STABILITY OF TWO CONSERVATIVE, HIGH-ORDER
FLUID-FLUID COUPLING METHODS

J. M. CONNORS AND R. D. DOLAN

ABSTRACT. This paper investigates two methods of coupling fluids across an
interface, motivated by air-sea interaction in application codes. One method
is for sequential configurations, where the air code module is invoked over
some time interval prior to the sea module. The other method is for concur-
rent setups, in which the air and sea modules run in parallel. The focus is
the temporal representation of air-sea fluxes. The methods we study conserve
moments of the fluxes, with an arbitrary order of accuracy possible in time.
Different step sizes are allowed for the two fluid codes. An a posteriori sta-
bility indicator is defined, which can be computed efficiently on-the-fly over
each coupling interval. For a model of two coupled fluids with natural heat
convection, using finite elements in space, we prove the sufficiency of our sta-
bility indicator. Under certain conditions, we also prove that stability can be
enforced by iteration when the coupling interval is small enough. In particular,
for solutions in a certain class, we show that the step size scaling is no worse
than $O(h)$ in three dimensions of space, where $h$ is a mesh parameter. This
is a sharper result than what has been shown previously for related algorithms
with finite element methods. Computational examples illustrate the behavior
of the algorithms under a wide variety of configurations.

1. Introduction

This paper concerns algorithms used to resolve interactions between the atmo-
sphere and ocean for applications like climate research, hurricane modeling and
regional weather forecasting. Since the physical properties of the air and sea sys-
tems are very different, separate code modules are used to simulate these with
independent, internal numerical methods. Another code module, called a flux cou-
pler, handles communication between the air and sea modules. The coupler receives
state data from the fluid modules that is used to compute fluxes of conserved physical
quantities, which are then transferred between the fluids in the form of boundary
conditions. Important examples are the Model Coupling Toolkit (MCT) used to
connect the Weather Research and Forecasting (WRF) and Regional Oceanic Mod-
eling System (ROMS) models (e.g. [12] [19]), and the related CPL7 coupler used in
the Community Earth Systems Model (CESM) [7].

There are challenges to couple the modules. Direct resolution of the boundary
layers of the fluids is prohibitively expensive and the air model bottom need not
match the sea model top. Thus, special bulk formulae have been derived to model
the fluxes in terms of the state values at the air and sea boundaries. Examples
we consider for the fluxes of momentum and heat are given later by (12) and (13).

Key words and phrases. air-sea, atmosphere-ocean, fluid-fluid, partitioned time stepping, con-
servative coupling.
which are motivated by the paper \cite{14}. Also, the computational grids for the air and sea are different, so some conservative remapping of fluxes is needed in space. The aforementioned coupling software includes some conservative, spherical remapping functionality. Time stepping methods are a challenge since the internal time steps of the air and sea codes are different but the bulk flux formulae use simultaneous state values. This paper is focused on the latter issue of the calculation of fluxes for time stepping purposes.

The temporal representation of fluxes is complicated by a lengthy list of mathematical, scientific and computational considerations. We study two coupling methods as reference points to help advance our understanding regarding these issues. The precise algorithms are defined in Section 4. Due to the complexity of the problem, we shall first provide a high-level explanation of the algorithms in Section 1.1. We will then provide more background and fully explain our current goals in Section 1.2.

1.1. Two coupling algorithms. We study two methods to couple the air and sea modules that we refer to as the sequential and concurrent modes. In the sequential mode, the air code module is run on the same set of processors as the sea code. Since the atmospheric dynamics are thought of as driving the ocean surface conditions, we assume the air code is run first. For the concurrent mode, the air and sea modules run simultaneously on different sets of processors.

We explain the methods to advance the air and sea modules from a time $t$ to time $t + \Delta t$, but they may advance internally using different numbers of substeps. Assume that the algorithms have been run on a previous interval $[t - \Delta t, t]$ already. The flux computations depend on least-squares data reconstructions in time, using polynomials of a chosen order $k \geq 0$, which can then be evaluated at any desired time. Enough data points (substeps) must be available to form the chosen reconstructions, which are defined rigorously in Section 1.2. To illustrate the algorithms, consider the case $k = 1$, with 8 air substeps and 2 sea substeps. In the following figures, flux and state values are pictured abstractly as heights along the vertical axes, with time on the horizontal axes. An arrow head points to data that is functionally dependent on the corresponding data at the tail; double arrows (with no tail) denote a two-way (implicit) dependence.
Figure 2. the concurrent mode. Left: extrapolated fluxes are used to compute new air and sea states, thus conserving moments of the flux. Right: the fluxes are updated after each new air state is computed by using extrapolated sea state data. This avoids CPU idle time, compared to using the new sea states.

In the sequential mode, the air code requires fluxes but the bulk formulae reference sea states at the same times. The sea states are provided by reconstructing the states on the previous interval, \([t - \Delta t, t]\), and extrapolating to the current interval, \([t, t + \Delta t]\). The fluxes are then computed semi-implicitly for each air step, using the current air state and extrapolated sea state. After the air code has reached time \(t + \Delta t\), the sea code is advanced from time \(t\) using the reconstructed fluxes on the current interval. The moments of the fluxes up to order \(k\) are therefore conserved in a time-integrated sense. This situation is depicted in Figure 1.

In order to achieve the same conservation property in the concurrent mode, we specify the fluxes before evolving the fluid modules. Thus, the reconstruction of the fluxes from the previous time interval is used as an extrapolation to the current interval. The fluxes must be updated on the current time interval before the method can be continued beyond time \(t + \Delta t\). Fluxes are updated as the new air states are computed, using current air states but extrapolated sea states. In order to use the current sea states, the flux coupler would need to wait until the sea module was done, which would increase CPU idle time. This method is illustrated in Figure 2.

One of the questions we address is the numerical stability of these algorithms. Previous work has shown that coupling algorithms can be designed with unconditional stability, but these methods were low-order accurate and not conservative \[6, 5\]. More recently, it has been shown in \[1\] that higher order can be achieved using deferred correction techniques, but these methods require multiple invocations of each module at each time step, using different equations for the internal dynamics. Furthermore, they have not yet incorporated conservation and multirate time stepping. Another approach is developed in the papers \[11, 12\], in which a minimally-intrusive iteration on a coupling interval is used to enforce conditional stability and accuracy. These methods have already been used with coupled WRF-ROMS simulations. We consider the same type of iterations in this paper; deferred correction methods are not as far along in development and so are left for future investigation.

Iteration for the sequential method means that after computing on the current interval, the air code is re-run starting from time \(t\). The reconstruction of the sea states from the previous iteration is used to compute the fluxes on the current
The reconstruction of the new fluxes is then used to compute new sea states, and the process can be repeated until some stopping criterion is achieved; see Figure 3.

For the concurrent method, new fluxes are computed corresponding to each air time step using the most recent air state and the last available reconstruction of sea states. We do not use the most recent sea states to compute the fluxes because it would introduce CPU idle time and increase storage requirements. New air and sea states can then be computed using the reconstruction of the new fluxes. Again, the process can be repeated as many times as desired on the current coupling interval; see Figure 4.

1.2. Further background and current goals. Since we have chosen to investigate the sequential coupling method using semi-implicit fluxes (implicit for the air), our sequential implementation does not fit immediately into existing code structures like CESM or coupled WRF-ROMS regional models. In these latter codes, data must first be passed to the flux coupler, then fluxes can be computed and returned to the air (or sea) code. However, we choose to explore this method to provide a reference point, as it may have superior accuracy or stability properties.
In practice, algorithms could approximate semi-implicit fluxes directly by iterating between the flux coupler and air module at a given air substep, until the flux at the current time is consistent with the current air state, before proceeding to the next air substep. This might be too expensive. More likely, since many air substeps are commonly taken in application, higher-order air state extrapolations could be performed locally in time within a coupling interval to approximate the correct flux at each internal air step explicitly, which would limit to one invocation of the flux coupler per air substep (no air-flux iteration prior to communicating with the sea module). This approaches semi-implicit coupling in the sense that the set of fluxes over a coupling interval would be computed using the set of current air states for that interval (and perhaps some also from a previous interval), not from a previous iteration between the air and sea codes.

The algorithms in this paper fit loosely into the global Schwarz framework outlined in [11, 12] and the recent thesis of Pelletier [16]. However, there are some important technical distinctions between our algorithms and methods implemented previously that point to some of the questions we wish to investigate. One issue is the method of data reconstruction. Least-squares may avoid aliasing errors generated by anchoring flux computations to data at specific previous times. Also, we can filter out some high-order fluctuations that may induce instabilities, without dropping the formal order of accuracy down by setting $k = 0$. Each reconstruction requires $k + 1$ storage vectors that can be calculated on-the-fly using some precomputed coefficients that do not depend on space or time. In addition, the algorithms in this paper are formally order $\Delta t^{k+1}$ accurate for the fluxes and conserve moments of the fluxes up to order-k in time. In previous investigations, flux calculations have not had all of these properties.

Although our algorithms have differences with others in the literature, our goal is not to claim the development of novel or superior coupling methods. Flux computation for coupled fluids is a very complicated problem that requires considerable on-going research. The goal of this paper is to identify possible efficiency gains by comparing different coupling constructions. Future algorithmic development will likely focus on adaptivity and load balancing issues. There is a need to better understand the differences between sequential and concurrent coupling methods. Also, iteration is quite expensive in application. It merits asking to what extent higher-order extrapolation can be used to gain accuracy without losing stability, which may depend on $\Delta t$. A reduction of $\Delta t$ could enable high-order accuracy and stability without any iteration. On the other hand, order reduction (reducing $k$) might mitigate stability issues locally in time without decreasing $\Delta t$, and possibly without any significant loss of accuracy. Therefore, we seek to understand the relationships between $k$, $\Delta t$ and iteration requirements as they pertain to accuracy and stability, for both sequential and concurrent modes. To help with this, we derive an \textit{a posteriori} stability indicator that can be computed on-the-fly for each coupling interval.

