expanding the solutions in terms of the finite element bases:

\[
u_h = \sum_{i=1}^{N_u} U_i \Psi_i, \quad p_h = \sum_{i=1}^{N_p} P_i \xi_i,
\]

where \( \{\Psi_i\}_{i=1}^{N_u} \) and \( \{\xi_i\}_{i=1}^{N_p} \) are bases for \( U_h \) and \( V_h \) respectively. Here, \( U_i \) and \( P_i \) are the coefficients to be determined. As per standard Galerkin-based finite element methods, taking \( w = \Psi_j, j \in \{1, \ldots, N_u\}, j \in \{1, \ldots, N_u\} \) and \( \phi = \xi_j, j \in \{1, \ldots, N_p\}, j \in \{1, \ldots, N_u\} \) in (29)–(30) produces the discrete saddle point system:

\[
\begin{bmatrix}
A & -B^T \\
B & D
\end{bmatrix}
\begin{bmatrix}
U \\
P
\end{bmatrix}
= \begin{bmatrix}
F_0 \\
F_1
\end{bmatrix},
\]

where \( U = \{U_i\}, P = \{P_i\} \) \( U = \{U_i\}_{i=1}^{N_u}, P = \{P_i\}_{i=1}^{N_p} \) are the coefficient vectors, and

\[
\begin{align*}
A_{ij} &= \left( \Psi_i, \mu \Psi_j \right)_{\Omega_h}, \\
B_{ij} &= \left( \xi_i, \nabla \cdot \Psi_j \right)_{\Omega_h}, \\
C_{ij} &= \left( \xi_i, c \Psi_j \right)_{\Omega_h}, \\
F_{0,j} &= - \left( \Psi_j \cdot n, p_0 \right)_{\partial \Omega_D}, \\
F_{1,j} &= \left( \xi_j, f \right)_{\Omega_h}.
\end{align*}
\]

Methods to efficiently invert such systems include \( H(\text{div})H(\text{div}) \)-multigrid (Arnold et al., 2000) (requiring complex overlapping-Schwarz smoothers), global Schur-complement factorizations (which require an approximation to the inverse of the dense\(^3\) elliptic Schur-complement \( C + B A^{-1} B^T D + B A^{-1} B^T \)), or auxiliary space multigrid (Hiptmair and Xu, 2007). Here, we focus on a solution approach using a hybridized mixed method (Arnold and Brezzi, 1985; Brezzi and Fortin, 1991) (Arnold and Brezzi, 1985; Brezzi and Fortin, 1991; Boffi et al., 2013).

The hybridization technique replaces the original system with a discontinuous variant, decoupling the velocity degrees of freedom between cells. This is done by replacing the discrete solution space for \( u_h \) with the “broken” space \( U_h : U_h \), defined

\(^3\)The Schur-complement, while elliptic, is globally dense due to the fact that \( A^{-1} A \) has a dense inverse. This is a result of velocities in \( U_h \) having continuous normal components across cell-interfaces.
as:

\[ U^d_h = \{ w \in [L^2(\Omega)]^n : w|_K \in U(K), \forall K \in \mathcal{T}_h \}. \]  \hfill (38)

The vector finite element space \( U^d_h \) is a subspace of \([L^2(\Omega)]^n\) consisting of local \( H(\text{div}) - H(\text{div}) \) functions, but normal components are no longer required to be continuous on \( \partial \mathcal{T}_h \). The approximation space for \( p_h \) remains unchanged.

Next, Lagrange multipliers are introduced as an auxiliary variable in the space \( M_h \), defined only on cell-interfaces:

\[ M_h = \{ \gamma \in L^2(\partial \mathcal{T}_h) : \gamma|_e \in M(e), \forall e \in \partial \mathcal{T}_h \}, \]  \hfill (39)

where \( M(e) \) denotes a polynomial space defined on each facet. We call \( M_h \) the space of approximate traces. Functions in \( M_h \) are discontinuous across vertices in two-dimensions, and vertices/edges in three-dimensions.

Deriving the hybridizable mixed system is accomplished through integration by parts over each element \( K \). Testing with \( w \in U^d_h(K), w \in U^d_h(K) \) and integrating (23) over the cell \( K \) produces:

\[
\left( w, \mu u^d_h \right)_K - \left( \nabla \cdot w, p_h \right)_K + \left( w \cdot n, \lambda_h \right)_{\partial K} = -\left( w \cdot n, p_0 \right)_{\partial K \cap \partial \Omega_D}.
\]  \hfill (40)

The trace function \( \lambda_h \) is introduced in surface integrals and approximates \( p_h \) on elemental boundaries: the surface integral as an approximation to \( p|_{\partial K} \). An additional constraint equation, called the transmission condition, is added to close the system. The resulting hybridizable formulation reads: find \((u^d_h, p_h, \lambda_h) \in U^d_h \times V_h \times M_h \) such that \((u^d_h, p_h, \lambda_h) \in U^d_h \times V_h \times M_h \) such that

\[
\left( w, \mu u^d_h \right)_{\mathcal{T}_h} - \left( \nabla \cdot w, p_h \right)_{\mathcal{T}_h} + \left( [w], \lambda_h \right)_{\partial \mathcal{T}_h \setminus \partial \Omega_D} = -\left( w \cdot n, p_0 \right)_{\partial \Omega_D}, \quad \forall w \in U^d_h, \]  \hfill (41)

\[
\left( \phi, \nabla \cdot u^d_h \right)_{\mathcal{T}_h} + \left( \phi, c p_h \right)_{\mathcal{T}_h} = \left( \phi, f \right)_{\mathcal{T}_h}, \quad \forall \phi \in V_h, \]  \hfill (42)

\[
\left( \gamma, [u^d_h] \right)_{\partial \mathcal{T}_h \setminus \partial \Omega_D} = \left( \gamma, [g] \right)_{\partial \Omega_N}, \quad \forall \gamma \in M_{h,0}, \]  \hfill (43)

where \( M_{h,0} \) denotes the space of traces vanishing on \( \partial \Omega_D \). The constraint in transmission condition (43) enforces both continuity of the normal components of \( u^d_h \) across element and \( u^d_h \cdot n \) across element boundaries, as well as the boundary condition: \( u^d_h \cdot n = g \) on \( \partial \Omega_N \). If the space of Lagrange multipliers \( M_h \) is chosen appropriately, then the “broken” velocity \( u^d_h \), albeit sought a priori in a discontinuous space, will coincide with its \( H(\text{div}) - H(\text{div}) \)-conforming counterpart. Specifically, the formulations in (41)–(42) and (29)–(30) are solving equivalent problems if the normal components of \( w \in U_h, w \in U_h \) lie in the same polynomial space as the trace functions (Arnold and Brezzi, 1985).

The discrete matrix system arising from (41)–(43) has the general form:

\[
\begin{bmatrix}
A_{00} & A_{01} & A_{02} \\
A_{10} & A_{11} & A_{12} \\
A_{20} & A_{21} & A_{22}
\end{bmatrix}
\begin{bmatrix}
U^d \\
P \\
\Lambda
\end{bmatrix} =
\begin{bmatrix}
F_0 \\
F_1 \\
F_2
\end{bmatrix},
\]  \hfill (44)
where the discrete system is produced by expanding functions in terms of the finite element bases for \( U_{h}^{d} \), \( V_{h} \), and \( M_{h} \) like before. Upon initial inspection, it may not appear to be advantageous to replace our original formulation with this augmented equation-set; the hybridizable system has substantially more total degrees of freedom. However, (44) has a considerable advantage over (32) in the following ways:

1. Since both \( U_{h}^{d} \) and \( V_{h} \) are discontinuous spaces, \( U_{h}^{d} \) and \( P_{h} U_{h}^{d} \) are coupled only within the cell. This allows us to simultaneously eliminate both unknowns via local static condensation to produce a significantly smaller global (hybridized) problem for the trace unknowns, \( A \Lambda \Lambda \):

\[
S \Lambda S A = F E,
\]

where \( S = \{ S_{K} \} \) and \( E = \{ E_{K} \} \) are assembled by gathering the local contributions: \( S \leftarrow \{ S_{K} \} \) and \( E \leftarrow \{ E_{K} \} \) are assembled via the local element tensors:

\[
SS_{K} = A A_{22}^{K} - \begin{bmatrix} A_{20}^{K} & A_{21}^{K} \\ A_{10}^{K} & A_{11}^{K} \end{bmatrix} \begin{bmatrix} A_{00}^{K} & A_{01}^{K} \\ A_{10}^{K} & A_{11}^{K} \end{bmatrix}^{-1} \begin{bmatrix} A_{02}^{K} \\ A_{12}^{K} \end{bmatrix},
\]

\[
EE_{K} = F F_{2}^{K} - \begin{bmatrix} A_{20}^{K} & A_{21}^{K} \\ A_{10}^{K} & A_{11}^{K} \end{bmatrix} \begin{bmatrix} A_{00}^{K} & A_{01}^{K} \\ A_{10}^{K} & A_{11}^{K} \end{bmatrix}^{-1} \begin{bmatrix} F_{0}^{K} \\ F_{1}^{K} \end{bmatrix}.
\]

Note that the inverse of the block matrix in (46) and (47) is never evaluated globally; the elimination can be performed locally by performing a sequence of Schur-complement reductions within each cell.

2. The matrix \( S \) is sparse, symmetric, positive-definite, and spectrally equivalent to the dense Schur-complement \( C + B A^{-1} B^{T} D + B A^{-1} B^{T} \) from (32) of the original mixed formulation (Cockburn et al., 2009a; Gopalakrishnan, 2003; Cockburn et al., 2009b).

3. Once \( A \Lambda \Lambda \) is computed, both \( U_{h}^{d} \) and \( P_{h} U_{h}^{d} \) can be recovered locally in each element. This can be accomplished in a number ways. One way is to compute \( P K P K \) by solving:

\[
(A A_{11}^{K} - A A_{10}^{K} (A A_{00}^{K})^{-1} A A_{01}^{K}) P K = F F_{1}^{K} - A A_{10}^{K} (A A_{00}^{K})^{-1} F F_{0}^{K} - (A A_{12}^{K} - A A_{10}^{K} (A A_{00}^{K})^{-1} A A_{02}^{K}) A \Lambda^{K},
\]

followed by solving for \( \{ U_{h}^{d} \}^{K} \):

\[
A A_{00}^{K} (U U_{h}^{d})^{K} = F F_{0}^{K} - A A_{01}^{K} P K - A A_{02}^{K} A \Lambda^{K}.
\]

Similarly, one could rearrange the order in which each variable is reconstructed.

4. If desired, the solutions can be improved further through local post-processing. We highlight two such procedures for \( U_{h}^{d} \) and \( P_{h} \) respectively, in Section 3.3.
Listing Figure 2 displays the corresponding Slate code for assembling the trace system, solving (45), and recovering the eliminated unknowns. For a complete reference on how to formulate the hybridized mixed system (41)–(43) in UFL, we refer the reader to Alnæs et al. (2014). Complete Firedrake code using Slate to solve a hybridizable mixed system is also publicly available in Zenodo/Tabula-Rasa (2019, “Code verification”). We remark that, in the case of this hybridizable system, (44) contains zero-valued blocks which can simplify the resulting expressions in (46)–(47) and (48)–(49). This is not true in general and therefore the expanded form using all sub-blocks of (44) is presented for completeness.

3.2 Hybridization of discontinuous Galerkin methods

The hybridized discontinuous Galerkin (HDG) method is a natural extension of discontinuous Galerkin (DG) discretizations. Here, we consider a specific HDG discretization, namely the LDG-H method (Cockburn et al., 2010b). Other forms of HDG that involve local lifting operators can also be implemented in this software framework by the introduction of additional local (i.e., discontinuous) variables in the definition of the local solver.