The remainder of this paper is organized as follows. In Section 2 we introduce a surrogate model of two coupled fluids with natural heat convection. This model retains the key mathematical aspects of the coupling conditions but removes many other complications of the full atmosphere-ocean system, simplifying our analysis. In Section 3 we provide some preliminary mathematical results and notation. The full algorithms are defined precisely in Section 4, then we provide an analysis of
Consider two fluids that are coupled across a flat interface. Let $\Omega_A = \Gamma_T \times (0, H)$ and $\Omega_S = \Gamma_T \times (-H, 0)$ be the two fluid domains, where $H > 0$ is a fixed height and $\Gamma_T = (0, L)^{d-1}$, for $d = 2$ or $d = 3$. Here $L > 0$ is the horizontal domain length parameter. The velocity, pressure and temperature will be denoted by $(u_A, p_A, \theta_A)$ on $\Omega_A$ and $(u_S, p_S, \theta_S)$ on $\Omega_S$. We use a nondimensionalized model to make the role of various parameters explicit. Given the following reference quantities: $L$ (length), $U$ (speed) and $\theta_{REF}$ (temperature), we present our model in dimensionless coordinates and variables. After rescaling, we redefine the interface as $\Gamma = (0, H)$. The rescaled fluid domains are $\Omega_A = \Gamma_T \times (0, \tilde{H})$ and $\Omega_S = \Gamma_T \times (-\tilde{H}, 0)$, with $\tilde{H} = H / L$. Due to the relative scaling of $H$ and $L$ in application, we assume here that $\tilde{H} \ll 1$. Given $i \in \{A, S\}$, the governing equations are:

$$
\begin{align*}
(1) & \quad \partial_t u_i - \nabla_i (u_i) + u_i \cdot \nabla u_i + \nabla p_i = f_i \quad \text{on } \Omega_i, \ t > 0, \\
(2) & \quad \left< 0, -\tilde{g}(1 - \tilde{\beta}_i \theta_i - \overline{\theta_i}) \right> = f_i, \\
(3) & \quad \partial_t \theta_i - D_i(\theta_i) + u_i \cdot \nabla \theta_i = 0 \quad \text{on } \Omega_i, \ t > 0, \\
(4) & \quad \nabla \cdot u_i = 0 \quad \text{on } \Omega_i, \ t > 0,
\end{align*}
$$

plus initial conditions. Here, $\tilde{g} = gL / (U^2)$ where $g$ is the magnitude of gravitational acceleration. The coefficients of thermal expansion for air and sea water are taken as the positive constants $\beta_A$ and $\beta_S$, respectively, and $\tilde{\beta}_i = \beta_i \theta_{REF}$, for $i \in \{A, S\}$. The temperature average over a domain $\Omega_i$ is denoted by $\overline{\theta_i}$. The viscosity terms are $\nu_i(u_i)$ and the heat diffusion terms are $D_i(\theta_i)$.

The model (1)-(4) represents two coupled, incompressible fluids with buoyancy-driven motions induced by small density gradients. Although we fix constant reference densities for our fluids, buoyancy forces are modeled by the famous Boussinesq approximation (2) as being proportional to thermal deviations from a reference background temperature; here we use the average temperature for the bulk fluid as the reference value. The proportionality constants $\beta_A$ and $\beta_S$ have been measured empirically; we use the values in Appendix A for our computations. Roughly speaking, where the fluid is hotter (cooler) than average a local decrease (increase) in the downward gravitational pull is induced in the momentum equation (1), resulting in rising (falling) fluid trajectories. In turn, the term $u_i \cdot \nabla \theta_i$ in (3) causes heat convection.

Due to the nature of flow features that result from the scale discrepancy $H \ll L$, it is typical in practice to treat horizontal and vertical viscosity and diffusion processes differently (see, e.g. [12, 8]). Let $V(u)$ represent any $d \times d$ tensor or...
matrix, with entries \( V(u)_{ij} \), \( 1 \leq i, j \leq d \). Define a decomposition by

\[
V(u) = V(u)^H + V(u)^\perp,
\]

\((5)\)

\[
(V(u)^H)_{ij} = \begin{cases} (V(u))_{ij} & \text{for } i = 1, \ldots, d-1, \ j = 1, \ldots, d, \\ 0, & \text{otherwise} \end{cases}
\]

\[
(V(u)^\perp)_{ij} = \begin{cases} (V(u))_{ij} & \text{for } i = d, \ j = 1, \ldots, d, \\ 0, & \text{otherwise} \end{cases}
\]

Now let \( V(u) = (\nabla u + \nabla u^T)/2 \) be specifically the viscous part of the Cauchy stress tensor. The viscosity terms are decomposed into horizontal and vertical components:

\[
V_i(u_i) = 2\nabla \cdot \left( \frac{1}{\text{Re}_H^i} V(u_i)^H + \frac{1}{\text{Re}_T^i} V(u_i)^\perp \right).
\]

These are given in terms of the following nondimensionalized parameters:

- Horizontal Reynolds numbers: \( \text{Re}_H^i = \frac{UL}{\nu_H^i}, \ i \in \{A, S\} \)
- Vertical Reynolds numbers: \( \text{Re}_T^i = \frac{UL}{\nu_T^i}, \ i \in \{A, S\} \).

The constants \( \nu_H^i > 0 \) and \( \nu_T^i > 0 \) are horizontal and vertical kinematic viscosity parameters, respectively.

Let the coordinates be \( x = (x_1, \ldots, x_d) \). The heat diffusion terms are:

\[
D_i(\theta_i) = \frac{1}{\text{Re}_H^i \text{Pr}_i} \sum_{j=1}^{d-1} \frac{\partial^2 \theta_i}{(\partial x_j)^2} + \frac{1}{\text{Re}_T^i \text{Pr}_i} \frac{\partial^2 \theta_i}{(\partial x_d)^2}.
\]

The Prandtl numbers for air (\( \text{Pr}_A \)) and seawater (\( \text{Pr}_S \)) are dimensionless, positive constants (see [10]).

All variables are horizontally periodic. Denote by \( \Gamma_b \) the model bottom, where \( x_d = -\tilde{H} \), and by \( \Gamma_t \) the model top, \( x_d = \tilde{H} \). Let \( \hat{n}_i \) denote the outward-pointing unit normal for \( \Omega_i, \ i = A, S \). Then we impose the following boundary conditions:

\[
u_S = 0 \quad \text{and} \quad \frac{\partial \theta_S}{\partial x_d} = 0 \text{ on } \Gamma_b,
\]

\[
u_A \cdot \hat{n}_A = 0 \quad \text{and} \quad \hat{n}_A \cdot V(u_A) = 0 \text{ on } \Gamma_t,
\]

\[
(\text{Re}_A^{\perp} \text{Pr}_A)^{-1} \frac{\partial \theta_A}{\partial x_d} = \frac{C_t}{U c_A \rho_A} (\theta_t - \theta_A) \text{ on } \Gamma_t,
\]

where \( c_A \) is a specific heat and \( \rho_A \) a constant density. The value \( C_t > 0 \) is just a proportionality constant and \( \theta_t \) is a fixed reference temperature for the model top. Heat energy is allowed to transfer across the model top in a way that drives the temperature toward \( \theta_t \) there, which supports convective currents and allows heat energy to be put into the system without driving the temperature up indefinitely. The vector \( \tau \) represents anything tangent to \( \Gamma_t \), so that a no-horizontal-stress condition results at the model top.

So far, the boundary conditions are chosen to enable flows with some features similar to those found in an air-sea system. But our primary interest is in the
coupling at the fluid-fluid interface $\Gamma_I$. Here we impose a zero-penetration and horizontal slip-with-friction condition. That is,
\[
\rho_A \frac{\partial u_A}{\partial x_d} \cdot \tau = \rho_S \frac{\partial u_S}{\partial x_d} \cdot \tau = \rho_A \kappa |u_A - u_S| (u_A - u_S) \cdot \tau,
\]
for every vector $\tau$ tangent to $\Gamma_I$. Here $\rho_S$ is a constant density. We assume $\rho_A \ll \rho_S$ due to the relative densities of air and water. The value $\kappa$ is a friction parameter. Heat flux across the interface is given by
\[
\left( \text{Re}_i^\perp \text{Pr}_i \right)^{-1} \frac{\partial \theta_i}{\partial x_d} = (c_i \rho_i)^{-1} F(u_A, u_S; \theta_A, \theta_S), \quad i \in \{A, S\}.
\]
where
\[
F(u_A, u_S; \theta_A, \theta_S) = \frac{C_{\text{sol}}}{U \theta_{\text{REF}}} + \frac{C_{\text{ir}}}{U} (\theta_A - \theta_S) + c_A \rho_A C_{\text{sen}} |u_A - u_S| (\theta_A - \theta_S).
\]
We assume $c_A < c_S$ due to the relative sizes of the specific heats for air and sea water. The remaining parameters are for the net downward solar radiative flux ($C_{\text{sol}} > 0$), infra-red radiation ($C_{\text{ir}}$) and sensible heat convection ($C_{\text{sen}}$). These are taken to be positive constants.

3. MATHEMATICAL PRELIMINARIES

**Definition 1.** Standard $L^2$-inner-products are defined on each domain by
\[
(u, \tilde{u})_A \equiv \int_{\Omega_A} u \cdot \tilde{u} \, dx \quad \text{and} \quad (v, \tilde{v})_S \equiv \int_{\Omega_S} v \cdot \tilde{v} \, dx.
\]
The domain subscript will be suppressed for the $L^2$ norm; if $u \in L^2(\Omega_i)^d$ we set $\|u\| = (u, u)^{1/2}_i$. The next notation helps to represent various boundary integrals.

**Definition 2.** Let $\Sigma_i \subset \partial \Omega_i$ for any $i \in \{A, S\}$ be $\sigma$-measurable (for some appropriate boundary measure $\sigma$). Let $s$ be any positive integer and $(u, \tilde{u}) \in L^2(\Sigma_i)^s \times L^2(\Sigma_i)^s$. Then we denote the usual inner-product by
\[
(u, \tilde{u})_{\Sigma_i} \equiv \int_{\Sigma_i} u \cdot \tilde{u} \, d\sigma,
\]
where $u \cdot \tilde{u} \equiv u \tilde{u}$ in case $s = 1$.

The lateral boundaries are $\Gamma^A_L = \partial \Omega_A \setminus (\Gamma_t \cup \Gamma_I)$ and $\Gamma^S_L = \partial \Omega_S \setminus (\Gamma_b \cup \Gamma_I)$. The velocities $u_i$ will have components denoted by $(u_{1,i}, \ldots, u_{d,i})$ for $i = A, S$. Generic vectors $u$ have components $(u_1, \ldots, u_d)$.

**Definition 3.** Denote by $e_i$, $1 \leq i \leq d$, the reference points $e_i = x \iff x_j = \delta_{ij}$, where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. Weak spaces are defined with
Lemma 1. Let \( q \in C^\infty(\overline{\Omega_i}) \) \( |q(x) = q(x + e_j), \) if \( x_j = 0, \ 1 \leq j < d \), \( i \in \{A, S\} \).

\[ \Theta_i \equiv \text{cl}_{H^1} \{C_{\text{per}, i}\}, \ i \in \{A, S\}. \]

\( X_A \equiv \text{cl}_{(H^1)^d} \left\{ u \in (C_{\text{per}, A})^d \mid u_d = 0 \text{ on } \Gamma_i \cup \Gamma_T \right\}. \)

\( X_S \equiv \text{cl}_{(H^1)^d} \left\{ v \in (C_{\text{per}, S})^d \mid v = 0 \text{ on } \Gamma_b, \ v_d = 0 \text{ on } \Gamma_T \right\}. \)

\( P_i \equiv \left\{ p \in L^2(\Omega_i) \mid \int_{\Omega_i} p \, dx = 0 \right\}, \ i \in \{A, S\}. \)

\( V_i \equiv \{ u \in X_i \mid (\nabla \cdot u, p)_i = 0, \ \forall p \in P_i \}, \ i \in \{A, S\}. \)

Some bilinear forms are needed for the weak problem formulation.

**Definition 4.** Let \( i = A \) or \( i = S \). Given functions \( u, \tilde{u} \in (H^1(\Omega_i))^d \), define

\[ a_i(u, \tilde{u}) = \frac{2}{\text{Re}_i} \int_{\Omega_i} \text{V}(u)^H : \nabla \tilde{u} \, dx + \frac{2}{\text{Re}_i} \int_{\Omega_i} \text{V}(\tilde{u})^\perp : \nabla \tilde{u} \, dx. \]

**Definition 5.** Let \( i = A \) or \( i = S \). Given functions \( \theta, \tilde{\theta} \in H^1(\Omega_i) \), define

\[ b_i(\theta, \tilde{\theta}) = \frac{1}{\text{Re}_i^H \text{Pr}_i} \sum_{j=1}^{d-1} \int_{\Omega_i} \frac{\partial \theta}{\partial x_j} \frac{\partial \tilde{\theta}}{\partial x_j} \, dx + \frac{1}{\text{Re}_i^\perp \text{Pr}_i} \int_{\Omega_i} \frac{\partial \theta}{\partial x_d} \frac{\partial \tilde{\theta}}{\partial x_d} \, dx. \]

Coercivity for \( a_i(\cdot, \cdot) \) follows from the next result. The proof appears in [4].

**Lemma 1.** Let \( i = A \) or \( i = S \). Given functions \( \tilde{u} \in (H^1(\Omega_i))^d \) and \( u \in V_i \), it holds that

\[ a_i(u, \tilde{u}) = \frac{1}{\text{Re}_i^H} \int_{\Omega_i} (\nabla u)^H : (\nabla \tilde{u})^H \, dx + \frac{1}{\text{Re}_i^\perp} \int_{\Omega_i} (\nabla u)^\perp : (\nabla \tilde{u})^\perp \, dx. \]

Coercivity-related constants are defined as follows.

**Definition 6.** Let \( i = A \) or \( i = S \). Constants \( \epsilon_i > 0 \) and \( \delta_i > 0 \) are defined as

\[ \epsilon_i = \min \left\{ \frac{1}{\text{Re}_i^H}, \frac{1}{\text{Re}_i^\perp} \right\} \quad \text{and} \quad \delta_i = \min \left\{ \frac{1}{\text{Re}_i^H \text{Pr}_i}, \frac{1}{\text{Re}_i^\perp \text{Pr}_i} \right\} = \frac{\epsilon_i}{\text{Pr}_i}. \]

The following identity is useful to analyze discrete time derivatives later.

**Lemma 2.** Given equal-size vectors \( u, v \) and \( w \),

\[ \frac{1}{2} (3u - 4v + w) \cdot u = \frac{1}{4} (|u|^2 + |2u - v|^2) - \frac{1}{4} (|v|^2 + |2v - w|^2) + \frac{1}{4} |u - 2v + w|^2. \]

The next result is used to analyze stability. The proof appears in [4].

**Lemma 3** (Friction monotonicity). Define \( T : \mathbb{R}^d \to \mathbb{R}^d \) by \( T(x) = |x|x \), for all \( x \in \mathbb{R}^d, \ d \in \mathbb{N} \). Then \( T \in C^1(\mathbb{R}^d, \mathbb{R}^d) \) and

\[ (T(x) - T(y)) \cdot (x - y) \geq \frac{1}{4} |x - y|^3, \quad \forall x, y \in \mathbb{R}^d. \]

We conclude with some well-known technical lemmas.
Lemma 4 (Poincaré inequalities). There exist fixed positive constants $C_p^0$ and $C_p$ such that
\[
\|u\|_S \leq C_p^0 \|\nabla u\|_S, \forall u \in X_S,
\]
\[
\|u_d\|_A \leq C_p^0 \|\nabla u_d\|_A, \forall u = (u_1, \ldots, u_d) \in X_A,
\]
\[
\|\theta_i - \bar{\theta}_i\|_i \leq C_p \|\nabla \theta_i\|_i, \forall \theta_i \in \Theta_i, \ i \in \{A,S\}.
\]

Lemma 5 (Skew-symmetry). Given $i \in \{A,S\}$, $u \in V_i$, $v \in X_i$, $w \in X_i$, $\theta \in \Theta$ and $\psi \in \Theta$,
\[
(u \cdot \nabla v, w)_i = -(u \cdot \nabla w, v)_i \quad \text{and} \quad (u \cdot \nabla \psi, \psi)_i = -(u \cdot \nabla \psi, \theta)_i.
\]
In particular, $(u \cdot \nabla v, v)_i = (u \cdot \nabla w, w)_i = 0$ and $(u \cdot \nabla \psi, \theta)_i = (u \cdot \nabla \psi, \psi)_i = 0$.

3.1. Weak convection model. Here we present the weak model of heat convection, upon which we will build our finite-element-based numerical methods. Due to differences in parameters and boundary conditions, we present the equations for weak solutions separately for each fluid domain. We search for $(u_A, p_A, \theta_A) \in (X_A, P_A, \Theta_A)$ such that for all $t > 0$,
\[
(\frac{\partial u_A}{\partial t} + u_A \cdot \nabla u_A, \bar{u})_A + a_A (u_A, \bar{u}) - (p_A, \nabla \cdot \bar{u})_A
\]
\[
(\frac{\partial \theta_A}{\partial t} + u_A \cdot \nabla \theta_A, \bar{\theta})_A + (\frac{c_A p_A}{\rho_A})^{-1} (F(u_A, u_S; \theta_A, \theta_S), \bar{\theta})_I \Gamma_1
\]
\[
+ C_A (U \cdot \bar{u})_A - \theta_A, \bar{\theta})_{\Gamma_1} + b_A (\theta_A, \bar{\theta})_A = 0, \forall \bar{\theta} \in \Theta_A.
\]
We search for $(u_S, p_S, \theta_S) \in (X_S, P_S, \Theta_S)$ such that for all $t > 0$,
\[
(\frac{\partial u_S}{\partial t} + u_S \cdot \nabla u_S, \bar{u})_S + a_S (u_S, \bar{u}) - (p_S, \nabla \cdot \bar{u})_S
\]
\[
(\frac{\partial \theta_S}{\partial t} + u_S \cdot \nabla \theta_S, \bar{\theta})_S - (\frac{c_S p_S}{\rho_S})^{-1} (F(u_A, u_S; \theta_A, \theta_S), \bar{\theta})_I \Gamma_1
\]
\[
+ b_S (\theta_S, \bar{\theta})_S = 0, \forall \bar{\theta} \in \Theta_S.
\]

Note that energy is put into the model through boundary fluxes; on $\Gamma_I$ the flux $F(u_S, u_S; \theta_S, \theta_S)$ includes solar heating, and the term with $\theta_i$ in $[19]$ allows heat to enter or escape through the model top. Weak solutions must be stable, in the sense that solutions cannot blow up in finite time. The proof is excluded here for brevity.

4. Full model discretization

We choose a standard, conforming finite element discretization in space. Since our interest is the temporal representation of fluxes, we circumvent any need for interpolation in space by restricting to equal-order elements and matching meshes.
on the fluid-fluid interface. After introducing the finite-element framework, the remainder of this section rigorously defines the sequential and concurrent time stepping algorithms.

4.1. Finite element method for space. For \( i = \mathcal{A} \) and \( i = \mathcal{S} \), let \( \tau_i^h \) denote conforming, quasi-uniform families of meshes for \( \Omega_i \), consisting of affine-equivalent triangles for \( d = 2 \) or tetrahedra for \( d = 3 \), with maximum element diameter \( 0 < h_i < 1 \). We define \( h = \max\{h_\mathcal{A}, h_\mathcal{S}\} \). Our periodic boundary conditions require that the mesh nodes on opposing lateral boundaries are compatible. A point \( \mathbf{x} \) is a mesh node if and only if \( \mathbf{x} + e_j \) is also a mesh node (refer to the notation in Definition 3). As mentioned above, we require the meshes for both subdomains to match on the interface, \( \Gamma_\mathcal{A} \).

Let \( P_k(S) \) denote polynomials of order \( k \) on a domain \( S \subset \mathbb{R}^d \) and define
\[
Z_k(\tau_i^h) \equiv \{p : \Omega_i \to \mathbb{R} | p|_E \in P_k(E), \forall E \in \tau^h_i \}.
\]
The necessary finite element spaces are defined as follows.
\[
\Theta_i^h \equiv Z_2(\tau_i^h) \cap \Theta_i, \quad X_i^h \equiv (Z_2(\tau_i^h))^d \cap X_i, \quad P_i^h \equiv Z_1(\tau_i^h) \cap P_i \cap C^0(\overline{\Omega}_i).
\]
The pairing of velocity and pressure is the well-known Taylor-Hood space, which is known to provide stability for the pressure.

The spaces \( X_i^h \) are not divergence-free, as required for exact mass conservation. The divergence is controlled using grad-div stabilization (see e.g. [13]). Define parameters \( \gamma_i \gg 1 \); these appear in Section 4.6 as penalization weights for the divergence of the velocity. Also, the skew-symmetry property for the standard convection terms does not hold for Taylor-Hood, which can lead to an incorrect energy balance. This is handled in a standard way, by using the following operators.