To construct the LDG-H discretization formulation follows exactly from standard DG methods. All prognostic variables are sought in the discontinuous spaces $U_h \times V_h \subset [L^2(\Omega)]^n \times L^2(\Omega)$. Within a cell $K$, integration by parts yields:

\[
(w, \mu u_h)_K - (\nabla \cdot w, p_h)_K + \langle w \cdot n, \hat{p} \rangle_{\partial K} = 0, \quad \forall w \in U(K),
\]

\[
-(\nabla \phi, u_h)_K + \langle \phi \cdot n, \hat{p} \rangle_{\partial K} + (\phi, cp_h)_K = (\phi, f)_K, \quad \forall \phi \in V(K),
\]

where $U(K)$ and $V(K)$ are vector and scalar polynomial spaces respectively. Now, we define the DG numerical fluxes $\hat{p}$ and $\hat{u}$ to be functions of the trial unknowns and a new independent unknown in the trace space $M_h$:

\[
\hat{u}(u_h, p_h, \lambda_h; \tau) = u_h + \tau (p_h - \hat{p}) n,
\]

\[
\hat{p}(\lambda_h) = \lambda_h,
\]

where $\lambda_h \in M_h$ is a function approximating the trace of $p$ on $\partial T_h$ and $\tau$ is a positive stabilization function that may vary on each facet $e \in \partial T_h$. We further require that $\lambda_h$ satisfies the Dirichlet condition for $p$ on $\partial \Omega_D$ in an $L^2$-projection sense. The full LDG-H formulation reads as follows. Find $(u_h, p_h, \lambda_h) \in U_h \times V_h \times M_h$ such that

\[
(w, \mu u_h)_{\tau_h} - (\nabla w, p_h)_{\tau_h} + \langle [[w]], \lambda_h \rangle_{\partial \tau_h} = 0, \quad \forall w \in U_h,
\]

\[
-(\nabla \phi, u_h)_{\tau_h} + \langle \phi, \left[ u_h + \tau (p_h - \lambda_h) n \right] \rangle_{\partial \tau_h} + (\phi, cp_h)_{\tau_h} = (\phi, f)_{\tau_h}, \quad \forall \phi \in V_h,
\]

\[
\langle \gamma_\partial \left[ u_h + \partial \Omega_D (p_h - \lambda_h) n \right] \rangle_{\partial \tau_h \setminus \partial \Omega_D} = \langle \gamma_\partial g \rangle_{\partial \Omega_N} + \langle \gamma_\partial p_0 \rangle_{\partial \Omega_D}, \quad \forall \gamma \in M_h,
\]

Equation (56) enforces continuity of the numerical flux $\hat{u}$ is the transmission condition, which enforces the continuity of $\hat{u} \cdot n$ on $\partial T_h$, which in turn produces a flux that and (57) ensures $\lambda_h$ satisfies the Dirichlet condition. This ensures that the numerical
# Element tensors defining the local 3-by-3 block system

```python
_A = Tensor(a)
_F = Tensor(L)
```

# Extracting blocks for Slate expression of the reduced system

```python
A = _A.blocks
F = _F.blocks
S = assemble(Sexp, bcs=[...])  # Assemble S
E = assemble(Eexp)  # Assemble E
lambda_h = Function(M)  # Function to store the result: $\Lambda$
```

# Solve for the Lagrange multipliers: $\Lambda$

```python
p_h = Function(V)  # Function to store the result: $P$
Lambda = AssembledVector(lambda_h)  # Local coefficient vector: $\Lambda^K$
P = AssembledVector(p_h)  # Local coefficient vector: $P^K$
```

# Intermediate expressions

```python
Sd = A[1, 1] - A[1, 0] * A[0, 0].inv * A[0, 1]
Sl = A[1, 2] - A[1, 0] * A[0, 0].inv * A[0, 2]
```

# Slate expressions for local recovery

```python
p_sys = Sd.solve(F[1] - A[1, 0] * A[0, 0].inv * F[0] - Sl * Lambda,
               decomposition="PartialPivLu")
_u_sys = A[0, 0].solve(F[0] - A[0, 1] * P - A[0, 2] * Lambda,
               decomposition="PartialPivLu")
assemble(p_sys, p_h)  # Solve for $P$
assemble(_u_sys, u_h)  # Solve for $U^d$
```

**Figure 2.** Firedrake code for solving (44) via static condensation and local recovery, given UFL expressions $a, L$ for (41)–(43). Arguments of the mixed space $U^d_h \times V_h \times M_h$ are indexed by $0, 1, 2$ respectively. Lines 8 and 9 are symbolic expressions for (46) and (47) respectively. Any vanishing conditions on the trace variables should be provided as boundary conditions during operator assembly (line 10). Lines 26 and 28 are expressions for (48) and (49) (using LU). Code-generation occurs in lines 10, 11, 30, and 31. A global linear solver for the reduced system is created and used in line 15. Configuring the linear solver is done by providing an appropriate Python dictionary of solver options for the PETSc library.

The **LDG-H method retains the advantages of standard DG methods while also enabling the assembly of reduced linear systems through static condensation.** The matrix system arising from (54)–(57) has the same general form as that of the hybridized mixed method in (44), except all sub-blocks are now populated with non-zero entries due to the coupling of trace functions with both $p_h$ and $u_h$. However, all previous properties of the discrete matrix system from Section 3.1 still apply.

---

**Figure 2.** Firedrake code for solving (44) via static condensation and local recovery, given UFL expressions $a, L$ for (41)–(43). Arguments of the mixed space $U^d_h \times V_h \times M_h$ are indexed by $0, 1, 2$ respectively. Lines 8 and 9 are symbolic expressions for (46) and (47) respectively. Any vanishing conditions on the trace variables should be provided as boundary conditions during operator assembly (line 10). Lines 26 and 28 are expressions for (48) and (49) (using LU). Code-generation occurs in lines 10, 11, 30, and 31. A global linear solver for the reduced system is created and used in line 15. Configuring the linear solver is done by providing an appropriate Python dictionary of solver options for the PETSc library.

**Figure 2.** Firedrake code for solving (44) via static condensation and local recovery, given UFL expressions $a, L$ for (41)–(43). Arguments of the mixed space $U^d_h \times V_h \times M_h$ are indexed by $0, 1, 2$ respectively. Lines 8 and 9 are symbolic expressions for (46) and (47) respectively. Any vanishing conditions on the trace variables should be provided as boundary conditions during operator assembly (line 10). Lines 26 and 28 are expressions for (48) and (49) (using LU). Code-generation occurs in lines 10, 11, 30, and 31. A global linear solver for the reduced system is created and used in line 15. Configuring the linear solver is done by providing an appropriate Python dictionary of solver options for the PETSc library.

**Figure 2.** Firedrake code for solving (44) via static condensation and local recovery, given UFL expressions $a, L$ for (41)–(43). Arguments of the mixed space $U^d_h \times V_h \times M_h$ are indexed by $0, 1, 2$ respectively. Lines 8 and 9 are symbolic expressions for (46) and (47) respectively. Any vanishing conditions on the trace variables should be provided as boundary conditions during operator assembly (line 10). Lines 26 and 28 are expressions for (48) and (49) (using LU). Code-generation occurs in lines 10, 11, 30, and 31. A global linear solver for the reduced system is created and used in line 15. Configuring the linear solver is done by providing an appropriate Python dictionary of solver options for the PETSc library.

**Figure 2.** Firedrake code for solving (44) via static condensation and local recovery, given UFL expressions $a, L$ for (41)–(43). Arguments of the mixed space $U^d_h \times V_h \times M_h$ are indexed by $0, 1, 2$ respectively. Lines 8 and 9 are symbolic expressions for (46) and (47) respectively. Any vanishing conditions on the trace variables should be provided as boundary conditions during operator assembly (line 10). Lines 26 and 28 are expressions for (48) and (49) (using LU). Code-generation occurs in lines 10, 11, 30, and 31. A global linear solver for the reduced system is created and used in line 15. Configuring the linear solver is done by providing an appropriate Python dictionary of solver options for the PETSc library.

**Figure 2.** Firedrake code for solving (44) via static condensation and local recovery, given UFL expressions $a, L$ for (41)–(43). Arguments of the mixed space $U^d_h \times V_h \times M_h$ are indexed by $0, 1, 2$ respectively. Lines 8 and 9 are symbolic expressions for (46) and (47) respectively. Any vanishing conditions on the trace variables should be provided as boundary conditions during operator assembly (line 10). Lines 26 and 28 are expressions for (48) and (49) (using LU). Code-generation occurs in lines 10, 11, 30, and 31. A global linear solver for the reduced system is created and used in line 15. Configuring the linear solver is done by providing an appropriate Python dictionary of solver options for the PETSc library.

**Figure 2.** Firedrake code for solving (44) via static condensation and local recovery, given UFL expressions $a, L$ for (41)–(43). Arguments of the mixed space $U^d_h \times V_h \times M_h$ are indexed by $0, 1, 2$ respectively. Lines 8 and 9 are symbolic expressions for (46) and (47) respectively. Any vanishing conditions on the trace variables should be provided as boundary conditions during operator assembly (line 10). Lines 26 and 28 are expressions for (48) and (49) (using LU). Code-generation occurs in lines 10, 11, 30, and 31. A global linear solver for the reduced system is created and used in line 15. Configuring the linear solver is done by providing an appropriate Python dictionary of solver options for the PETSc library.
The Slate expressions for the local elimination and reconstruction operations will be identical to those illustrated in Figure 2. For the interested reader, a unified analysis of hybridization methods (both mixed and DG) for second-order elliptic equations is presented in Cockburn et al. (2009a); Cockburn (2016).

### 3.3 Local post-processing

For both mixed (Arnold and Brezzi, 1985; Bramble and Xu, 1989; Stenberg, 1991)-(Arnold and Brezzi, 1985; Brezzi et al., 1985; Bramble and Xu, 1989; Stenberg, 1991) and discontinuous Galerkin methods (Cockburn et al., 2010b, 2009b), it is possible to locally post-process solutions to obtain superconvergent approximations (gaining one order of accuracy over the unprocessed solution). These methods can be expressed as local solves on each element, and so, in addition to static condensation, the Slate language also provides access to code generation for local post-processing of computed solutions.

Here and are straightforward to implement using Slate. In this section, we present two post-processing techniques: one for scalar fields, and another for the vector unknown. The Slate code follows naturally from previous discussions in Sections 3.1 and 3.2, using the standard set of operations on local element tensors summarized in Section 2.1.

#### 3.3.1 Post-processing of the scalar solution

Our first example is a modified version of the procedure presented by Stenberg (1991) for enhancing the accuracy of the scalar solution. This was also highlighted within the context of hybridizing eigenproblems by Cockburn et al. (2010a). This post-processing technique can be used for both the hybridized hybridizable mixed and LDG-H methods. We proceed by posing the finite element systems cell-wise.

Let $\mathcal{P}_k(K)$ denote a polynomial space of degree $\leq k$ on an element—a cell $K \in T_h$. Then for a given pair of computed solutions $u_h, p_h$ of the hybridized methods, we define the post-processed scalar $p_h^* \in \mathcal{P}_{k+1}(K)$ as the unique solution of the local problem:

\[
\left( \nabla w, \nabla p_h^* \right)_K = - \left( \nabla w, \kappa^{-1} u_h \right)_K, \quad \forall w \in \mathcal{P}_{k+1}^{l,l}(K),
\]

\[
\left( v, p_h^* \right)_K = \left( v, p_h \right)_K, \quad \forall v \in \mathcal{P}_l(K),
\]

where $0 \leq l \leq k$. Here, the space $\mathcal{P}_{k+1}^{l,l}(K)$ denotes the $L^2$-orthogonal complement of $\mathcal{P}_l(K)$. This post-processing method directly uses the definition of the flux $u_h = -\kappa \nabla p_h$ to construct the local problem $u_h$, the approximation of $-\kappa \nabla p$. In practice, the space $\mathcal{P}_{k+1}^{l,l}(K)$ may be constructed using an orthogonal hierarchical basis, and solving (58)–(59) amounts to inverting a local symmetric positive definite system in each cell of the mesh.

At the time of this work, Firedrake does not support the construction of such a finite element basis. However, we can introduce Lagrange multipliers to enforce the orthogonality constraint. The resulting local problem then becomes the following
# Define spaces for the higher-order pressure approximation and Lagrange multipliers
DGk1 = FunctionSpace(mesh, "DG", degree + 1)
DG0 = FunctionSpace(mesh, "DG", 0)
W = DGk1 * DG0
p, psi = TrialFunctions(W)
w, phi = TestFunctions(W)
DG0 = FunctionSpace(mesh, "DG", 0)
DGk1 = FunctionSpace(mesh, "DG", degree + 1)

# Create local Slate tensors for the post-processing system
K = Tensor((inner(grad(p), grad(w)) + inner(psi, w) + inner(p, phi)) * dx)
# Use the computed pressure \( p_h \) and flux \( u_h \) in the right-hand side
F = Tensor((-inner(u_h, grad(w)) + inner(p_h, phi)) * dx)
E = K.inv * F

# Function for the post-processed scalar \( p_h^* \)
p_star = Function(DGk1, name="Post-processed scalar")

assemble(E.blocks[0], p_star) # Assemble only the first field (pressure)

---

**Figure 3.** Example of local post-processing using Firedrake and Slate. Here, we locally solve the mixed system defined in (58)–(59). The corresponding symbolic local tensors are defined in lines 9 and 11. The Slate expression for directly inverting the local system is written in line 12. In line 16, a Slate-generated kernel is produced which solves the resulting linear system in each cell. Since we are not interested in the multiplier, we only return the block corresponding to the new pressure field.

The mixed system: find \((p_h^*, \psi) \in \mathcal{P}_{k+1}(K) \times \mathcal{P}_l(K)\) such that

\[
\left( \nabla w, \nabla p_h^* \right)_{K} + \left( w, \psi \right)_{K} = - \left( \nabla w, \kappa^{-1} u_h \right)_{K}, \quad \forall w \in \mathcal{P}_{k+1}(K),
\]

\[
\left( \phi, p_h^* \right)_{K} = \left( \phi, p_h \right)_{K}, \quad \forall \phi \in \mathcal{P}_l(K),
\]

where \(0 \leq l \leq k\). The local problems (60)–(61) and (58)–(59) are equivalent, with the Lagrange multiplier \( \psi \) enforcing orthogonality of test functions in \( \mathcal{P}_{k+1}(K) \) with functions in \( \mathcal{P}_l(K) \).

This post-processing method produces a new approximation which superconverges at a rate of \( k + 2 \) for hybridized mixed methods (Arnold and Brezzi, 1985; Cockburn et al., 2010a; Stenberg, 1991) (Stenberg, 1991; Cockburn et al., 2010a). For the LDG-H method, \( k + 2 \) superconvergence is achieved when \( \tau = O(1) \) and \( \tau = O(h) \), but only \( k + 1 \) convergence is achieved when \( \tau = O(1/h) \) (Cockburn et al., 2009b, 2010b). We demonstrate the increased accuracy in computed solutions in Section 5.1. An abridged example using Firedrake and Slate to solve the local linear systems is provided in Listing Figure 3.
3.3.2 Post-processing of the flux

Our second example illustrates a procedure that uses the numerical flux of an HDG discretization for (23)–(26). Within the context of the LDG-H method, we can use the numerical trace in (52) to produce a vector field that is $H(\text{div})$-conforming. The technique we outline here follows that of Cockburn et al. (2009b).