**Definition 7.** Given \( i \in \{\mathcal{S}, \mathcal{A}\} \) and any velocities \( u_i, v_i \) and \( w_i \) in \( X_i \), we define
\[
\Lambda_i(u_i, v_i, w_i) = \frac{1}{2}(u_i \cdot \nabla v_i, w_i)_i - \frac{1}{2}(u_i \cdot \nabla w_i, v_i)_i.
\]
Given any \( \theta_i \) and \( \phi_i \) in \( \Theta_i \), the same notation is also used to denote
\[
\Lambda_i(u_i, \theta_i, \phi_i) = \frac{1}{2}(u_i \cdot \nabla \theta_i, \phi_i)_i - \frac{1}{2}(u_i \cdot \nabla \phi_i, \theta_i)_i.
\]
When \( u_i \in V_i \), \( \Lambda_i(u_i, v_i, w_i) = (u_i \cdot \nabla v_i, w_i)_i \) and \( \Lambda_i(u_i, \theta_i, \phi_i) = (u_i \cdot \nabla \theta_i, \phi_i)_i. \)

4.2. Time reconstruction operators. Let \( N_\mathcal{A} \) and \( N_\mathcal{S} \) be positive integers. We define time step sizes \( \Delta t_\mathcal{A}, \Delta t_\mathcal{S} \) and
\[
(\text{coupling time step}) \quad \Delta t = N_\mathcal{A}\Delta t_\mathcal{A} = N_\mathcal{S}\Delta t_\mathcal{S}.
\]
We restrict \( \Delta t_\mathcal{A} \leq \Delta t_\mathcal{S} \) to be consistent with air-sea dynamics. Let \( t^n = n\Delta t \) for any integer \( n \). We refer to \([t^n, t^{n+1}]\) as a coupling interval. Locally, we define
\[
t^{n+j/N_i} = \left(n + \frac{j}{N_i}\right)\Delta t = t^n + j\Delta t_i, \quad 0 \leq j \leq N_i, \quad i \in \{\mathcal{A}, \mathcal{S}\}.
\]
Given \( \eta_i : \Omega_i \times [t^n, t^{n+1}] \to \mathbb{R} \), we denote local approximations by \( \eta_i^{n,j} \approx \eta_i(t^{n+j/N_i}) \).

Our algorithms also require some extrapolation and interpolation operators. In terms of the data \( \eta_i^{n,j} \), we first define a local, second-order extrapolation operator
by
\[
E_n^{i,j+1} = \begin{cases} 
2 \eta_i^{n,j} - \eta_i^{n-1,N_i-1}, & j = 0, \\
2 \eta_i^{n,j} - \eta_i^{n,j-1}, & 1 \leq j < N_i.
\end{cases}
\]
The Lagrange interpolant, denoted by \( P_i \eta_i^n : [t^n, t^{n+1}] \rightarrow \mathbb{R} \), is the unique polynomial of order \( N_i \) that satisfies
\[
P_i \eta_i^n \left( t^{n+j/N_i} \right) = \eta_i^{n,j}, \quad j = 0, 1, \ldots, N_i.
\]
This interpolant is used to define another, least-squares reconstruction operator, denoted by \( R_i^{(k)} \eta_i^n : [t^n, t^{n+1}] \rightarrow \mathbb{R} \). It is the unique polynomial of order \( k \) that satisfies
\[
\int_{t^n}^{t^{n+1}} t^j R_i^{(k)} \eta_i^n(t) \, dt = \int_{t^n}^{t^{n+1}} t^j P_i \eta_i^n(t) \, dt, \quad j = 0, 1, \ldots, k.
\]
The coefficients of \( R_i^{(k)} \eta_i^n \) may be computed on-the-fly as each \( \eta_i^{n,j} \) is calculated by using pre-computed weights, without storing the data at multiple times. Legendre polynomials form a basis for the least-squares reconstruction. See [12] for details.

### 4.3. State, numerical flux and iteration notations.

Given a coupling interval \([t^n, t^{n+1}]\), the state approximations at the times \( t^{n+j/N_i} \), \( j = 0, 1, \ldots, N_i \), are \((w_i^{n,j}, q_i^{n,j}, \phi_i^{n,j}) \in (X_i^h, P_i^h, \theta_i^h)\). Let \( m = 0, 1, \ldots \) be an iteration index. We reserve \( m = 0 \) for flux initialization data, as defined in Section 4.4. States \((w_i^{n,j}, q_i^{n,j}, \phi_i^{n,j}) \in (X_i^h, P_i^h, \theta_i^h)\) are computed for \( m = 1, 2, \ldots \); we specify details of the iterations later. After the last iteration, we denote the final approximation by dropping the \( m \)-notation, yielding the desired data \((w_i^{n,j}, q_i^{n,j}, \phi_i^{n,j})\).

We employ the following notation for downward numerical fluxes used to compute on \( \Omega_i \times [t^n, t^{n+1}] \) upon iteration \( m \):
\[
(\widehat{\mathcal{G}}_{i(m)}^{n,j}, \widehat{\mathcal{F}}_{i(m)}^{n,j}) \quad \text{(momentum, heat), for } j = 1, \ldots, N_i.
\]

### 4.4. Data initializations.

Given data over \([t^{n-1}, t^n]\), we provide the initializations needed to compute on \([t^n, t^{n+1}]\). Since the current states on \( \Omega_S \) will not be known when the fluxes are updated, this data is extrapolated from the previous interval. Choose a non-negative integer \( k \). Given data \((w_S^{n-1,j}, \phi_S^{n-1,j}) \in (X_S^h, \theta_S^h)\) for \( j = 0, 1, \ldots, N_S \), we initialize
\[
(w_S^{n,j}, \phi_S^{n,j}) = (R_S^{(k)}w_S^{n-1}, R_S^{(k)}\phi_S^{n-1})|_{t=t^{n+j/N_S}}
\]
for \( j = 1, \ldots, N_S \). Note \( j = 0 \) corresponds to the initial time \( t^n \) for the coupling interval, where the states are known already.

For the concurrent method, the fluxes are specified over the interval \([t^n, t^{n+1}]\) before computing the current states. Initially, this is done by extrapolating from flux data on the previous interval, which is not the same data used to compute the previous states. For accuracy, fluxes are updated after new states are computed; the recomputed fluxes are denoted using the following \( (\cdot) \)-notation (see Section 4.5).

Given some flux data \((\widehat{\mathcal{G}}_{A}^{-n,j}, \widehat{\mathcal{F}}_{A}^{-n,j})\) for \( j = 0, 1, \ldots, N_A \), we initialize (extrapolate)
\[
(\widehat{\mathcal{G}}_{A(0)}^{-n,j}, \widehat{\mathcal{F}}_{A(0)}^{-n,j}) = (R_A^{(k)}\widehat{\mathcal{G}}_{A}^{-n-1}, R_A^{(k)}\widehat{\mathcal{F}}_{A}^{-n-1})|_{t=t^{n+j/N_A}}, \quad \text{for } j = 1, \ldots, N_A.
\]
4.5. Numerical fluxes. Computations are performed internally for each subdomain using standard, implicit time stepping described in Section 4.3. Therefore, in order to compute states on $\Omega_i$ at time step $j$, the numerical fluxes must be defined at time $t^{n+j/N_i}$. We assume the initializations of Section 4.4 and define the fluxes now for an iteration index $m \geq 1$.

In the sequential configuration, the states on $\Omega_A$ are computed before those for $\Omega_S$. Fluxes are computed semi-implicitly in this case. States $(w_{A(m)}^{n,j}, q_{A(m)}^{n,j})$ are computed using the numerical flux

$$G_{A(m)}^{n,j} = \kappa \left| w_{A(m)}^{n,j} - w \right| \left( w_{A(m)}^{n,j} - w \right),$$

for $j = 1, \ldots, N_A$. Similarly, states $\phi_{A(m)}^{n,j}$ are computed using the numerical flux

$$F_{A(m)}^{n,j} = F(w_{A(m)}^{n,j}, w; \phi_{A(m)}^{n,j}, \phi),$$

where $\phi = R_S(\phi_{S(m-1)}^{n}) (t^{n+j/N_A})$. The sea states are then computed using the fluxes

$$G_{S(m)}^{n,j} = R_A(F_{A(m)}^{n,j}) (t^{n+j/N_S}),$$

for $j = 1, \ldots, N_S$. In the sequential case, the $k$-th order least-squares reconstruction in (31)-(32) guarantees that the first $k$ moments of the fluxes are conserved in the sense of (23) on each coupling interval.

The concurrent method uses an explicit specification of the fluxes to guarantee conservation even if there is no iteration. To compute states at times $t^{n+j/N_i}$ on either subdomain at iteration $m \geq 1$, we use fluxes

$$\left( \hat{G}_{i(m)}^{n,j}, \hat{F}_{i(m)}^{n,j} \right) = (R_A(\hat{G}_{A(m-1)}^{n}), R_A(\hat{F}_{A(m-1)}^{n})) |_{t = t^{n+j/N_i}},$$

for $j = 1, \ldots, N_i$. In case $m = 1$, the right-hand side of (33) is defined by (26). We then recompute the fluxes for the next iteration, or to extrapolate to the next coupling interval, by using these formulæ:

$$\hat{G}_{A(m)}^{n,j} = \kappa \left| w_{A(m)}^{n,j} - w \right| \left( w_{A(m)}^{n,j} - w \right),$$

for $j = 1, \ldots, N_A$. The data in (35) and (37) is evaluated from iteration $m - 1$, not $m$, for two reasons: (1) to avoid waiting to compute fluxes until the current sea states are available (minimizing CPU idle time in a concurrent setup), and (2) to avoid storing many air states until a new sea state is computed in case $N_A \gg N_S$. This method also conserves fluxes in the integral sense indicated by (33) and (23).
4.6. Full computational algorithms. This section provides the internal discretizations for each subdomain and the remaining details of the iteration scheme. In order to investigate some stability properties we introduce a posteriori stability parameters, defined below, that can be computed on-the-fly and only require the storage of a few scalars. We will use these to define a stopping criterion for iteration that guarantees energetic stability. Although the motivation for doing this is primarily to investigate the stability properties of the algorithms, it could also lead to a reduction in overall iteration, compared to using an accuracy-based criterion. We return to this point in the computational section.

Algorithm 1 (Iterate for stability). Choose a non-negative integer $k$. Given data $(w_{S}^{n-1,j}, q_{S}^{n-1,j}, \phi_{S}^{n-1,j}) \in (X_{S}^{h}, P_{S}^{h}, \theta_{S}^{h})$ for $j = 0, 1, \ldots, N_{S}$, we calculate the states $(w_{A}^{n,j}, q_{A}^{n,j}, \phi_{A}^{n,j}) \in (X_{A}^{h}, P_{A}^{h}, \theta_{A}^{h})$ for $j = 1, \ldots, N_{A}$, as follows.

1. For $m \geq 1$, initiate stability indicators $\eta^{(m)}_{i} = \psi^{(m)}_{i} = 0$ for $i = A, S$.
2. For $j = 0, \ldots, N_{A} - 1$ calculate $(w_{A}^{n,j+1}, q_{A}^{n,j+1})$ by solving

\[
\frac{1}{2 \Delta t_{A}} \left( 3w_{A}^{n,j+1} - 4w_{A}^{n,j} + w_{A}^{n-1,j}, \tilde{u} \right)_{A} + a_{A} \left( w_{A}^{n,j+1}, \tilde{u} \right)_{A} + \lambda_{A} \left( E_{A}w_{A}^{n,j+1}, w_{A}^{n,j+1}, \nabla \cdot \tilde{u} \right)_{A} - \left( \rho_{A}^{n,j+1}, \nabla \cdot \tilde{u} \right)_{A} + \gamma_{A} \left( \nabla \cdot \tilde{w}_{A}^{n,j+1}, \nabla \cdot \tilde{u} \right)_{A} = \left( E_{A}f_{A}^{n,j+1}, \tilde{u} \right)_{A}, \forall \tilde{u} \in X_{A}^{h},
\]

and $(\nabla \cdot w_{A}^{n,j+1}, q)_{A} = 0, \forall q \in P_{A}^{h}$,

where $w_{A}^{n,j} \equiv w_{A}^{n-1,N_{A}+j}$ for $j \leq 0$ is given.

3. For each $j$, update the stability parameter $\eta_{A}^{(m)} = \eta_{A}^{(m)} + \left( \rho_{A}^{n,j+1}, w_{A}^{n,j+1} \right)_{\Gamma_{I}}$.

4. Calculate $\phi_{A}^{n,j+1}$ by solving

\[
\frac{1}{2 \Delta t_{A}} \left( 3\phi_{A}^{n,j+1} - 4\phi_{A}^{n,j} + \phi_{A}^{n-1,j}, \tilde{\theta} \right)_{A} + a_{A} \left( w_{A}^{n,j+1}, \phi_{A}^{n,j+1} \right)_{A} + \lambda_{A} \left( E_{A}w_{A}^{n,j+1}, \phi_{A}^{n,j+1}, \nabla \cdot \tilde{u} \right)_{A} - \left( \rho_{A}^{n,j+1}, \phi_{A}^{n,j+1} \right)_{A} + \gamma_{A} \left( \nabla \cdot \tilde{w}_{A}^{n,j+1}, \nabla \cdot \tilde{u} \right)_{A} = \left( E_{A}f_{A}^{n,j+1}, \tilde{u} \right)_{A}, \forall \tilde{u} \in X_{A}^{h},
\]

where $\phi_{A}^{n,j} \equiv \phi_{A}^{n-1,N_{A}+j}$ for $j \leq 0$ is given.

5. For each $j$, update the stability parameter $\psi_{A}^{(m)} = \psi_{A}^{(m)} + \left( \rho_{A}^{n,j+1}, w_{A}^{n,j+1} \right)_{\Gamma_{I}}$.

6. For $j = 0, \ldots, N_{S} - 1$ calculate $(w_{S}^{n,j+1}, q_{S}^{n,j+1})$ by solving

\[
\frac{1}{2 \Delta t_{S}} \left( 3w_{S}^{n,j+1} - 4w_{S}^{n,j} + w_{S}^{n-1,j}, \tilde{u} \right)_{S} - \left( \rho_{S}^{n,j+1}, \nabla \cdot \tilde{u} \right)_{S} + a_{S} \left( w_{S}^{n,j+1}, \tilde{u} \right)_{S} - \frac{\rho_{S}}{\rho_{A}} \left( \theta_{A}^{n,j+1}, \tilde{u} \right)_{\Gamma_{I}} + \gamma_{S} \left( \nabla \cdot \tilde{w}_{S}^{n,j+1}, \nabla \cdot \tilde{u} \right)_{S} = \left( E_{S}f_{S}^{n,j+1}, \tilde{u} \right)_{S}, \forall \tilde{u} \in X_{S}^{h},
\]
where \( \mathbf{w}^{n+j}_{j(m)} = \mathbf{w}^{n+j}_{j} \) for \( j \leq 0 \) is given,

\[
\begin{align*}
(7) & \quad \text{For each } j \text{, update the stability parameter } \eta^{(m)}_{j} = \eta^{(m)}_{j} + (\mathbf{g}^{n+j}_{j(m)}, \mathbf{w}^{n+j}_{j})_{\Gamma_{z}}. \\
(8) & \quad \text{Calculate } \phi^{n+j+1}_{j} \text{ by solving} \\
& \quad \frac{1}{2} \Delta t_{S} \left( 3\phi^{n+j+1}_{S(m)} - 4\phi^{n+j}_{S(m)} + \phi^{n+j-1}_{S(m)} \right)_{S} + \Delta t_{S} \left( \mathbf{w}^{n+j+1}_{S(m)} + \phi^{n+j+1}_{S(m)} \mathbf{b} \right)_{\Gamma_{z}} \\
& \quad \text{where } \phi^{n+j}_{S(m)} = \phi^{n+j}_{S} \text{ for } j \leq 0 \text{ is given.} \\
(9) & \quad \text{For each } j \text{, update the stability parameter } \psi^{(m)}_{j} = \psi^{(m)}_{j} + (\mathbf{f}^{n+j+1}_{S(m)}, \phi^{n+j+1}_{S})_{\Gamma_{z}}. \\
(10) & \quad \text{Set } m \leftarrow m + 1 \text{ and repeat Steps 1-9 until} \\
(11) & \quad \text{Evolution over } [t^{n}, t^{n+1}] \text{ is complete; for } i \in \{\mathcal{A}, \mathcal{S}\} \text{ define} \\
& \quad (\mathbf{w}^{n+j}_{i}, q^{n+j}_{i}, \phi^{n+j}_{i}) = (\mathbf{w}^{n+j}_{i(m)}, q^{n+j}_{i(m)}, \phi^{n+j}_{i(m)}), \ j = 0, 1, \ldots, N_{i}.
\end{align*}
\]

The numerical fluxes are computed as defined in Section 4.5 so that for the concurrent method Steps 2-5 can be performed at the same time as Steps 6-9. The algorithm above conserves moments of the fluxes up to order \( k \) in time, in the sense of (23). The time stepping method used for the bulk fluids is known as BDF-2 (see e.g. [24]), which is expected to provide second-order accuracy in time for smooth solutions, not accounting for the accuracy of the flux computations. The time accuracy associated with the fluxes should be formally of order \( \Delta t^{k+1} \), in the sense of the truncation error associated with those terms. This stems from the accuracy of the underlying least-squares approach, which is well-known from the theory of orthogonal polynomials (see e.g. [18]). As a result, we expect that the time accuracy scales globally for the method as (roughly) \( \mathcal{O}(\Delta t^{2}) \), with \( q = \min\{2k+1\} \) for smooth solutions.

**Remark 1.** We emphasize that the main conclusions of our stability analysis, in particular the scaling of the time step size relative to the spatial grid size, could also have been made with other methods for the bulk fluids like backward-Euler or Crank-Nicolson. However, we wanted the higher accuracy often associated with BDF-2 over backward-Euler and we felt that the sort of half-time evaluations built into Crank-Nicolson would have made our analysis even more lengthy. Thus, BDF-2 is a preference, but not one that we feel is very critical for the main points of this paper.

The stability criterion is motivated by the properties of the governing equations and would hold automatically for monolithic algorithms, meaning the air and sea states being implicitly coupled at each time step. More generally, a time step restriction may be required. In the next section we study the stability and convergence of the iteration for a special case to provide intuition regarding the general algorithm. We also compare this algorithm in our computations to the following variant.
Algorithm 2 (Iterate for convergence). Choose a convergence tolerance $\delta > 0$. Modify Algorithm 1 as follows:

- Skip the computations of the stability indicators $\eta_i^{(m)} = \psi_i^{(m)}$.
- Replace the stopping criterion (44) with

\[
\int_{t_n}^{t_{n+1}} \int_{\Gamma_x} \left| R_S^{(k)} \mathbf{w}_S^{(m)} - R_S^{(k)} \mathbf{w}_S^{(m-1)} \right|^2 \, d\sigma \, dt \\
+ \int_{t_n}^{t_{n+1}} \int_{\Gamma_x} \left| R_S^{(k)} \phi_S^{(m)} - R_S^{(k)} \phi_S^{(m-1)} \right|^2 \, d\sigma \, dt < \delta^2.
\]

The convergence criterion (45) is chosen because it is relatively cheap to compute and leads to convergence of the fluxes, thus the states. Other tests would be stronger, but this test has proved sufficient in our computations.

5. Algorithm stability and convergence

In their full generality, we show our algorithms are stable when criterion (44) is satisfied. Note that while the criterion is not explicitly checked for Algorithm 2, stability is still assured under the same condition. However, the stability condition may not hold in general. We prove stability under a time step restriction, but only for a special case of the algorithms due to their complexity. The proof illustrates that, if the size of the coupling interval $\Delta t$ is small enough, then the iteration process must converge to a limit for which the fluxes are implicitly coupled and satisfy (44) trivially.

5.1. A time step restriction. Our stability analysis requires bounds for data on the interface in terms of the global norms of $L^2(\Omega_i)$. We use techniques of finite element analysis to lift off the interface, and some time step restrictions. In Section 5.1, let $g_i, i \in \{A, S\}$, be a finite element function of the form

\[
g \in \left\{ q \in (C^0(\overline{\Omega_i}))^d \mid q \in P_k(E), \forall E \in \tau_h^i \right\},
\]

for some positive integer $k$.

Cubic flux terms are bounded later in part by using the following result. It is an immediate consequence of the proof in [3].

Lemma 6. There exists a constant $C_{32}$ that depends on $i$, $k$ and $d$ but is independent of $h$ and $g$, such that

\[
\|g\|_{L^3(\Gamma_x)} \leq C_{32} h^{1-d/6} \|g\|_{L^2(\Gamma_x)}.
\]

The lifts also use the following well-known result (see e.g. [17]).