Let $T_h$ be a mesh consisting of simplices. On each element $K \in T_h$, we define a new function $u_h^\star$ to be the unique element of the local Raviart-Thomas space $[P_k(K)]^n + xP_k(K)$ satisfying

\[
\begin{align*}
\left\langle r, u_h^\star \right\rangle_K &= \left\langle r, u_h \right\rangle_K, \quad \forall r \in [P_{k-1}(K)]^n, \quad (62) \\
\left\langle \mu, u_h^\star \cdot n \right\rangle_e &= \left\langle \mu, \hat{u} \cdot n \right\rangle_e, \quad \forall \mu \in P_k(e), \forall e \in \partial K. \quad (63)
\end{align*}
\]

for all facets $e$ on $\partial K$, where $\hat{u}$ is the numerical flux defined in (52). This local problem produces a new velocity $u_h^\star$ with the following properties:

1. $u_h^\star$ converges at the same rate as $u_h$ for all choices of $\tau$ producing a solvable system for (54)–(57). However,

2. $u_h^\star \in H(\text{div}; \Omega) \Rightarrow u_h^\star \in H(\text{div}; \Omega)$. That is,

\[
[u_h^\star]_e = 0, \quad \forall e \in E_h^o.
\]

\[
[[u_h^\star]]_e = 0, \quad \forall e \in E_h^r.
\]

3. Additionally, the divergence of $u_h^\star$ converges at a rate of $k + 1$.

The Firedrake implementation using Slate is similar to the scalar post-processing example (see Listing Figure 3); the element-wise cell-wise linear systems (62)–(63) can be expressed in UFL, and therefore the necessary Slate expressions to invert the local systems follows naturally from the set of operations presented in Section 2.1. We use the very sensitive parameter dependency in the post-processing methods to validate our software implementation in Zenodo/Tabula-Rasa (2019).""". The Firedrake implementation using Slate is similar to the scalar post-processing example (see Listing Figure 3); the element-wise cell-wise linear systems (62)–(63) can be expressed in UFL, and therefore the necessary Slate expressions to invert the local systems follows naturally from the set of operations presented in Section 2.1. We use the very sensitive parameter dependency in the post-processing methods to validate our software implementation in Zenodo/Tabula-Rasa (2019)."

4 Static condensation as a preconditioner

Slate enables static condensation approaches to be expressed very concisely. Nonetheless, application of a particular approach to different variational problems using Slate still requires a certain amount of code repetition. By formulating each form of static condensation as a preconditioner, code can be written once and then applied to any mathematically suitable problem. Rather than writing the static condensation by hand, in many cases, it is sufficient to just select the appropriate, Slate-based, preconditioner.

\[\text{This particular post-processing strategy only works on triangles and tetrahedra.}\]
For context, it is helpful to frame the problem in the particular context of the solver library: PETSc. Firedrake uses PETSc to provide linear solvers, and we implement our preconditioners as PETSc-as its main solver abstraction framework, and can provide operator-based preconditioners for solving linear systems as objects. These are defined to act on the problem residual, and return a correction to the solution. Specifically, we expressed in Python via petsc4py (Dalcin et al., 2011). For a comprehensive overview on solving linear systems using PETSc, we refer the interested reader to Balay et al. (2019, Chapter 4).

Suppose we wish to solve a linear system: $Ax = b$. We can think of (left) preconditioning the matrix equation system in residual form:

$$rr = rr(A, b) \equiv b - Ax = 00$$

by an operator $P \mathcal{P}$ (which may not necessarily be linear) as a transformation into an equivalent system of the form

$$PPr = (bPb - Ax)\mathcal{P}Ax = 00.$$  

(65)

Given a current iterate $x_i$, the residual at the $i$-th iteration is simply $r_i \equiv b - Ax_i$, and $PPr_i$ acts on the residual to produce an approximation to the error $e_i \equiv x - x_i$. If $P Pr_i \equiv x - x_i$. If $P$ is an application of an exact inverse, the residual is converted into an exact (up to numerical round-off) error.

We will denote the application of a particular Krylov subspace method (KSP) for the linear system (64) as $Kx(r(A, b))Kx(r(A, b))$. Upon preconditioning the system via $P \mathcal{P}$ as in (65), we write

$$KxKx(rPr(A, b)).$$

(66)

If (66) is solved directly via the application of $A^{-1}$, then $PPr(A, b) = A^{-1}b - x$. So, we have that $KxKx(rPr(A, b)) = KxKx(r(I, A^{-1}b))$ produces $P = A^{-1}$, then $PPr(A, b) = A^{-1}b - x$. So (66) then becomes $KxKx(r(I, A^{-1}b))$, producing the exact solution of (64)

4.1 Interfacing with PETSc via custom preconditioners

The implementation of preconditioners for these systems—the systems considered in this paper requires manipulation not of assembled matrices, but rather their symbolic representation. To do this, we use the preconditioning infrastructure developed by Kirby and Mitchell (2018), which gives preconditioners written in Python access to the symbolic problem description. In Firedrake, this means all derived preconditioners have direct access to the UFL representation of the PDE system. From this mathematical specification, we manipulate this appropriately via Slate and provide operators assembled from Slate expressions to PETSc for further algebraic preconditioning. Using this approach, we have developed a static condensation interface for the hybridization of $H(div) \times L^2$-$H(div) \times L^2$ mixed problems, and a generic interface for statically condensing hybridized finite element systems. The advantage of writing even the latter as a preconditioner is the ability to switch out the solution scheme for the system, even when nested inside a larger set of coupled equations or nonlinear solver (Newton-based methods) at runtime.
4.1.1 A static condensation interface for hybridization

As discussed in sections 3.1 and 3.2, one of the main advantages of using a hybridizable variant of a DG or mixed method is that such systems permit the use of element-wise condensation and recovery cell-wise static condensation. To facilitate this, we provide a PETSc PC static condensation interface, \texttt{firedrake.SCPC}. This preconditioner takes the discretized system as in (44), and performs the local elimination and recovery procedures. Slate expressions are generated from the underlying UFL problem description.

More precisely, the incoming system has the form:

\[
\begin{bmatrix}
A_{e,e} & A_{e,c} \\
A_{c,e} & A_{c,c}
\end{bmatrix}
\begin{bmatrix}
X_e \\
X_c
\end{bmatrix} =
\begin{bmatrix}
R_e \\
R_c
\end{bmatrix},
\]

(67)

where \(X_e = X_{e,c} \) is the vector of unknowns to be eliminated, \(X_c = X_{c,e} \) is the vector of unknowns for the condensed field, and \(R_e, R_c \) are the incoming right-hand sides. The partitioning in (67) is determined by the \texttt{scpc solver} option: \texttt{pc_sc_eliminate_fields}. Field indices are provided in the same way one configures solver options to PETSc. These indices determine which field(s) to statically condense into. For example, on a three-field problem (with indices 0, 1, and 2), setting \texttt{-pc_sc_eliminate_fields 0,1} will configure \texttt{firedrake.SCPC} to cell-wise eliminate field-fields 0 and 1; the resulting condensed system is associated with field 2.

In exact arithmetic, the \texttt{firedrake.SCPC} preconditioner applies the inverse of the can be interpreted as a Schur-complement factorization of method for (67) of the form:

\[
P = \begin{bmatrix}
I & -A_{e,e}^{-1}A_{e,c} \\
0 & I
\end{bmatrix}
\begin{bmatrix}
A_{e,e}^{-1} & 0 \\
0 & S^{-1}
\end{bmatrix}
\begin{bmatrix}
I & 0 \\
-A_{c,e}A_{e,e}^{-1} & I
\end{bmatrix},
\]

(68)

where \(S = A_{c,c} - A_{c,e}A_{e,c}^{-1}A_{e,e} \) and \(S = A_{c,c} - A_{c,e}A_{e,c}^{-1}A_{e,e} \) is the Schur-complement operator for the \(X_e = X_{e,c} \) system. The distinction here from block preconditioners via the PETSc fieldsplit option (Brown et al., 2012), for example, is that \(P \) does not require global actions; by design \(A_{e,e}^{-1}A_{e,c}^{-1} \) can be inverted locally and \(S \) is sparse. As a result, \(S \) can be assembled or applied exactly up to numerical round-off, via Slate-generated kernels.

In practice, the only globally coupled system requiring iterative inversion is \(S \):

\[
K_{X_e}K_{X_e}(P_1 r((SS, R_E R_s))),
\]

(69)

where \(R_E R_s \) is the condensed right-hand side and \(P_1 \) is another possible choice of preconditioner for \(S \). Once \(X_{e,c} = P_1 \) is a preconditioner for \(S \), once \(X_{e,c} \) is computed, \(X_{e,c} \) is reconstructed element-wise via inverting the local systems \(X_e \) is reconstructed by inverting the system \(X_e = A_{e,e}^{-1}(R_e - A_{e,c} X_c) \) cell-wise.

By construction, this preconditioner is suitable for both hybridized mixed and HDG discretizations. It can also be used within other contexts, such as the static condensation of continuous Galerkin discretizations (Guyan, 1965; Irons, 1965) or primal-hybrid methods (Devloo et al., 2018). As with any PETSc preconditioner, solver options can be specified for inverting...
\( S \cdot S \) via the appropriate options prefix (condensed_field). The resulting \texttt{KSP-KSP} for (69) is compatible with existing solvers and external packages provided through the PETSc library. This allows users to experiment with a direct method and then switch to a more parallel-efficient iterative solver without changing the core application code.

### 4.1.2 Preconditioning mixed methods via hybridization

The preconditioner \texttt{firedrake.HybridizationPC} expands on the previous one, this time taking an \( \mathbf{H}(\text{div}) \times L^2 \times \mathbf{H}(\text{div}) \times L^2 \) system and automatically forming the hybridizable problem. This is accomplished through manipulating the UFL objects representing the discretized PDE. This includes replacing argument spaces with their discontinuous counterparts, introducing test functions on an appropriate trace space, and providing operators assembled from Slate expressions in a similar manner as described in section Section 4.1.1.

More precisely, let \( AX = R \) be the incoming mixed saddle point problem, where \( R = \left\{ R_U - R_P \right\}^T \), \( X = \left\{ U^d - P \right\}^T \), and \( U \) and \( P \) are the velocity and scalar unknowns respectively. Then this preconditioner replaces \( AX = R \) with the augmented system: \( AX = R \) with the extended problem:

\[
\begin{bmatrix}
\hat{A} & C^T \\
C & 0
\end{bmatrix}
\begin{bmatrix}
\hat{X} \\
\Lambda
\end{bmatrix}
= \begin{bmatrix}
\hat{R} \\
R_g
\end{bmatrix}
\tag{70}
\]

where \( \hat{R} = \left\{ \hat{R}_U - \hat{R}_P \right\}^T, \hat{R}_U, \hat{R}_P \) are the right-hand sides for the flux and scalar equations respectively, and \( \hat{\cdot} \) indicates modified matrices and co-vectors with discontinuous functions. Here, \( \hat{X} = \left\{ U^d^\cdot - P \right\}^T \) are the hybridized (discontinuous) unknowns to be determined, and \( C\hat{X} = R_g \) is the matrix representation of the transmission condition for the hybridizable mixed method (see (43)).

The preconditioning operator for the hybrid mixed system \textbf{application of \texttt{firedrake.HybridizationPC}} can be interpreted as the Schur-complement reduction of (70) has the form:

\[
\hat{\mathbf{P}} = \begin{bmatrix}
I & -\hat{A}^{-1} C^T \\
0 & I
\end{bmatrix}
\begin{bmatrix}
\hat{\mathbf{A}}^{-1} & 0 \\
0 & S^{-1}
\end{bmatrix}
\begin{bmatrix}
I & 0 \\
-\hat{C}^{-1} \hat{A}^{-1} & I
\end{bmatrix},
\tag{71}
\]

where \( S \) is the Schur-complement matrix \( S = K\hat{A}^{-1} K^T S = -\hat{C} \hat{A}^{-1} C^T \). As before, a single globally coupled system for \( \Lambda \) is required. global system for \( \Lambda \) can be assembled cell-wise using Slate-generated kernels. Configuring the solver for inverted \( \hat{S} \) is done via the PETSc options prefix: \texttt{-hybridization}. The recovery of \( U^d \) and \( P \cdot U^d \) and \( P \) happens in the same manner as \texttt{firedrake.SCPC}.

Since the flux- \textit{hybridizable} flux solution is constructed in a discontinuous space \( U^d \) the “broken” \( \mathbf{H}(\text{div}) \) space \( U^d \), we must project the computed solution into \( U_n \subset \mathbf{H}(\text{div}) U_b \subset \mathbf{H}(\text{div}) \). This can be done cheaply via local facet averaging. The
resulting solution is then updated via \( U \leftarrow \Pi_{\text{div}} U^d \), where \( \Pi_{\text{div}} : U^d_n \rightarrow U^d_h \) is the projection mapping. This ensures the residual for the original mixed problem is properly evaluated to test for solver convergence. With \( \hat{\mathcal{P}} \hat{\mathcal{P}} \) as in (71), the preconditioning operator for the original \( AX = R \) system is: system \( AX = R \) then has the form:

\[
\mathcal{P} = \Pi_{\text{div}} \hat{\mathcal{P}} \Pi^T, \quad \Pi_{\text{div}} = \begin{bmatrix} \Pi_{\text{div}} & 0 & 0 \\ 0 & I & 0 \end{bmatrix}.
\] (72)

We note here that assembly of the right-hand side for the \( A \) side for the \( \mathbf{A} \) system requires special attention. Firstly, when Neumann conditions are present, then \( R_{\mathbf{\Lambda}} \rightarrow R_{\mathbf{h}} \) is not necessarily \( 0 \rightarrow 0 \). Since the hybridization preconditioner has access to the entire Python context (which includes a list of boundary conditions and the spaces in which they are applied), surface integrals on the exterior boundary are added where appropriate and incorporated in the generated Slate expressions. A more subtle issue that requires extra care is the incoming right-hand side tested in \( U_n \) the \( H(\text{div}) \) space \( U_h \).

The situation we are given is that we have \( R_U = R_U(w) \) for \( w \in U^d_n \), \( R_U = R_U(w) \) for \( w \in U^d_h \), but require \( \hat{R}_U(w^d) \) for \( w^d \in U^d_n \hat{R}_U(w^d) \) for \( w^d \in U^d_h \). For consistency, we also require for any \( w \in U^d_n \) that \( w \in U^d_h \) that

\[
\hat{R}_U(w) = R R_U(w).
\] (73)

We can construct such a \( \hat{R}_U \hat{R}_U \) satisfying (73) in the following way. By construction of the space \( U^d_n \), we have for \( \Psi_i \in U^d_n \) each basis function \( \Psi_i \in U^d_h \):

\[
\Psi_i = \begin{cases} 
\Psi_i^d & \text{\( \Psi_i \) associated with an exterior facet node}, \\
\Psi_i^{d,+} + \Psi_i^{d,-} & \text{\( \Psi_i \) associated with an interior facet node}, \\
\Psi_i^d & \text{\( \Psi_i \) associated with a cell interior node},
\end{cases}
\] (74)

where \( \Psi_i^d, \Psi_i^{d,+}, \Psi_i^{d,-} \in U^d_n \), \( \Psi_i^d, \Psi_i^{d,+} \in U^d_h \), and \( \Psi_i^{d,+} \) and \( \Psi_i^{d,-} \) are basis functions corresponding to the positive and negative restrictions associated with the \( i \)-th facet node. We then define our “broken” right-hand side via the local definition:

\[
\hat{R}_U(\Psi_i^d) = \frac{R_U(\Psi_i^d)}{N_i} \frac{R_U(\Psi_i^d)}{N_i},
\] (75)

where \( N_i \) is the number of cells that the degree of freedom corresponding to the basis function \( \Psi_i \in U^d_n \) touches \( \Psi_j \in U^d_h \) is topologically associated with. Using (74), (75), and the fact that \( R_U R_U \) is linear in its argument, we can verify that our construction of \( \hat{R}_U \hat{R}_U \) satisfies (73).

---

4 These are the two “broken” parts of \( \Psi_i^d \) on a particular facet connecting two elements. That is, for two adjacent cells, a basis function in \( U^d_n \) for a particular facet node can be decomposed into two basis functions in \( U^d_h \) defined on their respective sides of the facet.

4 These are the two “broken” parts of \( \Psi_i^d \) on a particular facet connecting two elements. That is, for two adjacent cells, a basis function in \( U^d_h \) for a particular facet node can be decomposed into two basis functions in \( U^d_h \) defined on their respective sides of the facet.
5 Numerical studies

We now present results utilizing the Slate DSL and our static condensation preconditioners for a set of test problems. Since we are using the interfaces outlined in Section 4, Slate is accessed indirectly and requires no manually-written solver code for hybridization or static condensation/local recovery. All parallel results were obtained on a single fully-loaded compute node of dual-socket Intel E5-2630v4 (Xeon) processors with $2 \times 10$ cores (2 threads per core) running at 2.2GHz. In order to avoid potential memory effects due to the operating system migrating processes between sockets, we pin MPI processes to cores.

Verification of the generated code is performed using parameter-sensitive convergence tests. The study consists of running a variety of discretizations spanning the methods outlined in Section 3. Details and numerical results are made public and can be viewed in Zenodo/Tabula-Rasa (2019) (see “Code and data availability”). All results are in full agreement with the theory.

5.1 LDG-H-HDG method for a three-dimensional elliptic equation

In this section, we take a closer look at the LDG-H method for the model elliptic equation (sign-definite Helmholtz):

\[-\nabla \cdot \nabla p + p = f, \quad \text{in } \Omega = [0, 1]^3, \tag{76}\]

\[p = g, \quad \text{on } \partial \Omega, \tag{77}\]

where $f$ and $g$ are chosen such that the analytic solution is $p = \exp\{\sin(\pi x) \sin(\pi y) \sin(\pi z)\}$. We use a regular mesh consisting of $6 \cdot N^3$ tetrahedral elements ($N \in \{4, 8, 16, 32, 64\}$). First, we reformulate (76)–(77) as the mixed problem:

\[\mathbf{u} + \nabla p = 0, \tag{78}\]

\[\nabla \cdot \mathbf{u} + p = f, \tag{79}\]

\[p = g, \quad \text{on } \partial \Omega. \tag{80}\]

We start with linear polynomial approximations, up to cubic, for the LDG-H discretization of (78)–(80). Additionally, we compute a post-processed scalar approximation $p_h^\ast$ of the HDG solution. This raises the approximation order of the computed solution by an additional degree. In all numerical studies here, we set the HDG parameter $\tau = 1$. All results were computed in parallel, utilizing a single compute node (described previously).

A continuous Galerkin (CG) discretization of the primal problem (76)–(77) serves as a reference for this experiment. Due to the superconvergence in the post-processed solution for the HDG method, we use CG discretizations of polynomial order 2, 3, and 4. This takes into account the enhanced accuracy of the HDG solution, despite being initially computed as a lower-order approximation. We therefore expect both methods to produce equally accurate solutions to the model problem.

Our aim here is not to compare the performance of HDG and CG, which has been investigated elsewhere (for example, see Kirby et al. (2012); Yakovlev et al. (2016)). Instead, we provide a reference that the reader might be more familiar with in order to evaluate whether our software framework produces a sufficiently performant HDG implementation relative to what might be expected.
For the CG discretization, we invert the CG system with a matrix-explicit iterative solver consisting of the conjugate gradient method preconditioned conjugate gradient solver with hypre’s boomerAMG implementation of algebraic multigrid (AMG) [Falgout et al. (2006)]. We use a preconditioner (Falgout et al., 2006). For the HDG method, we use the preconditioner described in Section 4.1.1 to statically condense the LDG-H system, using the same iterative solver and the same solver setup as the CG method for the Lagrange multipliers. This is sensible, as the global trace operator defines a symmetric trace system. While indeed the trace operator is symmetric and positive-definite, one should keep in mind that conclusions regarding the performance of off-the-shelf AMG packages on the HDG trace system is still relatively unclear. As a result, efforts on developing more efficient multigrid strategies is a topic of on-going interest (Cockburn et al., 2014; Kronbichler and Wall, 2018).

To avoid over-solving, we iterate to a relative tolerance such that the discretization error is minimal for a given mesh. In other words, the solvers are configured to terminate when there is no further reduction in the $L^2$-error of the computed solution compared with the analytic solution. This means we are not iterating to a fixed solver tolerance across all mesh resolutions. Therefore, we can expect the total number of Krylov iterations (for both the CG and HDG methods) to increase as the mesh resolution becomes finer. The rationale behind this approach is to directly compare the execution time to solve for the best possible approximation to the solution given a fixed resolution.

### 5.1.1 Error versus execution time

The total execution time is recorded for the CG and HDG solvers, which includes the setup time for the AMG preconditioner, matrix assembly, and the time-to-solution for the Krylov method. In the HDG case, we include all setup costs, the time spent building the Schur-complement for the traces, local recovery of the scalar and flux approximations, and post-processing. The $L^2$-error against execution time and Krylov iterations to reach discretization error for each mesh are summarized in Figure 4.

The HDG method of order $k-1$ ($H DG_{k-1}$) with post-processing, as expected, produces a solution which is as accurate as the CG method of order $k$ ($C G_k$). While the full HDG system is never explicitly assembled, the larger execution time is a result of several factors. The primary factor is that the total number of trace unknowns for the $H DG_1$, $H DG_2$, and $H DG_3$ discretizations is roughly four, three, and two times larger (resp.) than the corresponding number of CG unknowns. Therefore, each iteration is more expensive. Moreover, we also observe that the trace system requires more Krylov iterations to reach discretization error. The gap in total number of iterations starts to close, which appears to improve relative to the CG method as the approximation degree increases (see Figure 5). The extra cost of HDG due to the larger degree of freedom count and the need to perform local tensor inversion is offset by the local conservation order increases. Further analysis on multigrid methods for HDG systems is required to draw further conclusions. The main computational bottleneck in HDG methods is the global linear solver. We therefore expect our implementation to be dominated by the cost associated with inverting the trace operator. If one considers just the time-to-solution, the CG method is clearly ahead of the HDG method. However, the superior scaling, local conservation, and stabilization properties which are useful of the HDG method make it a
Figure 4. Error Comparison of continuous Galerkin and LDG-H solvers for the model three-dimensional positive-definite Helmholtz equation. Left: a log-log plot showing the error against execution time for the CG and HDG with post-processing ($\tau = 1$) methods. Right: A log-linear plot showing Krylov iterations of the AMG-preconditioned conjugate gradient algorithm (to reach discretization error) against number of cells.

Krylov iterations of the AMG-preconditioned conjugate gradient algorithm (to reach discretization error) against number of cells.

Comparison of continuous Galerkin and LDG-H solvers for the model three-dimensional positive-definite Helmholtz equation.

particularly appealing choice for fluid dynamics applications (Yakovlev et al., 2016; Kronbichler and Wall, 2018). Therefore, the development of good preconditioning strategies for the HDG method is critical for its competitive use.

5.1.2 Break down of solver time

The HDG method requires many more degrees of freedom than CG or primal DG methods. This is largely due to the fact that the HDG method simultaneously approximates the primal solution and its velocity. The global matrix for the traces is significantly larger than the one for the CG system at low polynomial order. The execution time for HDG is then compounded by a more expensive global solve. We remind the reader that our goal here is to verify that the solve time for HDG is as might be expected given the problem size.

Figure 5 displays the a break down of total execution times on a simplicial mesh consisting of 1.5 million elements. The execution times have been normalized by the CG total time, showing that the HDG method is roughly 3 times the execution time of more expensive than the CG method. This is expected given the larger degree-of-freedom count and expensive global solve. The raw numerical breakdown of the HDG and CG solvers are shown in Table 1. We isolate each component of the HDG method contributing to the total execution time. Local operations include static condensation (trace operator assembly), forward elimination (right-hand side assembly for the trace system), backwards substitution to recover the scalar and velocity
Table 1. Breakdown of the raw timings for the $H DG_{k-1}$ ($\tau = 1$) and $CG_k$ methods, $k = 2, 3,$ and $4$. Each method corresponds to a mesh size $N = 64$ on a fully-loaded compute node.

<table>
<thead>
<tr>
<th>Stage</th>
<th>$H DG_1$</th>
<th>$H DG_2$</th>
<th>$H DG_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t_{stage}$ (s)</td>
<td>% $t_{total}$</td>
<td>$t_{stage}$ (s)</td>
</tr>
<tr>
<td>Matrix assembly (static cond.)</td>
<td>1.05</td>
<td>7.49 %</td>
<td>6.95</td>
</tr>
<tr>
<td>Forward elimination</td>
<td>0.86</td>
<td>6.13 %</td>
<td>6.32</td>
</tr>
<tr>
<td>Trace solve</td>
<td>10.66</td>
<td>76.24 %</td>
<td>43.89</td>
</tr>
<tr>
<td>Back substitution</td>
<td>1.16</td>
<td>8.28 %</td>
<td>8.71</td>
</tr>
<tr>
<td>Post processing</td>
<td>0.26</td>
<td>1.86 %</td>
<td>0.98</td>
</tr>
<tr>
<td>HDG Total</td>
<td>13.98</td>
<td>66.85 %</td>
<td>308.37</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$CG_2$</th>
<th>$CG_3$</th>
<th>$CG_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{stage}$ (s)</td>
<td>% $t_{total}$</td>
<td>% $t_{total}$</td>
<td>% $t_{total}$</td>
</tr>
<tr>
<td>Matrix assembly (monolithic)</td>
<td>0.50</td>
<td>12.01 %</td>
<td>2.91</td>
</tr>
<tr>
<td>Solve</td>
<td>3.63</td>
<td>87.99 %</td>
<td>22.67</td>
</tr>
<tr>
<td>CG Total</td>
<td>4.12</td>
<td>25.59 %</td>
<td>109.36</td>
</tr>
</tbody>
</table>

Figure 5. Break down of the $CG_k$ and $H DG_{k-1}$ execution times on a $6 \cdot 64^3$ simplicial mesh.
unknowns, and local post-processing of the primal scalar solution. For all $k$, our HDG implementation is solver-dominated as expected.

Both trace operator and right-hand side assembly are dominated by the costs of inverting a local square mixed matrix coupling the primal and dual variables scalar and velocity unknowns, which is performed directly via an LU factorization. They should This is also the case for backwards substitution. They should all therefore be of the same magnitude in time spent. We observe that this is the case across all degrees, with times ranging between approximately 6—11% of total execution time for both local operations assembling the condensed system. Back-substitution takes roughly the same time as the static condensation and forward elimination stages (between 8—15 approximately 12% of execution time across all $k$). This is expected, as these operations are all dominated by the cost of inverting the local matrix coupling the scalar and velocity degrees of freedom. The slight increase in time is due splitting the local equations into two local solvers: one for $p_h$ and another for the velocity, on average. Finally, the additional cost of post-processing accrues negligible time (roughly 2% of execution time across all degrees). This is a small cost for an increase in order of accuracy.