Lemma 7. There exists a constant $C_L$ that depends on $i$, $k$ and $d$ but is independent of $h$ and $g$, such that for any side $e$ of any element $E \in \tau_h^i$,

\[
\|g\|_{L^2(e)} \leq C_L h^{-1/2} \|g\|_{L^2(E)}.
\]

An analogous inverse inequality holds globally, but for functions within a certain class the $h$-dependence is more mild.
Definition 8. Given $i \in \{A, S\}$, denote by $S^h_i$ the set of elements with a side on the interface:

$$S^h_i \equiv \{ E \in \tau^h | |E \cap \Gamma_x| > 0 \}.$$

Given $d$, $k$ and some size parameter $\omega > 0$, we define classes $B_\omega = B_\omega(d, k, i)$ as

$$B_\omega = \left\{ g \mid \frac{1}{|S^h_i|} \int_{S^h_i} |g|^2 \, dx \leq \omega \frac{1}{|\Omega_i|} \int_{\Omega_i} |g|^2 \, dx \right\}.$$

In practice, the horizontal and vertical mesh scalings are quite different. To sharpen our analysis, we introduce the following parameter.

Definition 9. The thickness of the layer of elements along the interface is defined as

$$(48) \quad h^i \equiv \text{sup } \{ z \in \mathbb{R}^+ | S^h_A \cup S^h_S \subset \Gamma_x \times [-z, z] \}.$$

Since $|\Gamma_x| = 1$, the volumes of the sets $S^h_i$ are thus bounded by

$$(49) \quad |S^h_i| \leq h^i \leq h, \ i \in \{A, S\}.$$

For $g \in B_\omega$ it holds that

$$(50) \quad \|g\|^2_{L^2(S^h_i)} = \frac{\|g\|^2_{L^2(S^h_i)}}{\|g\|^2} \leq \frac{|S^h_i|}{|\Omega_i|} \omega \|g\|^2 \leq \frac{h^i}{\beta} \omega \|g\|^2.$$

Lemma 8. Choose any $\beta > 0$. Under the time step restriction

$$(51) \quad \Delta t \leq \beta^{-3/2} C_T h^{\frac{d+2}{2}}$$

it holds that

$$(52) \quad \|g\|_{L^2(\Gamma_x)} \leq \frac{1}{\sqrt{\beta} \Delta t^{1/3}} \|g\|.$$

Here, $C_T > 0$ is a finite constant that depends on $\Omega_i$ and $k$. Furthermore, if $g \in B_\omega$ then $(51)$ may be replaced by

$$(53) \quad \Delta t \leq \beta^{-3/2} C_T \left( \frac{\bar{h}}{\omega} \right)^{3/2} \frac{h^{d+2}}{(h^i)^{3/2}}.$$

Proof. Observe that by using $(46)$ and then $(47)$ we have

$$(54) \quad \Delta t^{2/3} \|g\|^2_{L^2(\Gamma_x)} \leq C_{32}^2 \Delta t^{2/3} h^{-\frac{d+2}{2}} \sum_{E \in \Gamma_x} \|g\|^2_{L^2(E \cap \Gamma_x)}$$

$$\leq (C_{32} C_L)^2 \Delta t^{2/3} h^{-\frac{d+2}{2}} \sum_{E \in \Gamma_x} \|g\|^2_{L^2(E)} = (C_{32} C_L)^2 \Delta t^{2/3} h^{-\frac{2d+4}{2}} \|g\|^2_{L^2(S^h_i)}.$$

Since $S^h_i \subset \Omega_i$ we may bound

$$\Delta t^{2/3} \|g\|^2_{L^2(\Gamma_x)} \leq \frac{1}{\beta} \left( \beta (C_{32} C_L)^2 \Delta t^{2/3} h^{-\frac{2d+4}{2}} \right) \|g\|^2.$$

The inequality $(52)$ follows under the restriction $(51)$ with $C_T = (C_{32} C_L)^{-3}$. One may apply the alternative bound $(50)$ for $(54)$ and carry the result through the above analysis to show the result holds under $(53)$. \hfill \Box
An important point of Lemma 8 is that since \( h_9 \leq h \), the time step condition for data \( g \in B_i \) is even less restrictive than the rougher bound

\[
\Delta t \leq \frac{\beta^{3/2} C_I}{\omega} \left( \frac{\tilde{H}}{\omega} \right)^{3/2} h^{\frac{d-1}{2}}.
\]

In terms of the scaling relationship between \( \Delta t \) and \( h \), the condition looks like either \( \Delta t = O(\sqrt{\tilde{H}}) \) for \( d = 2 \) or \( \Delta t = O(h) \) for \( d = 3 \). In our stability analysis, only the computed temperature or velocity states need to be in some classes \( B_\omega \), where \( \omega \) then takes on the meaning of a ratio between a mathematical, average state energy in the boundary layer of elements versus the average over the entire domain. This assumption can only fail if the ratio is unbounded, which would require an unbounded concentration of energy at the interface relative to the bulk fluid, which seems counter-intuitive. Therefore, we think it is a reasonable assumption to differentiate between the more standard time step restriction (51) versus the more realistic condition (53). The implied scaling (55) is sharpened significantly compared to the earlier analysis for fluid-fluid coupling methods with finite elements in [20].

5.2. Stability. We assume initialization data is available on a coupling interval \([t^n, t^{n+1}]\) with \( n = -1 \). The first of our two main theoretical results shows that the criterion (44) is sufficient to guarantee stability.

Theorem 1 (Algorithm 1 Stability). For \( i \in \{A, S\} \), let \( \phi_i^{l,j} = \phi_i^{(m)} \in \Theta_i^h \) satisfy (40) and (43), respectively, and let \( w_i^{l,j} = w_i^{(m)} \in X_i^h \) satisfy (38) and (41), respectively, for each \( j \in \{1, 2, \ldots, N_i-1\} \) and \( l \in \{0, 1, 2, \ldots, n_i\} \), for any \( 1 \leq n \) such that \( t^n \leq t_f < \infty \). Here, \( t_f \) is some final time of simulation. If the stability criteria (44) holds, then the states computed using Algorithm 1 satisfy:

\[
\sum_{i \in \{A, S\}} c_i \rho_i \left[ \| \phi_i^{n,N_i} \|^2 + \| 2\phi_i^{n,N_i} - \phi_i^{n,N_i-1} \|^2 \right]
+ 4 \sum_{l=0}^{n} \sum_{i \in \{A, S\}} c_i \rho_i \Delta t_i \sum_{j=0}^{N_i-1} \delta_i \left\| \nabla \phi_i^{l,j+1} \right\|^2 \leq C_\Theta
\]

and

\[
\sum_{i \in \{A, S\}} \rho_i \left[ \| w_i^{n,N_i} \|^2 + \| 2w_i^{n,N_i} - w_i^{n,N_i-1} \|^2 \right]
+ 2 \sum_{i \in \{A, S\}} \Delta t_i \rho_i \sum_{l=0}^{n} \sum_{j=0}^{N_i-1} \left\| \nabla w_i^{l,j+1} \right\|^2 \leq C_w,
\]

with the stability constants given by

\[
C_\Theta = t_f \left( \frac{2C_\theta^2}{U} + \frac{2C_{sol}^2}{U^2 C_{t\gamma}^2} \right) + \sum_{i \in \{A, S\}} c_i \rho_i \left[ \| \phi_i^{1,N_i} \|^2 + \| 2\phi_i^{1,N_i} - \phi_i^{1,N_i-1} \|^2 \right].
\]
For temperature, polynomial that satisfies

\[ C_w = \sum_{i \in \{\mathcal{A}, \mathcal{S}\}} \rho_i \left[ \left\| \mathbf{w}^{-1,N_i} \right\|^2 + \left\| 2\mathbf{w}^{-1,N_i} - \mathbf{w}^{-1,N_{i-1}} \right\|^2 \right] + 80t_f (C_p^0 \tilde{g})^2 \sum_{i \in \{\mathcal{A}, \mathcal{S}\}} \frac{\rho_i}{\epsilon_i} + 20 \max_{i \in \{\mathcal{A}, \mathcal{S}\}} \frac{\left( \hat{\beta}_i \right)^2}{c_i \delta_i \epsilon_i} \left( C_p + 2 \sum_{i \in \{\mathcal{A}, \mathcal{S}\}} \Delta t_i c_i \rho_i \delta_i \left( \left\| \nabla \phi^-_{i-1,N_i} \right\|^2 + \left\| \nabla \phi^-_{i-1,N_{i-1}} \right\|^2 \right) \right). \]

The proof is given in Appendix B.1

5.3. Iteration convergence. We will now analyze the special case for \( k = 1 \) with \( N_A = N_S = 1 \). In this case, \( R_i^{(1)} c_i^n : [t^n, t^{n+1}] \to \mathbb{R} \) is the unique linear polynomial that satisfies

\[
\int_{t^n}^{t^{n+1}} t^j R_i^{(k)} c_i^n(t) \, dt = \int_{t^n}^{t^{n+1}} t^j P_i c_i^n(t) \, dt, \quad j = 0, 1.
\]

In particular, this implies that

\[
R_i^{(1)} c_i^n(t^{j+1}) = c_i^{j+1}, \quad j = 0, 1.
\]

Algorithm 1 then reduces to solving the following:

Given data

\[
(w_i^{n-1}, q_i^{n-1}, \phi_i^{n-1}), (w_i^n, q_i^n, \phi_i^n) \in (X_i^h, P_i^h, \theta_i^h)
\]

we calculate

\[
(w_i^{n+1}, q_i^{n+1}, \phi_i^{n+1}) \in (X_i^h, P_i^h, \theta_i^h)
\]

by solving the following two sets of equations and iterating over \( m = 1, 2, \ldots \) until the stability criteria are satisfied.

For the velocity and pressure in \( \Omega_A \):

\[
\frac{1}{2 \Delta t} \left( 3w_{A(m)}^{n+1} - 4w_{A}^{n} + w_{A(n-1)}^{n}, \tilde{u} \right)_{\mathcal{A}} + \Lambda_A \left( 2w_{A}^{n} - w_{A(n-1)}^{n}, w_{A(m)}^{n+1}, \tilde{u} \right) + \gamma_A \left( \nabla \cdot w_{A(m)}^{n+1}, \nabla \cdot \tilde{u} \right)_{\mathcal{A}} + a_A \left( w_{A(m)}^{n+1}, \tilde{u} \right) - \left( q_{A(m)}^{n+1}, \nabla \cdot \tilde{u} \right)_{\mathcal{A}} + \left( G_{A(m)}^{n+1}, \tilde{u} \right)_{\Gamma} = \left( 2f_{A}^{n} - f_{A}^{n-1}, \tilde{u} \right)_{\mathcal{A}}, \quad \forall \tilde{u} \in X_A^h,
\]

and \( \left( \nabla \cdot w_{A(m)}^{n+1}, q \right)_{\mathcal{A}} = 0, \quad \forall q \in P_A^h \).