We note that caching of local tensors does not occur. Each pass to perform the local eliminations and backwards reconstructions rebuilds the local element tensors. It is not clear at this time whether the performance gained from avoiding rebuilding the local operators will offset the memory costs of storing the local matrices. Moreover, in time-dependent problems where the operators may contain state-dependent variables, rebuilding local matrices will be necessary in each time-step regardless.

5.2 Hybridized-mixed-solver Hybridizable mixed methods for the shallow water equations

A primary motivator for our interest in hybridized hybridizable methods revolves around developing efficient solvers for problems in geophysical flows. In this section, we present some results resulting integrating the nonlinear, rotating shallow water equations on the sphere using test case 5 (flow past an isolated mountain) from Williamson et al. (1992). For our discretization approach, we use the framework of compatible finite elements (Cotter and Shipton, 2012; Cotter and Thuburn, 2014).

5.2.1 Semi-implicit finite element discretization

We start with The model equations we consider are the vector-invariant rotating nonlinear shallow water system defined on a two-dimensional spherical surface $\Omega$ embedded in $\mathbb{R}^3$:

\[
\frac{\partial \mathbf{u}}{\partial t} + (\nabla \cdot \mathbf{u} + f) \mathbf{u} + \nabla \left( g(D + b) + \frac{1}{2} |\mathbf{u}|^2 \right) = 0, \tag{81}
\]

\[
\frac{\partial D}{\partial t} + \nabla \cdot (\mathbf{u} D) = 0, \tag{82}
\]

where $\mathbf{u}$ is the fluid velocity, $D$ is the depth field, $f$ is the Coriolis parameter, $g$ is the acceleration due to gravity, $b$ is the bottom topography, and $(\cdot)^\perp \equiv \mathbf{z} \times \cdot$, with $\mathbf{z} \equiv \mathbf{z} \times \cdot$, with $\mathbf{z}$ being the unit normal to the surface $\Omega$.

After discretizing in space and time using a semi-implicit scheme and Picard linearization, following Natale and Cotter (2017) and based on Melvin et al. (2019) (see Appendix A for a summary of the entire solution strategy) Natale et al. (2016).
Table 2. The number of unknowns to be determined are summarized for each compatible finite element method. Resolution is the same for both methods.

<table>
<thead>
<tr>
<th>Mixed method</th>
<th># cells</th>
<th>$\Delta x$</th>
<th>Velocity unknowns</th>
<th>Depth unknowns</th>
<th>Total (millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$RT_1 \times DG_0$</td>
<td>327,680</td>
<td>$\approx 43$ km</td>
<td>491,520</td>
<td>327,680</td>
<td>0.8 M</td>
</tr>
<tr>
<td>$BDM_2 \times DG_1$</td>
<td></td>
<td></td>
<td>2,457,600</td>
<td>983,040</td>
<td>3.4 M</td>
</tr>
</tbody>
</table>

, we must solve an indefinite a sequence of saddle point system at each step time step of the form:

\[
\begin{bmatrix}
A & -g \frac{\Delta t}{2} B^T \\
H \frac{\Delta t}{2} B & M
\end{bmatrix}
\begin{bmatrix}
\Delta U \\
\Delta D
\end{bmatrix}
= \begin{bmatrix}
R_u \\
R_D
\end{bmatrix},
\]

(83)

The approximations $\Delta U$ and $\Delta D$ See Appendix A for a complete description of the entire discretization strategy. The system (83) is the matrix equation corresponding to the linearized equations in (A3)–(A4).

The Picard updates $\Delta U$ and $\Delta D$ are sought in the mixed finite element spaces $U_h \subset H(\text{div})$, $U_h \subset H(\text{div})$ and $V_h \subset L^2$ respectively. Pairings including the RT or BDM spaces such as Stable mixed finite element pairings correspond to the well-known RT and BDM mixed methods, such as: RT $k \times DG_{k-1}$ or BDM $k \times DG_{k-1}$. These also fall within the set of compatible mixed spaces ideal for geophysical fluids fluid dynamics ( Cotter and Shipton, 2012; Natale et al., 2016; Melvin et al., 2019). In particular, the lowest-order RT method ($RT_1 \times DG_0$) on a structured quadrilateral grid (such as the latitude-longitude grid used many operational dynamical cores) corresponds is analogous to the Arakawa C-grid finite difference discretization.

In staggered finite difference models, the standard approach for solving (83) is to neglect the Coriolis term and eliminate the velocity unknowns $\Delta U$ unknown $\Delta U$ to obtain a discrete elliptic problem for which equation for $\Delta D$, where smoothers like Richardson iterations or relaxation methods are convergent. This is more problematic in the compatible finite element framework, since $A$ has a dense inverse. Instead, we use the preconditioner described in Section 4.1.2 to form the hybridized problem and eliminate both $\Delta U$ and $\Delta D$ locally-equivalent hybridizable formulation, where both $\Delta U$ and $\Delta D$ are eliminated locally to produce a sparse elliptic equation for the Lagrange multipliers.

5.2.1 Atmospheric flow over a mountain

As a test problem, we solve test case 5 of Williamson et al. (1992), on the surface of an Earth-sized sphere a sphere with radius $R = 6371$ km. We refer the reader to Cotter and Shipton (2012); Shipton et al. (2018) for a more comprehensive study on mixed finite elements for shallow water systems of this type. We use the mixed finite element pairs ($RT_1 , DG_0$) (lowest-order RT method) and ($BDM_2 , DG_1$) (next-to-lowest order BDM method) for the velocity and depth spaces. The sphere mesh A mesh of the sphere is generated from 7 refinements of an icosahedron, resulting in a triangulation $T_h$ consisting of 327,680 elements in total. The discretization information is grid information for both mixed methods are summarized in Table 2.
We run for a total of 25 time-steps, with a fixed number of 4 Picard iterations in each time-step. We compare the overall simulation time using two different solver configurations for the implicit linear system. First, we use a flexible variant of GMRES\(^5\) acting on the outer system (83) with an approximate Schur complement preconditioner:

\[
P_{SC} = \begin{bmatrix} I & g \frac{\Delta t}{2} A^{-1} B^T \\ 0 & I \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & \tilde{S}^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -H \frac{\Delta t}{2} BA^{-1} & I \end{bmatrix},
\]

where \(\tilde{S} = \tilde{D} - C \text{diag}(A) - B\), and \(\text{diag}(A) \tilde{S} = M + gH \frac{\Delta t}{2} B \text{diag}(A)^{-1} B\), and \(\text{diag}(A)\) is a diagonal approximation to the velocity mass matrix (plus the addition of a Coriolis matrix). The Schur-complement system is inverted via GMRES due to the asymmetry from the Coriolis term, with the inverse of \(\tilde{S}^{-1}\) as the preconditioning operator. The sparse approximation \(\tilde{S}^{-1}\) is inverted using PETSc’s smoothed aggregation multigrid (GAMG). The Krylov method is set to terminate once the preconditioned residual norm is reduced by a factor of 10\(^8\). \(A^{-1} A^{-1}\) is computed approximately using a single application of incomplete LU (zero fill-in).

Next, we use only the application of our hybridization preconditioner (no outer iterations Krylov method), which replaces the original linearized mixed system with its hybrid mixed equivalent. After hybridization, we have the problem: find \((\Delta u_h^d, \Delta D_h, \lambda_h) \in U_h^d \times V_h \times M_h\) such that the following extended problem for the Picard updates: find \((\Delta u_h^d, \Delta D_h, \lambda_h) \in U_h^d \times V_h \times M_h\) satisfying

\[
\begin{aligned}
\left( w, \Delta u_h^d \right)_{T_h} + \frac{\Delta t}{2} \left( w, f(\Delta u_h^d) \right)_{T_h} - \frac{\Delta t}{2} \left( \nabla \cdot w, g \Delta D_h \right)_{T_h} + \left\langle [w], \lambda_h \right\rangle_{\partial T_h} &= \hat{R}_u, \quad \forall w \in U_h^d, \\
\left( \phi, \Delta D_h \right)_{T_h} + \frac{\Delta t}{2} \left( \phi, H \nabla \cdot \Delta u_h^d \right)_{T_h} &= R_D, \quad \forall \phi \in V_h, \\
\left\langle \gamma, \left[ \Delta u_h^d \right] \right\rangle_{\partial T_h} &= 0, \quad \forall \gamma \in M_h.
\end{aligned}
\]

Note that the space \(M_h\) is chosen such that the trace functions, when restricted to a facet \(e \in \partial T_h\), are from the same polynomial space as \(\Delta u_h^d\) restricted to that same facet. Additionally \(\Delta u_h \cdot n\rceil_e\). Moreover, it can be shown that the Lagrange multiplier \(\lambda_h\) is an approximation to the depth unknown \(\Delta t g \Delta D/2\) restricted to \(\partial T_h\).

The resulting three-field problem for in (85)–(87) has the matrix form: produces the following matrix equation:

\[
\begin{bmatrix} \tilde{A} & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} \Delta X \\ \Lambda \end{bmatrix} = \begin{bmatrix} \hat{R}_{\Delta X} \\ 0 \end{bmatrix},
\]

where \(\tilde{A}\) is the discontinuous operator coupling \(\Delta X = \left\{ \Delta U^d, \Delta D \right\}^T\), and \(R_{\Delta X} = \left\{ \hat{R}_u, R_D \right\}^T\) \(\Delta X = \left\{ \Delta U^d, \Delta D \right\}^T\) and \(R_{\Delta X} = \left\{ \hat{R}_u, R_D \right\}^T\) are the problem residuals. An exact Schur-complement factorization is performed on (88), using Slate

\(^5\)We use a flexible version of GMRES on the outer system since we use an additional Krylov solver to iteratively invert the Schur-complement.
Table 3. Preconditioner solve times for a 25-step run with $\Delta t = 100s$. These are cumulative times in each stage of the two preconditioners throughout the entire profile run. We display the average iteration count (rounded to the nearest integer) for both the outer and the inner Krylov solvers. The significant speedup when using hybridization is a direct result of eliminating the outer-most solver.

<table>
<thead>
<tr>
<th>Mixed method Preconditioner</th>
<th>$t_{total}$ (s)</th>
<th>Avg. outer its.</th>
<th>Avg. inner its.</th>
<th>$\frac{t_{sc}}{t_{total}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT$<em>1 \times$ DG$<em>0$ approx. Schur. ($P</em>{sc}P</em>{SC}$)</td>
<td>15.137</td>
<td>2</td>
<td>8</td>
<td>3.413</td>
</tr>
<tr>
<td>hybridization ($P_{hybrid}$)</td>
<td>4.434</td>
<td>None</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>BDM$<em>2 \times$ DG$<em>1$ approx. Schur. ($P</em>{sc}P</em>{SC}$)</td>
<td>300.101</td>
<td>4</td>
<td>9</td>
<td>5.556</td>
</tr>
<tr>
<td>hybridization ($P_{hybrid}$)</td>
<td>54.013</td>
<td>None</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Breakdown of the cost (average) of a single application of the preconditioned flexible GMRES algorithm and hybridization preconditioner. Hybridization takes approximately the same time per iteration.

<table>
<thead>
<tr>
<th>Preconditioner</th>
<th>Stage</th>
<th>RT$_1 \times$ DG$_0$</th>
<th>BDM$_2 \times$ DG$_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>approx. Schur ($P_{SC}$)</td>
<td>Schur solve</td>
<td>0.07592</td>
<td>0.78405</td>
</tr>
<tr>
<td></td>
<td>mass. $A$ operator: $A$</td>
<td>0.00032</td>
<td>0.00678</td>
</tr>
<tr>
<td></td>
<td>apply inverse: $A^{-1}A^{-1}$</td>
<td>0.00041</td>
<td>0.00703</td>
</tr>
<tr>
<td></td>
<td>gmres other</td>
<td>0.00652</td>
<td>0.04041</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.08317</td>
<td>0.83827</td>
</tr>
<tr>
<td>hybridization ($P_{hybrid}$)</td>
<td>Transfer: $\mathcal{R}<em>{\Delta X \to \hat{\mathcal{R}}</em>{\Delta X}}$</td>
<td>0.00322</td>
<td>0.00597</td>
</tr>
<tr>
<td></td>
<td>Forward elim.: $-K\hat{A}^{-1} \hat{R}<em>{\Delta X} - C\hat{A}^{-1} \hat{R}</em>{\Delta X}$</td>
<td>0.00561</td>
<td>0.12308</td>
</tr>
<tr>
<td></td>
<td>Trace solve</td>
<td>0.02289</td>
<td>0.28336</td>
</tr>
<tr>
<td></td>
<td>Back sub.</td>
<td>0.00986</td>
<td>0.12220</td>
</tr>
<tr>
<td></td>
<td>Projection: $\mathcal{H}<em>{div} \Delta \hat{U} \leftarrow \Pi</em>{div} \Delta U_d$</td>
<td>0.00264</td>
<td>0.00516</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.04434</td>
<td>0.54013</td>
</tr>
</tbody>
</table>

to generate the local elimination kernels. We use the same set of solver options for the inversion of $\tilde{S}^{-1}\tilde{S}$ in (84) to invert the Lagrange multiplier system. The increments $\Delta U_d$ and $\Delta D$ are recovered locally, using Slate-generated kernels. Once recovery is complete, $\Delta U_d$ is injected $\Delta U_d$ is projected back into the conforming $\mathcal{H}(div)-\mathcal{H}(div)$ finite element space via $\Delta U \leftarrow \Pi_{div} \Delta U_d$. Based on the discussion in section 4.1.2, we apply:

$$P_{\text{hybrid}} = \Pi \Pi^{T} \begin{bmatrix} I & -\hat{A}^{-1}C^T \hat{A}^{-1} & 0 \\ 0 & I & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & S^{-1} \end{bmatrix} \begin{bmatrix} I \\ -C\hat{A}^{-1} \\ I \end{bmatrix} \Pi \Pi^{T}. \quad (89)$$
Table 3 displays a summary of our findings. The advantages of a hybridizable method versus a mixed method are more clearly realized in this experiment. When using hybridization, we observe a significant reduction in time spent during the implicit solve stage in the implicit solver compared to the approximate Schur-complement approach. This is primarily because we have reduced the number of required “outer” iterations to zero; the hybridization preconditioner is performing an exact factorization of the global hybridized operator.