For temperature,

\[
\frac{1}{2 \Delta t} \left( 3\phi_{A(m)}^{n+1} - 4\phi_{A}^{n} + \phi_{A(n-1)}^{n}, \tilde{\theta} \right)_{\mathcal{A}} + \Lambda_A \left( 2\phi_{A}^{n} - \phi_{A(n-1)}^{n}, \phi_{A(m)}^{n+1}, \tilde{\theta} \right) + \left( c_{A} \rho_A A \right)^{-1} \left( f_{A(m)}^{n+1}, \tilde{\theta} \right)_{\Gamma} + b_A \left( \phi_{A(m)}^{n+1}, \tilde{\theta} \right) + C_t \left( U c_{A} \rho_A A \right)^{-1} \left( \phi_{A(m)}^{n+1} - \theta_i, \tilde{\theta} \right)_{\Gamma} = 0, \quad \forall \tilde{\theta} \in \Theta_A^h.
\]
For the velocity and pressure in $\Omega_S$:

$$\frac{1}{2\Delta t} \left( 3w_{S_{(m)}^{n+1}} - 4w_{S_{n-1}} + w_{S_{n-1}} \right) + \Lambda_{S} \left( 2w_{S_{n}} - w_{S_{n-1}} , w_{S_{(m)}^{n+1}} \right) + \gamma_{S} \left( \nabla \cdot w_{S_{(m)}^{n+1}} , \nabla \cdot \mathbf{u} \right) + a_{S} \left( w_{S_{n+1}}^{n+1} , \mathbf{u} \right) - \left( q_{S_{(m)}}^{n+1} , \nabla \cdot \mathbf{u} \right) = \left( 2\mathbf{f}_{S_{n}} - \mathbf{f}_{S_{n-1}}^{n-1} , \mathbf{u} \right), \forall \mathbf{u} \in \mathcal{X}_{S}^{h},$$

and

$$\left( \nabla \cdot w_{S_{(m)}^{n+1}} , q \right) = 0, \forall q \in \mathcal{P}_{S}^{h}. \tag{60}$$

For temperature,

$$\frac{1}{2\Delta t} \left( 2\rho_{S}^{n+1} - 2\rho_{S} + \phi_{S}^{n+1} , \mathbf{u} \right) + \Lambda_{S} \left( 2\rho_{S}^{n} - \phi_{S_{n}}^{n+1} , \mathbf{u} \right) - \left( \rho_{S_{(m)}}^{n+1} , \mathbf{u} \right) = \left( 2\mathcal{F}_{S_{n}} - \mathcal{F}_{S_{n-1}}^{n-1} , \mathbf{u} \right), \forall \mathbf{u} \in \Theta_{S}^{h}. \tag{61}$$

Note that the numerical fluxes for the states on $\Omega_S$ match the fluxes used on $\Omega_A$.

The numerical fluxes in the sequential mode are given by:

$$\mathcal{G}_{A_{(m)}^{n+1}} = \kappa \left( \mathbf{w}_{A_{(m)}^{n+1}} - \mathbf{w}_{S_{(m-1)}} \right) \left( \mathbf{w}_{A_{(m)}^{n+1}} - \mathbf{w}_{S_{(m-1)}^{n+1}} \right) \tag{62}$$

and in the concurrent mode, we list here the fluxes for $m \geq 2$ only:

$$\mathcal{F}_{A_{(m)}^{n+1}} = \mathcal{F}(\mathbf{w}_{A_{(m)}^{n+1}}, \mathbf{w}_{S_{(m-2)}^{n+1}}), \quad \mathcal{G}_{A_{(m)}^{n+1}} = \mathcal{G}(\mathbf{w}_{A_{(m)}^{n+1}}, \mathbf{w}_{S_{(m-2)}^{n+1}}) \tag{63}$$

The fluxes for $m = 1$ are different, but this detail is not important to discuss the limiting behavior as $m$ becomes large.

5.3.1. A Monolithic Iteration. We will now discuss an alternative, theoretical iteration that is used later for technical reasons in the proof of convergence of the iteration process in Algorithm 1. To this end, given data $\mathbf{w}_{j}^{n}$ and $\phi_{j}^{n}$, $j = 0, \ldots, n$, we define a monolithic version of the coupling by replacing the interface terms in (58) and (60) with

$$\left( \kappa \left| \mathbf{v}_{A_{(m-1)}}^{n+1} - \mathbf{v}_{S_{(m-1)}}^{n+1} \right| \left( \mathbf{v}_{A_{(m)}}^{n+1} - \mathbf{v}_{S_{(m)}}^{n+1} \right) , \mathbf{u} \right) \Gamma_{\mathcal{I}}$$

and in (59) and (61) with

$$\mathcal{F}(\mathbf{v}_{A_{(m-1)}^{n+1}}, \mathbf{v}_{S_{(m-1)}^{n+1}}), \mathcal{G}(\mathbf{v}_{A_{(m-1)}^{n+1}}, \mathbf{v}_{S_{(m-1)}^{n+1}}), \mathcal{H}(\mathbf{v}_{A_{(m-1)}^{n+1}}, \mathbf{v}_{S_{(m-1)}^{n+1}}) \Gamma_{\mathcal{I}}.$$

Here $(\mathbf{v}_{A_{(m)}^{n+1}}, \mathbf{s}_{(m)}^{n+1})$ are states generated using the monolithic formulation at the time $n + 1$.

We now show that the stability criterion (44) is satisfied trivially with this monolithic formulation, for all $m, n$. For the velocity states, note that

$$\frac{\Delta \mathcal{A}_{A_{(m)}^{n+1}}}{\Delta t} = \mathcal{A}_{A_{(m)}^{n+1}} - \mathcal{A}_{A_{(n)}^{n+1}} \tag{66}$$

$$= \mathcal{A}_{A_{(m)}} \kappa \left| \mathbf{v}_{A_{(m-1)}}^{n+1} - \mathbf{v}_{S_{(m-1)}}^{n+1} \right| \left| \mathbf{v}_{A_{(m)}}^{n+1} - \mathbf{v}_{S_{(m)}}^{n+1} \right| \Gamma_{\mathcal{I}} \frac{d \sigma}{d \sigma} \geq 0.$$
For the temperature states, use Young’s inequality to bound

\[ \Delta t_A\psi^{(m)}_A - \Delta t_S\psi^{(m)}_S = \Delta t \left( F(v^{(n+1)}_{A(m-1)}, v^{(n+1)}_{S(m-1)}; \xi^{n+1}_A(m), \xi^{n+1}_S(m), \xi^{n+1}_A - \xi^{n+1}_S) \right)_{\Gamma_x} \]

\[ = \Delta t \int_{\Gamma_x} \frac{C_{sol}}{U \theta_{REF}} \left( \xi^{n+1}_A(m) - \xi^{n+1}_S(m) \right) d\sigma + \Delta t \int_{\Gamma_x} \frac{C_{ir}}{U} \left( \xi^{n+1}_A(m) - \xi^{n+1}_S(m) \right)^2 d\sigma \]

\[ + \Delta t \int_{\Gamma_x} \frac{C_{ir} \rho A C_{scn}}{U} \left( v^{n+1}_A(m-1) - v^{n+1}_S(m-1) \right) \left( \xi^{n+1}_A(m) - \xi^{n+1}_S(m) \right)^2 d\sigma \]

\[ \geq \frac{\Delta t}{2} \int_{\Gamma_x} \frac{C_{ir}}{U} \left( \xi^{n+1}_A(m) - \xi^{n+1}_S(m) \right)^2 d\sigma - \Delta t \left| \Gamma_x \right| \frac{C_{sol}}{2U \theta_{REF}^2 C_{ir}} \geq -\Delta t \frac{C_{sol}}{2U \theta_{REF}^2 C_{ir}}. \]

The following lemma will be used in the proof of stability for Algorithm 1. The
monolithic states in the lemma are only theoretical reference points that depend
on a given set of data from previous times. It is critical to note that the previous
data represents the states computed with Algorithm 1 or Algorithm 2, not the
monolithic variant.

**Lemma 9.** Given data \( w^j_i \in X^h_i \) and \( \phi^j_i \in \Theta^h_i, \) \( j = 0, \ldots, n, \) satisfying the
assumptions of Theorem 4 for \( N_A = N_S = 1 \) with reconstruction order \( k = 1, \)
then \( v^{n+1}_{i(m)} \) and \( \xi^{n+1}_{i(m)} \) satisfy the following bounds:

\[ \sum_{i \in \{A,S\}} \rho_i \left( \|v_{i(m)}\|^2 + 2\|v_{i(m)} - w^i_n\|^2 \right) \leq C_v \]

and

\[ \sum_{i \in \{A,S\}} c_i \rho_i \left( \|\xi_{i(m)}\|^2 + \|2\xi_{i(m)} - \phi^i_n\|^2 \right) \leq C_\xi. \]

The bounding constants are given by

\[ C_v = 2C_w + 32g^2(\Delta t)^2 \sum_{i \in \{A,S\}} \left( \bar{\rho}_i + 4\tilde{\beta}_i^2 \right) \frac{4(\tilde{\beta}_i)^2}{c_i} C_{\phi} \]

and

\[ C_\xi = \frac{2\Delta t}{U} \left( C_{ir}^2 + \frac{C_{sol}^2}{\theta_{REF}^2} \right) C_\phi. \]

These constants are independent of \( m \) or \( n. \)

The proof is given in Appendix B.2.

### 5.3.2. Iteration convergence for Algorithm 2

Under a time step restriction, the following result proves that the stability criterion \( (44) \) must hold after a finite
number of iterations for our special case. That is because in the limit the fluxes
coincide with a certain (nonlinear) monolithic formulation. Then \( (44) \) holds by
arguments analogous to \( (66)-(67). \)

**Theorem 2.** Let \( N_A = N_S = 1 \) with reconstruction order \( k = 1, \) and let data
\( w^j_i \in X^h_i \) and \( \phi^j_i \in \Theta^h_i, \) \( j = 0, \ldots, n, \) be given that satisfy the assumptions
of Theorem 4. Assume that the monolithic reference states in Lemma 4 satisfy

\[ (70) \]

\[ 0 < b_1 \leq \left( \kappa \left| v_{A(m-1)}^{n+1} - v_{S(m-1)}^{n+1} \right| \left( v_{A(m)}^{n+1} - v_{S(m)}^{n+1} \right), v_{A(m)}^{n+1} - v_{S(m)}^{n+1} \right)_{\Gamma_x}, \]
and
\[
0 < b_2
\]
\[
\text{(71)} \quad \frac{C^2_{\text{sol}}}{2U\theta_{\text{REX}}C_{\text{ir}}}
\]
for each iteration, \( m \). If the time step \( \Delta t \) is small enough, then there exists an index \( m = m_0 < \infty \) for which the states \( w_{i(m)} = w_{i}^{n+1} \) and \( \phi_{i(m)} = \phi_{i}^{n+1} \) satisfy (44).