This is empirically supported when considering the per iteration solve times, summarized per-application solve times. The values reported in Table 4 show the average cost of a single outer GMRES iteration (which includes the application of $P_{SC}$) and a single application of $P_{hybrid}$. Hybridization and the approximate Schur complement preconditioner are comparable in terms of average execution time, with hybridization being slightly faster per application. This further demonstrates that the primary cause for the longer execution time of the latter is a direct result of the additional outer iterations induced from using an approximate factorization to achieve the same reduction in the overall residual. In terms of over all time-to-solution, the hybridizable methods are clearly ahead of the original mixed methods.

We also measure the relative reductions in the true residual problem residual of the linear system (83). Our hybridization preconditioner reduces the residual by a factor of $10^8$ on average, which coincides with the specified relative tolerance for the Krylov method on the trace system. Snapshots of a (coarser) In other words, the reduction in the residual for the trace system translates to an overall reduction in the residual for the mixed system by the same factor.

The test case was run up to day 15 on a coarser resolution (20,480 simplicial cells with $\Delta x \approx 210$ km) and a time-step size $\Delta t = 500$ seconds. Snapshots of the entire simulation are provided in Figure 6 using the semi-implicit scheme described in this paper Appendix A. The results we have obtained for days 5, 10, and 15 are comparable to the corresponding results of Nair et al. (2005); Ullrich et al. (2010); Kang et al. (2019). We refer the reader to Shipton et al. (2018) for an exposition further demonstrations of shallow water test cases featuring the use of a hybridized implicit solver (as the hybridization preconditioner described in Section 4.1.2).

5.3 Rotating-Hybridizable methods for a linear Boussinesq model

As a final example, we consider the simplified atmospheric model obtained from a linearization of the compressible Boussinesq equations in a rotating domain:

$$\frac{\partial \mathbf{u}}{\partial t} + 2\Omega \times \mathbf{u} = -\nabla p + b\hat{z},$$  
$$\frac{\partial p}{\partial t} = -c^2 \nabla \cdot \mathbf{u},$$  
$$\frac{\partial b}{\partial t} = -N^2 \mathbf{u} \cdot \hat{z},$$

where $\mathbf{u}$ is the fluid velocity, $p$ the pressure, $b$ is the buoyancy, $\Omega$ the planetary angular rotation vector, $c$ is the speed of sound ($\approx 343\text{ms}^{-1}$), and $N$ is the buoyancy frequency ($\approx 0.01\text{s}^{-1}$). Equations (90)–(92) permit fast-moving acoustic waves driven by perturbations in $b$. This is the same model presented in Skamarock and Klemp (1994), which uses a quadratic equation of
state to avoid some of the complications of the full compressible Euler equations (the hybridization of which we shall address in future work). We solve these equations subject to the rigid-lid condition $u \cdot n = 0$ on all boundaries.

Our domain consists of a spherical annulus, with the mesh constructed from a horizontal “base” mesh of the surface of a sphere of radius $R$, extruded upwards by a height $H$. The vertical discretization is a structured one-dimensional grid, which facilitates the staggering of thermodynamic variables, such as $b$. We consider two kinds of meshes: one obtained by extruding an icosahedral sphere mesh, and another from a cubed sphere.

Since our mesh has a natural tensor product tensor product structure, we construct suitable finite element spaces constructed by taking the tensor product of a horizontal space with a vertical space. To ensure our discretization is “compatible,” we use the one- and two-dimensional finite element de-Rham complexes: $V_0 \overset{\partial}{\rightarrow} V_1 \rightarrow U_0 \overset{\nabla}{\rightarrow} U_1 \rightarrow U_2 \rightarrow V_h^0 \overset{\partial}{\rightarrow} V_h^1 \rightarrow U_h^0 \overset{\nabla}{\rightarrow} U_h^1 \rightarrow U_h^2$.

![Figure 6. Snapshots (view from the northern pole) from the isolated mountain test case. The surface height (m) at days 5, 10, and 15. The snapshots were generated on a mesh with 20,480 simplicial cells, a BDM$_2 \times$ DG$_1$ discretization, and $\Delta t = 500$ seconds. The linear system during each Picard cycle was solved using the hybridization preconditioner.](image-url)
Table 5. Vertical and horizontal spaces for the three-dimensional compatible finite element discretization of the linear Boussinesq model. The RT\(_k\) and BDFM\(_{k+1}\) methods are constructed on triangular prism elements, while the RTCF\(_k\) method is defined on extruded quadrilateral elements.

<table>
<thead>
<tr>
<th>Mixed method</th>
<th>(V_0)-(V_0)</th>
<th>(V_0)-(V_1)</th>
<th>(U_0)-(U_0)</th>
<th>(U_1)-(U_1)</th>
<th>(U_2)-(U_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT(_k)</td>
<td>CG(_k)([0, H_2])</td>
<td>DG(_{k-1})([0, H_2])</td>
<td>CG(_k)((\triangle))</td>
<td>RT(_k)((\triangle))</td>
<td>DG(_{k-1})((\triangle))</td>
</tr>
<tr>
<td>BDFM(_{k+1})</td>
<td>CG(_{k+1})(([0, H_2]))</td>
<td>DG(_k)(([0, H_2]))</td>
<td>CG(_{k+1})((\triangle))</td>
<td>BDFM(_{k+1})((\triangle))</td>
<td>DG(_k)((\triangle))</td>
</tr>
<tr>
<td>RTCF(_k)</td>
<td>CG(_k)(([0, H_2]))</td>
<td>DG(_{k-1})([0, H_2])</td>
<td>Q(_k)((\square))</td>
<td>RTCF(_k)((\square))</td>
<td>DG(_{k-1})((\square))</td>
</tr>
</tbody>
</table>

We can then construct the three-dimensional complex:

\[
W_0 \xrightarrow{\nabla} W_1 \xrightarrow{\nabla} W_2 \xrightarrow{\nabla} W_3, \quad \text{where} \quad W_0 \subseteq W_1 \subseteq W_2 \subseteq W_3, \quad \text{where}
\]

\[
W_0 = U_0 \otimes V_0, \quad W_1 = H\text{Curl}(U_1 \otimes V_0) \oplus H\text{Curl}(U_0 \otimes V_1), \quad W_2 = H\text{Div}(U_2 \otimes V_0) \oplus H\text{Div}(U_1 \otimes V_2), \quad W_3 = U_2 \otimes V_2
\]

5 \hspace{1cm} W_2^3_h = U_2 \otimes V_1^h. \tag{96}

Here, H\text{Curl} and H\text{Div} denote operators which ensure the correct Piola transformations are applied when mapping from physical to reference element. We refer the reader to (McRae et al., 2016; McRae et al., 2016) for an overview of constructing tensor product finite element spaces in Firedrake. For the analysis of compatible finite element discretizations and their relation to the complex (93)–(96), we refer the reader to Natale et al. (2016); Cotter and Thuburn (2014); Natale et al. (2016). Each discretization used in this section is constructed from more familiar finite element families, shown in Table 5.

5.3.1 Compatible finite element discretization

We seek the unknown fields from:

A compatible finite element discretization of (90)–(92) constructs solutions in the following finite element spaces:

\[
u_h \in W_0^h \oplus W_2^h, \quad p_h \in W_3^h, \quad b_h \in W_2^h, \quad \text{where } W_2^h \subseteq W_2^h \text{ is the subspace of } W_2 \text{ satisfying the no-slip condition, and } W_0 = U_2 \otimes V_0 \oplus H(\text{div}) \text{ whose functions } w \text{ satisfy } w \cdot n = 0 \text{ on } \partial \Omega, W_3^h \subseteq L^2, \text{ and } W_2^h \equiv U_2 \otimes V_0^0. \text{ Note that } W_2^h \text{ is just the vertical part of the velocity space.}^6
\]

\(^6\) The choice for \(W_2\) in corresponds to a Charney-Phillips vertical staggering of the buoyancy variable, which is the desired approach for the UK Met Office’s Unified Model (Melvin et al., 2010). One could also collocate \(b\) with \(p\) (\(b \in W_2\)), which corresponds to a Lorenz staggering. This however supports a computational mode which is exacerbated by fast moving waves. We restrict our discussion to the former case.
version of the vertical velocity space.\footnote{The choice for $W_h^b$ in (97) corresponds to a Charney-Phillips vertical staggering of the buoyancy variable, which is the desired approach for the UK Met Office’s Unified Model (Melvin et al., 2010). One could also collocate $b_h$ with $p_h (b_h \in W_h^b)$, which corresponds to a Lorenz staggering. This however supports a computational mode which is exacerbated by fast-moving waves. We restrict our discussion to the former case.} That is, $W_h$, $W_h^v$, $W_h^b$, and $W_h^{2,v}$ have the same number of degrees of freedom, but differ in how they are pulled back to the reference element. For ease of notation, we write $W_2$ in place of $W_h^2$.

To obtain the discrete system, we simply multiply equations (90)–(92) by test functions $w \in W_2$, $\phi \in W_3$, and $\eta \in W_1$ and integrate by parts. We introduce the increments $\delta u = u^{(n+1)} - u^{(n)}$, $\delta u_h = u_h^{(n+1)} - u_h^{(n)}$, and $u_0 = u(0)$. Using an implicit midpoint rule discretization, we need to solve the following linear variational-mixed problem at each time-step: find $\delta u \in W_2$, $\delta p \in W_3$, and $\delta b \in W_1$ such that

\[
\left( w, \delta u_h \right)_{\mathcal{T}_h} + \frac{{\Delta}t}{2} \left( w, 2\Omega \times \delta u_h \right)_{\mathcal{T}_h} - \frac{{\Delta}t}{2} \left( \nabla \cdot w, \delta p \right)_{\mathcal{T}_h} - \frac{\Delta t}{2} \left( w, \delta b \cdot \hat{z} \right)_{\mathcal{T}_h} = r_u, \quad \forall w \in \tilde{W}_h^2
\]

\[
\left( \phi, \delta p \right)_{\mathcal{T}_h} + \frac{\Delta t}{2} c^2 \left( \phi, \nabla \cdot \delta u_h \right)_{\mathcal{T}_h} = r_p, \quad \forall \phi \in \tilde{W}_h^3
\]

\[
\left( \eta, \delta b \right)_{\mathcal{T}_h} + \frac{\Delta t}{2} N^2 \left( \eta, \delta u_h \cdot \hat{z} \right)_{\mathcal{T}_h} = r_b, \quad \forall \eta \in \tilde{W}_h^1
\]

for all $w \in W_2$, $\phi \in W_3$, and $\eta \in W_1$, where the residuals are: $r_u = -\Delta t (w, 2\Omega \times u_0)_{\mathcal{T}_h}$, $r_p = -c^2 \Delta t (\phi, \nabla \cdot u_0)_{\mathcal{T}_h}$, and $r_b = -N^2 \Delta t (\eta, u_0 \cdot \hat{z})_{\mathcal{T}_h}$.

The resulting matrix equations have the form:

\[
\begin{bmatrix}
A_u & -\frac{\Delta t}{2} D^T & -\frac{\Delta t}{2} Q^T \\
\frac{\Delta t}{2} c^2 D & M_p & 0 \\
\frac{\Delta t}{2} N^2 Q & 0 & M_b
\end{bmatrix}
\begin{bmatrix}
U \\
P \\
B
\end{bmatrix}
= 
\begin{bmatrix}
R_u \\
R_p \\
R_b
\end{bmatrix}
\]

\[\text{(101)}\]

where $A_u = M_u + \Delta t C \Omega$, $C \Omega$ is the asymmetric matrix associated with the Coriolis term, $M_u$, $M_p$, $M_b$, $M_p$, $M_b$, $M_b$, are mass matrices, $D$ is the weak divergence term, and $Q$ is the operator coupling $\delta u$ and $\delta b$.

To solve $Q$ is an operator containing the vertical components of $\delta u_h$. In the absence of orography, we can use the buoyancy equation to arrive at an elimination procedure by substituting point-wise expression for the buoyancy:

\[
\delta b = r_b - \frac{\Delta t}{2} N^2 \delta u \cdot \hat{z}
\]

\[\text{(102)}\]

and eliminate $\delta b_h$ from (101) by substituting (102) into (98). Note that in a planar geometry, is satisfied point-wise. However, this is not true when orography is present or on the surface of a sphere. Hence this is a further approximation of the equations. After solving for $u$ and $p$, we reconstruct $b$ from Equation which requires solving the mass matrix $M_p$. Fortunately, $M_p$ is
well-conditioned (independent of the grid resolution/time-step) and decoupled columnwise; it can be inverted with a small number of conjugate gradient iterations.

The full solution procedure is outlined as follows. First, we approximately eliminate the buoyancy to obtain a mixed system for the velocity and pressure. This produces the following mixed velocity-pressure system:

\[
\begin{align*}
\mathbf{A} \mathbf{A} \begin{pmatrix} \mathbf{U} \\ \mathbf{P} \end{pmatrix} &= \begin{pmatrix} \tilde{\mathbf{A}}_u & -\Delta t D^T \\ c^2 \Delta t D & M_p \end{pmatrix} \begin{pmatrix} \mathbf{U} \\ \mathbf{P} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{R}}_u \\ \mathbf{R}_p \end{pmatrix},
\end{align*}
\]

(103)

where \( \tilde{\mathbf{A}}_u = \mathbf{A}_u + \frac{\Delta t}{\Delta t} N^2 Q^T M_{\delta b}^{-1} Q \) and \( \tilde{\mathbf{R}}_u = \mathbf{R}_u + \frac{\Delta t}{\Delta t} Q^T M_{\delta b}^{-1} R_{\delta b} \tilde{\mathbf{A}}_u = \mathbf{A}_u + \frac{\Delta t}{\Delta t} N^2 Q^T M_{\delta b}^{-1} Q \) and \( \tilde{\mathbf{R}}_u = \mathbf{R}_u + \frac{\Delta t}{\Delta t} Q^T M_{\delta b}^{-1} R_{\delta b} \) is the modified velocity operator and right-hand side respectively. Note that in our elimination strategy, the expression for \( \tilde{\mathbf{A}}_u \) corresponds to the bilinear form obtained after substituting the point-wise expression for \( \delta b \) eliminating the buoyancy at the equation level.

\[
\begin{align*}
\tilde{\mathbf{A}}_u &\left( \mathbf{w}, \delta \mathbf{u} \right)_\mathcal{T} + \Delta t \frac{\Delta t}{2} \left( \mathbf{w}, 2\Omega \times \delta \mathbf{u} \right)_\mathcal{T} + \frac{\Delta t}{2} N^2 (\mathbf{w} \cdot \mathbf{\hat{z}}, \delta \mathbf{u} \cdot \mathbf{\hat{z}})_\mathcal{T}. 
\end{align*}
\]

(104)

A similar construction holds for \( \tilde{\mathbf{R}}_u \). Once (103) is solved, \( \delta \mathbf{b} \delta \mathbf{b}_b \) is reconstructed by solving:

\[
\begin{align*}
M_b B + \frac{\Delta t}{2} N^2 Q U &= \mathbf{R} M_b B = M_{\delta b}^{-1} R_{\delta b} - \frac{\Delta t}{2} \mathbf{Q} U.
\end{align*}
\]

(105)

Equation (105) can be efficiently inverted using a preconditioned conjugate gradient method (preconditioned by block ILU with zero in-fill).

15 5.3.2 Preconditioning the mixed velocity pressure system

The primary difficulty is finding efficient solvers for (103). This was studied by Mitchell and Müller (2016). Mitchell and Müller (2016) within the context of developing a robust preconditioner to withstand fast-moving acoustic waves. This is critical, as support fast waves driven by perturbations to the pressure field precomditioner which is robust against parameters, like \( \Delta t \) and mesh resolution. However, the implicit treatment of the Coriolis term was not taken into account. We consider two preconditioning strategies.

The first strategy follows from Mitchell and Müller (2016). As with the rotating shallow water system in Section 5.2, we can construct One strategy proposed by Mitchell and Müller (2016) is to build a preconditioner based on the Schur-complement factorization of \( \mathbf{A} \mathbf{A} \) in (103):

\[
\begin{align*}
\mathbf{A} \mathbf{A}^{-1} = \begin{bmatrix} I & \frac{\Delta t}{2} \tilde{\mathbf{A}}_u^{-1} \mathbf{D}^T \\ 0 & I \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{A}}_u^{-1} & 0 \\ 0 & \mathbf{H}^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -c^2 \frac{\Delta t}{2} D \tilde{\mathbf{A}}_u^{-1} & I \end{bmatrix},
\end{align*}
\]

(106)

where \( \mathbf{H} = M_p + c^2 \frac{\Delta t}{2} D \tilde{\mathbf{A}}_u^{-1} D^T \mathbf{H} = M_p + c^2 \frac{\Delta t}{2} D \tilde{\mathbf{A}}_u^{-1} D^T \) is the dense pressure Helmholtz operator. Because we have chosen to include the Coriolis term, the operator \( \mathbf{H} - \mathbf{H} \) is non-symmetric, and we require a sparse approximation to and has

35
\[ HH = MM_p + c^2 \frac{\Delta t^2}{4} DD \left( u \tilde{M}_u + \Delta t C_\Omega \Delta t \frac{1}{2} C_\Omega \right)^{-1} DD^T, \]  

(107)

where \( \tilde{M}_u \) is the modified velocity \( \tilde{M}_u \) is a modified mass matrix. As \( \Delta t \) increases, the contribution of \( C_\Omega \) becomes more prominent in \( HH \), making sparse approximations of \( HH \) more challenging. We shall elaborate on this further below when we present the results of our second solver strategy.

Our second strategy revolves around the hybridization preferred strategy solves the hybridizable formulation of the system defined in using Firedrake's HybridizationPC. The space of traces is generated on the faces of triangular prisms and extruded quadrilaterals, with appropriate polynomial degrees such that every trace function lies in the same polynomial space as \( u \| n \| f \) for all faces \( f \).

We locally eliminate \( U \) and \( P \) eqref:mixed-vel-pr-system. Let \( W_{2,d} \) denote the broken version of \( W_2^d \) and \( M_h \) the space of Lagrange multipliers. Then the hybridizable formulation for the velocity-pressure system reads as follows: find \( \delta u_h \in W_{2,d}^d \), \( \delta p_h \in W_{2,d}^d \), and \( \lambda_h \in M_h \) such that

\[
\begin{align*}
(w, \delta u_h)_{T_h} + \frac{\Delta t}{2} (w, 2\Omega \times \delta u_h)_{T_h} + \frac{\Delta t^2}{4} N^2 (w \cdot \hat{z}, \delta u_h \cdot \hat{z})_{T_h} + \frac{\Delta t}{2} (\nabla \cdot w, \delta p_h)_{T_h} + \langle [[w]], \lambda_h \rangle_{\partial T_h} &= \tilde{r}_u, \quad \forall w \in W_{2,d}^d \\
(\phi, \delta p_h)_{T_h} + \frac{\Delta t}{2} c^2 (\phi, \nabla \cdot \delta u_h)_{T_h} &= \tilde{r}_p, \quad \forall \phi \in W_2^3, \\
\langle \gamma, [\delta u_h] \rangle_{\partial T_h} &= 0, \quad \forall \gamma \in M_h,
\end{align*}
\]

(108)

(109)

(110)

The system (108)–(110) is automatically formed by the Firedrake preconditioner: firedrake.HybridizationPC. We then locally eliminate the velocity and pressure after hybridization, and produce the resulting system for the traces: producing the condensed problem:

\[
\begin{align*}
HH_\theta U = E, \quad HH_\theta = K \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} \hat{A}_u & -c^2 \frac{\Delta t}{2} \hat{D}^T \end{bmatrix}^{-1} \begin{bmatrix} C^T \\ 0 \end{bmatrix}, \quad E = K \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} \hat{A}_u & -c^2 \frac{\Delta t}{2} \hat{D}^T \end{bmatrix}^{-1} \begin{bmatrix} \tilde{R}_u \\ \tilde{R}_p \end{bmatrix}
\end{align*}
\]

(111)

where \( \hat{A} \) is the result of hybridizing \( A \) and \( HH_\theta \) is the statically condensed trace operator to be inverted; denotes matrices/vectors with discontinuous test and trial functions. The non-symmetric operator \( HH_\theta \) is inverted using a preconditioned generalized conjugate residual (GCR) Krylov method, as suggested in Thomas et al. (2003). For our choice of preconditioner, we follow strategies outlined in Elman et al. (2001) and employ an algebraic multigrid method (V-cycle) with GMRES (five iterations) smoothers on the coarse levels. The GMRES smoothers are preconditioned with block ILU on each level. For the finest level, block ILU produces a line smoother (necessary for efficient solution on thin domains) when the trace variable nodes are null.
numbered in vertical lines, as is the case in our Firedrake implementation. On the coarser levels, less is known about the properties of ILU under the AMG coarsening strategies, but as we shall see, we observe performance that suggests ILU is still behaving as a line smoother. More discussion on multigrid for non-symmetric problems can be found in Bramble et al. (1994, 1988); Mandel (1986). A gravity wave test using our solution strategy and hybridization preconditioner is illustrated in Figure 7 for a problem on a condensed Earth (radius scaled down by a factor of 125) and 10km lid.

![Figure 7. Buoyancy perturbation (y - z cross section) at t = 3600s from a simple gravity wave test (Δt = 100s). The initial conditions (in lat.-long. coordinates) for the velocity is a simple solid-body rotation: \( \mathbf{u} = 20 e \), where \( e \) is the unit vector pointing in the direction of decreasing longitude. A buoyancy anomaly is defined via \( b = \frac{d^2}{d^2 + q^2} \sin(\pi z/10000) \), where \( q = R \cos^{-1}(\cos(\phi) \cos(\lambda - \lambda_\phi)) \), \( d = 5000m \), \( R = 6371 \text{km} / 125 \) is the planet radius, and \( \lambda_\phi = 2/3 \). The equations are discretized using the lowest-order method RTCF_1, with 24,576 quadrilateral cells in the horizontal and 64 extrusion levels. The velocity-pressure system is solved using hybridization.

5.3.3 Robustness against acoustic Courant number with implicit Coriolis

A desired property of the implicit linear solver is robustness with respect to the time-step size. In particular, it is desirable for the execution time to remain constant across a wide range of \( \Delta t \). As \( \Delta t \) increases, the conditioning of the elliptic operator becomes worse. Therefore, iterative solvers need to work much harder to reduce the residual down to a specified tolerance.

In this final experiment, we repeat a similar study to that presented in Mitchell and Müller (2016). Our setup closely resembles the gravity wave test of Skamarock and Klemp (1994) extended to a spherical annulus. We initialize the velocity in a simple solid-body rotation and introduce a localized buoyancy anomaly. A Coriolis term is introduced as a function of the Cartesian coordinate \( z \), and is constant along lines of latitude (\( f \)-plane): \( 2\Omega = 2\Omega_0 \frac{\hat{z}}{\hat{\Omega}} \hat{z} \), with angular rotation rate \( \Omega_0 = 7.292 \times 10^{-5} \text{s}^{-1} \). We fix the resolution of the problem and run our solver for the solver over a range of \( \Delta t \). We measure this by adjusting the horizontal acoustic Courant number \( \lambda_C = c \frac{\Delta t}{\Delta x} \), where \( c \) is the speed of sound and \( \Delta x \) is the horizontal resolution. We remark that the range of Courant numbers used in this paper exceeds what is typical in operational forecast settings (typically between \( O(2) - O(10) \)). The grid set up mirrors that of actual global weather models; we extrude a spherical mesh of the Earth.
Table 6. Grid set up and discretizations for the acoustic Courant number study. The total unknowns (velocity and pressure) and hybridized unknowns (broken velocity, pressure, and trace) are shown in the last two columns (millions). The vertical resolution is fixed across all discretizations.

<table>
<thead>
<tr>
<th>Discretizations and grid information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed method</td>
</tr>
<tr>
<td>RT(_1)</td>
</tr>
<tr>
<td>RT(_2)</td>
</tr>
<tr>
<td>BDFM(_2)</td>
</tr>
<tr>
<td>RTCF(_1)</td>
</tr>
<tr>
<td>RTCF(_2)</td>
</tr>
</tbody>
</table>

upwards to a height of 85km. The set up for the different discretizations (including degrees of freedom for the velocity-pressure and hybridized systems) are presented in Table 6.

It was shown in Mitchell and Müller (2016) that using a sparse approximation of the pressure Schur-complement of the form:

\[
\widetilde{H} = MM_p + c^2 \Delta t^2 \frac{D D}{4} \left( \text{Diag}( \tilde{A}_u u ) \right)^{-1} DD^T
\]

served as a good preconditioner, leading to a system that was amenable to multigrid methods and resulted in a Courant number independent solver. However, when the Coriolis term is included, this is no longer the case: the diagonal approximation to \( \tilde{M}_u \text{Diag}( \tilde{A}_u ) \) becomes worse with increasing \( \lambda_C \). To demonstrate this, we solve the mixed problem gravity wave system on a low-resolution grid (10km lid, 10 vertical levels, maintaining the same cell aspect ratio as in Table 6) using the Schur-complement factorization (106) and use LU factorizations to apply both \( \tilde{A}_u^{-1} \) and \( \tilde{H}^{-1} \). \( \tilde{H}^{-1} \) is computed. LU factorizations are applied to invert both \( \tilde{A}_u^{-1} \) and \( \tilde{H}^{-1} \). Inverting the Schur-complement \( \tilde{H}^{-1} \) is done using preconditioned GMRES iterations, and a flexible-GMRES algorithm is used on the full velocity-pressure system. If \( \tilde{H}^{-1} = \tilde{H}^{-1} \) \( \tilde{H}^{-1} \) is a good approximation to \( \tilde{H}^{-1} \tilde{H}^{-1} \), then we should see low iteration counts when inverting \( \tilde{H} \) in the Schur-complement solve. Figure 8 shows the results of this study for a range of Courant numbers.

For the lower-order methods, the number of iterations to invert \( \tilde{H} \tilde{H}^{-1} \) grow slowly but remain under control. Increasing the approximation degree by one results in degraded performance. As \( \Delta t \) increases, the number of Krylov iterations needed to invert the system to a relative tolerance of \( 10^{-5} \) grows rapidly. It is clear that this sparse approximation is not robust against Courant number. This can be explained by the fact that diagonalizing the velocity operator fails to take into account the effects of the Coriolis term (which appear in off-diagonal positions in the operator). Even if one were to use traditional mass-lumping (row-wise sums), the Coriolis effects are effectively cancelled out due to asymmetry.

Hybridization avoids this problem entirely: we always construct an exact Schur complement, and only have to worry about solving the trace system (111). We now show that this approach (described in Section 5.3.2) is much more robust to increases in the Courant number when \( \Delta t \) changes. We use the same workstation as for the three-dimensional CG/HDG problem in Section 38.
Figure 8. Number of Krylov iterations to invert the Helmholtz system using $\tilde{H}^{-1}$ as a preconditioner. The preconditioner is applied using a direct LU factorization within a GMRES method on the entire pressure equation. While the lowest-order methods grow slowly over the Courant number range, the higher-order (by only one approximation order) methods quickly degrade and diverge after the critical range $\lambda_C = O(2)–O(10)$. At $\lambda_C > 32$, the solvers take over 150 iterations.

5.1 (executed with a total of 40 MPI processes). Figure 9 shows the parameter test for all the discretizations described in Table 6. We see that, in terms of total number of GCR iterations needed to invert the trace system, hybridization is far more robust against Courant number than the approximate Schur complement approach controlled as Courant number increases. They largely remain constant throughout the entire parameter range, only varying by an iteration or two. It is not until after $\lambda_C > 32$ that we begin to see a larger jump in the number of GCR iterations. This is expected, since the Coriolis operator causes the problem to become singular for very large Courant numbers. However, unlike with the approximate Schur-complement solver, iteration counts are still under control. In particular, each method (lowest and higher order) remains constant throughout the critical range (shaded in gray in Figures ?? and ??). GCR iterations vs Courant number: Figure 9).

In Figure ?? (right), we display the ratio of execution time and the time-to-solution at the lowest Courant number of two. We perform this normalization to better compare the lower and higher order methods (and discretizations on triangular prisms vs extruded quadrilaterals). The calculation of the ratios includes the time needed to eliminate and reconstruct the hybridized velocity/pressure variables. The fact that the hybridization solver remains close to one demonstrates that the entire solution procedure is largely $\lambda_C$-independent until around $\lambda_C = 32$. The overall trend is largely the same as what is observed in Figure ?? the number of Krylov iterations to reach solver convergence. This is due to our hybridization approach being solver dominated, with local operations like forward elimination together with local recovery taking approximately $1/3$ of the total execution time for each method. The percentage breakdown of the hybridization solver is similar to what is already presented in Section 5.1.2.
Figure 9. Work vs Courant number parameter test run on a fully-loaded compute node. Both figures display the hybridized solver for each discretization, described in Table 6. The left figure displays to total iteration count (preconditioned GCR) to solve the trace system to a relative tolerance of $10^{-5}$. The right figure displays the relative work of each solver, which takes into account the time required to forward eliminate and locally recover the velocity and pressure.

Implicitly treating the Coriolis term has been discussed for semi-implicit discretizations of large-scale geophysical flows (Temperton, 1997; Côté and Staniforth, 1988; Cullen, 2001; Nechaev and Yaremchuk, 2004). Incorporating (Temperton, 1997; Cullen, 2001) the addition of the Coriolis term presents a particular challenge to the solution finite element discretizations is difficult, as this makes the challenge of finding robust of these equations since it increases the difficulty of finding a good sparse approximation of the resulting elliptic operator even more difficult. The hybridization process mimics standard staggered finite difference elimination procedures used in many global circulation models. In other words, hybridization is the finite element analogue of point-wise elimination strategies in finite difference codes.

6 Conclusions

We have presented Slate, and shown how this language can be used to create concise mathematical representations of localized linear algebra on the tensors corresponding to finite element forms. We have shown how this DSL can be used in tandem with UFL in Firedrake to implement solution approaches making use of automated code generation for static condensation,
hybridization, and localized post-processing. Setup and configuration is done at runtime, allowing one to switch in different discretizations at will. In particular, this framework alleviates much of the difficulty in implementing such methods within intricate numerical, and paves the way for future low-level optimizations. In this way, the framework in this paper can be used to help enable the rapid development and exploration of new hybridization and static condensation techniques for a wide class of problems. We remark here that the reduction of global matrices via element-wise algebraic static condensation, as described in Guyan (1965); Irons (1965) is also possible using Slate, including other more general static condensation procedures outside the context of hybridization.

Our approach to preconditioner design revolves around its composable nature, in that these Slate-based implementations can be seamlessly incorporated into complicated solution schemes. In particular, there is current research in the design of dynamical cores for numerical weather prediction using implementations of hybridization and static condensation with Slate (Bauer and Cotter, 2018; Shipton et al., 2018). The performance of such methods for geophysical flows are a subject of ongoing investigation.

In this paper, we have provided some examples of hybridization procedures for compatible finite element discretizations of geophysical flows. These approaches avoid the difficulty in constructing sparse approximations of dense elliptic operators. Static condensation arising from hybridizable formulations can best be interpreted as producing an exact Schur-complement factorization on the global hybridizable system. This eliminates the need for outer iterations from a suitable Krylov method to solve the full mixed system, and replaces the original global mixed equations with a condensed elliptic system. More extensive performance benchmarks, which requires detailed analysis of the resulting operator systems arising from hybridization, is a necessary next-step to determine whether hybridization provides a scalable solution strategy for compatible finite elements in operational settings.

Code and data availability. The contribution in this paper is available through open-source software provided by the Firedrake Project: https://www.firedrakeproject.org/. We cite archives of the exact software versions used to produce the results in this paper. For all components of the Firedrake project used in this paper, see Zenodo/Firedrake (2019). The numerical experiments, full solver configurations, code-verification (including local-processing), and raw data are available in Zenodo/Tabula-Rasa (2019).
Appendix A: Semi-implicit method for the shallow water system

For some tessellation, \( T_h \), our semi-discrete mixed method for (81)–(82) seeks approximations \((u_h, D_h) \in U_h \times V_h \subset H(\text{div}) \times L^2\) satisfying: \((u_h, D_h) \in U_h \times V_h \subset H(\text{div}) \times L^2\) satisfying:

\[
\begin{align*}
(\omega, \frac{\partial u_h}{\partial t})_{T_h} - (\nabla \cdot (\omega \cdot u_h^\perp), u_h^\perp)_{T_h} + (w, f u_h^\perp)_{T_h} + \left[ \left[ n^\perp \omega \cdot u_h^\perp \right] , \tilde{u}_h^\perp \right]_{\partial T_h} & \quad - \left( \nabla \cdot w, g (D_h + b) + \frac{1}{2} |u_h|^2 \right)_{T_h} = 0, \quad \forall \omega \in U_h, \quad (A1) \\
(\phi, \frac{\partial D_h}{\partial t})_{T_h} - (\nabla \phi \cdot u_h D_h)_{T_h} + \left[ \left[ \phi u_h \right], \tilde{D}_h \right]_{\partial T_h} = 0, \quad \forall \phi \in V_h, \quad (A2)
\end{align*}
\]

where \( \tilde{\cdot} \) indicates that the value of the function should be taken from the upwind side of each facet. The discretisation of the velocity advection operator is an extension of the energy-conserving scheme of Natale and Cotter (2017) to the shallow-water equations.

The time-stepping scheme follows a Picard iteration semi-implicit approach, where predictive values of the relevant fields are determined via an explicit step of the advection equations, and corrective updates are generated by solving an implicit linear system (linearized about a state of rest) for \((\Delta u_h, \Delta D_h) \in U_h \times V_h\), given by

\[
\begin{align*}
(w, \Delta u_h)_{T_h} + \frac{\Delta t}{2} (w, f \Delta u_h^\perp)_{T_h} - \frac{\Delta t}{2} (\nabla \cdot w, g \Delta D_h)_{T_h} &= -R_u[u_h^{n+1}, D_h^{n+1}; w], \quad \forall w \in U_h, \quad (A3) \\
(\phi, \Delta D_h)_{T_h} + \frac{H \Delta t}{2} (\phi, \nabla \cdot \Delta u_h)_{T_h} &= -R_D[u_h^{n+1}, D_h^{n+1}; \phi], \quad \forall \phi \in V_h, \quad (A4)
\end{align*}
\]

where \( H \) is the mean layer depth, and \( R_u \) and \( R_D \) are residual linear forms that vanish when \( u_h^{n+1} \) and \( D_h^{n+1} \) are solutions to the implicit midpoint rule time discretization of (A1)–(A2). The residuals are evaluated using the predictive values of \( u_h^{n+1} \) and \( D_h^{n+1} \).

The implicit midpoint rule time discretization of the non-linear rotating shallow water equations (A1)–(A2) is:

\[
\begin{align*}
(w, u_h^{n+1} - u_h^n)_{T_h} - \Delta t \left( \nabla \cdot (w \cdot u_h^{n+1}^\perp), u_h^{n+1}^\perp \right)_{T_h} + \Delta t \left( w, f u_h^{n+1}^\perp \right)_{T_h} &
+ \Delta t \left[ \left[ n^\perp \omega \cdot u_h^{n+1}^\perp \right] , \tilde{u}_h^{n+1}^\perp \right]_{\partial T_h} \\
- \Delta t \left( \nabla \cdot w, g (D_h^n + b) + \frac{1}{2} |u_h^n|^2 \right)_{T_h} &= 0, \quad \forall w \in U_h, \quad (A5) \\
(\phi, D_h^{n+1} - D_h^n)_{T_h} - \Delta t \left( \nabla \phi \cdot u_h^{n+1} D_h^n \right)_{T_h} + \Delta t \left[ \left[ \phi u_h^n \right], \tilde{D}_h^n \right]_{\partial T_h} &= 0, \quad \forall \phi \in V_h, \quad (A6)
\end{align*}
\]

where \( u_h^n = (u_h^{n+1} + u_h^n) / 2 \) and \( D_h^n = (D_h^{n+1} + D_h^n) / 2 \).

One approach to construct the residual functionals \( R_u \) and \( R_D \) would be to simply define these from (A5)–(A6). However, this leads to a small critical time-step for stability of the scheme. To make the numerical scheme more stable, we define
residuals as follows. For $R_u$, we first solve for $\boldsymbol{v}_h \in U_h$ such that

$$
(w, v_h - u_h^n)_{T_h} - \Delta t \left( \nabla^\perp (w \cdot u_h^*), v^*_h \right)_{T_h} + \Delta t \left( w, f v^*_h \right)_{T_h} + \Delta t \left( \left[ \begin{array}{c} \Pi w \cdot u_h^* \\ \Pi v_h^* \end{array} \right], \tilde{v}^*_{\partial T_h} \right)_{\partial T_h} - \Delta t \left( \nabla \cdot w, g(D_h^n + b) + \frac{1}{2}|u_h^n|^2 \right)_{T_h} = 0, \quad \forall w \in U_h,
$$

(A7)

where $v^*_h = (v_h + u_h^n)/2$. This is a linear variational problem. Then,

$$
R_u[u_h^{n+1}, D_h^{n+1}; w] = (w, v_h - u_h^{n+1})_{T_h}.
$$

(A8)

Similarly, for $R_D$ we first solve for $E_h \in V_h$ such that

$$
(\phi, E_h - D_h^n)_{T_h} - \Delta t \left( \nabla \phi, u_h^* E^*_h \right)_{T_h} + \Delta t \left( \left[ \begin{array}{c} \phi u_h^* \\ \phi E_h^* \end{array} \right], E^*_{\partial T_h} \right)_{\partial T_h} = 0, \quad \forall \phi \in V_h,
$$

(A9)

where $E^*_h = (E_h + D_h^n)/2$. This is also a linear problem. Then,

$$
R_D[u_h^{n+1}, D_h^{n+1}; \phi] = (\phi, E_h - D_h^{n+1})_{T_h}.
$$

(A10)

This process can be thought of as iteratively solving for the average velocity and depth that satisfies the implicit midpoint rule discretisation. Both (A7) and (A9) can be solved separately, since there is no coupling between them. The fields $v_h$ and $E_h$ are then used to construct the right-hand side for the implicit linearized system in (A3)–(A4). Once the system is solved, the solution $(\Delta u_h, \Delta D_h)$ is then used to update the iterative values of $u_h^n$ and $D_h^n$ according to

$$
(u_h^{n+1}, D_h^{n+1}) \leftarrow (u_h^n, D_h^n) + (\Delta u_h, \Delta D_h),
$$

having initially chosen $(u_h^n, D_h^n) = (u_h^n, D_h^n)$.

Author contributions. T. H. Gibson is the principal author and developer of the software presented in this paper and main author of the text. Authors L. Mitchell and D. A. Ham assisted and guided the software abstraction as a domain-specific language, and edited text. C. J. Cotter contributed to the formulation of the geophysical fluid dynamics and the design of the numerical experiments, and edited text.

Competing interests. D. A. Ham is an executive editor of the journal. The other authors declare they have no other competing interests.

Acknowledgements. This work was supported by the Engineering and Physical Sciences Research Council (grant numbers: EP/M011054/1, EP/L000407/1, and EP/L016613/1), and the Natural Environment Research Council (grant number: NE/K008951/1).
References