The proof is given in Appendix B.3.

Remark 2. In practice, the conditions (70) - (71) will hold trivially, since they are equivalent to the assumption that the jump of the monolithic velocity and temperature reference states across \( \Gamma_I \) is nonzero, as measured in \( L^1(\Gamma_I) \) for example.

Remark 3. The time step restriction comes directly from Lemma 8 and the stability constant is determined from the numbers \( \alpha \) and \( \beta \) that appear in the proof given in Appendix B.3. The dependence of the time step on any parameter of interest is explicitly tracked through our proofs. The relationship between \( \Delta t \) and \( h \) was already discussed in the remainder of Section 5.1 after Lemma 8.

6. A computational study

We compare the algorithms in various configurations by measuring the errors in the quantities of interest \( KE_i(t) \equiv \rho_i ||w_i||^2 \) for \( i \in \{A, S\} \). A reference solution is computed to provide the target kinetic energy values. The algorithms are then compared by varying the size of the coupling interval, the number of internal substeps, the cutoff criterion for the iteration process and the order, \( k \), of the reconstruction operator. We discuss the observed stability and accuracy properties.

We refer to the dimensional form of the model with vertical scaling \( H = 200 \) m and horizontal scaling \( L = 2000 \) m. In all the computations we use uniform triangulations across both subdomains with horizontal and vertical mesh spacings of \( 10 \) m. Our algorithms are implemented using the finite element software FreeFem++ [9] with \( d = 2 \) dimensions in space.

6.1. A model problem and reference solution. All computations begin from time \( t = 0 \) with the velocities set to zero and a uniform temperature of \( 285K \) across both subdomains. The model top reference temperature is also \( \theta_t = 285K \). Energy is put into the system by introducing a heat flux through the solar term. We set \( C_{\text{sol}} = (1 - 0.05) \cdot 0.75 \cdot 340 \cdot \cos(\pi (x - 1000)/1000) \) Watts per meter, where 0.05, and 0.75 \cdot 340 are representative values of sea surface albedo and the net downward flux of solar radiation under cloudless conditions, respectively [14]. This induces non-uniform heating that drives convective currents.

The Reynolds numbers of the flows are initially zero and grow in time. We slow the growth by choosing viscosity parameters larger than the values for air and sea water. This delays the transition to turbulence so that we may perform a longer time integration, since we do not use any turbulence modeling. The viscosity values are:
\[
\nu_A^H = 0.1, \quad \nu_S^H = \frac{\nu_A^H}{15}, \quad \nu_A^\perp = 20\nu_A^H, \quad \nu_S^\perp = 20\nu_S^H.
\]
Other physical parameters are chosen using the values for air and sea water at sea level; these and other remaining model parameters are specified in Appendix A.

A reference solution was computed by using a nonlinear, monolithic version of the coupling that has zero local truncation error. That is, we set $\Delta t_A = \Delta t_S = \Delta t$ so that the air and sea codes evolve together, implicitly coupled at every time step by calculating the fluxes to satisfy precisely the coupling conditions in Section 2.

We set the final time of simulation to be 8 hours. After this time, there is a transition toward turbulence and we cannot resolve the solution with our code. The monolithic solution at the final time is shown in Figure 5 and Figure 6.

Plots of the kinetic energy over time were found to be visually identical when using time step sizes of $\Delta t = 0.5$ and $\Delta t = 2$ seconds. The kinetic energy plots are shown in Figure 7 for $\Delta t = 0.5 \text{s}$. Note that the fluid on $\Omega_A$ has a faster response to the forcing, but the total kinetic energy is eventually larger for the fluid on $\Omega_S$ because $\rho_A \ll \rho_S$. This test problem has an important dynamical similarity with an atmosphere-ocean system. The faster response of the fluid on $\Omega_A$ places a restriction on the time step size needed for accuracy, while the global time scales of the system dynamics are made larger by the slower response of the fluid on $\Omega_S$.

### 6.2. The sequential mode

We observe the errors in the kinetic energies and discuss the stability and accuracy for the sequential algorithm upon varying the
Figure 7. The kinetic energies of the fluids show the faster response of the top fluid and the larger energy eventually attained for the bottom fluid.

Figure 8. Sequential mode. Errors in kinetic energies of the fluids for \( k = 0 \) with \( \Delta t = 32 \) s (solid) and \( \Delta t = 8 \) s (dash).

Figure 9. Sequential mode. Errors (J) for \( k = 1 \) with \( \Delta t = 32 \) s (solid) and \( \Delta t = 8 \) s (dash).

Figure 10. Sequential mode. Errors for \( k = 2 \) with \( \Delta t = 32 \) s (solid) and \( \Delta t = 8 \) s (dash).

coupling step size, \( \Delta t \), and the reconstruction order, \( k \in \{0, 1, 2\} \). We also compare the strategies of iterating for stability (Algorithm 1) versus iterating for convergence (Algorithm 2). In every computation we use \( N_A = N_S \) with \( \Delta t = 8 \) s and 4 substeps, or \( \Delta t = 32 \) s and 16 substeps. Then the data reconstructions are not just pointwise interpolants and the internal time steps are 2 seconds even for different \( \Delta t \).

Given a coupling interval \([t^n, t^{n+1}]\), let the velocities computed using the monolithic method be denoted by \( \hat{w}_i^{n,j} \) for each internal time index \( j = 0, 1, \ldots, N_i \). Recall that the velocities computed in Section 4.6 are denoted by \( w_i^{n,j} \). We define the error in computing the kinetic energies for the fluids as

\[
\Delta K E_i \left( t^{n+j/N_i} \right) \equiv \rho_i \left\| \hat{w}_i^{n,j} \right\|_2 - \rho_i \left\| w_i^{n,j} \right\|_2, \quad j = 0, \ldots, N_i.
\]
The errors using Algorithm 1 are plotted in Figure 8 ($k = 0$), Figure 9 ($k = 1$) and Figure 10 ($k = 2$). No iteration was required for stability in any of these cases. There is a significant decrease in error with larger $k$, except near the final time for the fluid in $\Omega_A$. This is because the internal time step size is not in the asymptotic regime for the dynamics in $\Omega_A$ near this time, which can be verified by comparing the kinetic energy computed using the monolithic method with $\Delta t = 0.5 s$ versus $\Delta t = 2 s$. This comparison reveals an oscillatory error on the order of $0.08 J$ for the reference data with a 2-second time step (not shown here), whereas the corresponding errors in Figure 8-Figure 10 are around $0.2 - 0.3 J$. On the other hand, for the fluid on $\Omega_S$ the error decreases significantly when $k$ increases, for the entire simulation time.

We note that the scale of the kinetic energy is many orders of magnitude larger than the errors in all of the above cases, up to and including at the final time. This indicates that the length of the coupling interval can be much larger than the internal time step sizes of the fluid modules, which are restricted by the dynamics of the respective fluids, while still retaining accuracy and stability without any iteration.

We repeated the computations for $\Delta t = 32 s$ with $k = 0$ and $k = 2$, but this time using Algorithm 2 (iterate for convergence) with stopping parameters $\delta = 10^{-5}$ and $\delta = 10^{-11}$. The results are shown in Figure 11 for $k = 0$, and Figure 12 for $k = 2$. In our algorithms, we have used $m$ as an iteration index, but $m$ means specifically the number of invocations of the fluid solvers. Thus, the algorithms are only truly “iterative” for $m > 1$.

For case $k = 0$ and $\delta = 10^{-5}$ there is a significant decrease in error with $m \leq 2$. The effect is more pronounced for KE$_A$ than for KE$_S$. Since the fluxes are first-order accurate for $k = 0$, an analogous result could only be achieved with $m < 2$ by decreasing the time step size to less than half. This would require more work, so the iterations may improve the efficiency (not accounting for the increase in memory usage to implement iteration). On the other hand, with $\delta = 10^{-11}$ we found $4 \leq m \leq 21$ was needed on each coupling interval, without any significant decrease in the error compared to $\delta = 10^{-5}$.

For case $k = 2$ a significant decrease in error is observed when increasing from $m = 1$ to $m \leq 2$ ($\delta = 10^{-5}$), and again to $m \leq 6$ ($\delta = 10^{-11}$), but not near the final time. In contrast to the case $k = 0$, less iterations were required to reach the cutoff of $\delta = 10^{-11}$ and the error decreased more in proportion up until around $t = 4$ hours. At later times, there was no significant decrease in overall error upon iteration. In particular, we note that the oscillatory errors persist. Theoretically,
these might be suppressed by pushing $\delta$ closer to the size of numerical round-off error, but at great computational cost.

For $k = 0$ the effect of using little or no iteration was more significant than for $k = 2$. But further iteration for $k = 0$ pushes the result to a low-order reconstruction of the reference state, which places a limit on the accuracy. However, since the reconstruction error may be smaller with larger $k$, iteration can push the result closer to the reference state for higher-order reconstructions. Most importantly, the reduction in error associated with an increase in $k$ is much more than that associated with iteration.

6.3. The concurrent mode. Given the same values of $k$ and the same choices of internal time steps, the computed solutions for the concurrent and sequential modes have the same theoretical limit as the iterations converge. Thus, we focus instead on the differences in the stability and accuracy properties between the two modes when using Algorithm [1]. That is, when iterating just for stability. The concurrent method fails when $\Delta t = 32 \text{ s}$ (16 substeps), for the cases with $k > 0$. This is illustrated in Figure [13]. Although the low-order coupling remains stable in this case with only $m = 1$ required, the error is quite large.
The corresponding results for $\Delta t = 8\, s$ (4 substeps) are shown in Figure 14. The methods are stable, but only weakly for $k = 1$, which does not have good accuracy. The best accuracy is observed for $k = 2$. Cases $k = 0$ and $k = 2$ required only $m = 1$; case $k = 1$ required $1 \leq m \leq 2$. These results indicate that the coupling interval needs to be smaller to enter the asymptotic regime for the concurrent mode. In order to investigate this, we present results for $\Delta t = 2\, s$ with 4 substeps for both fluids. These are shown in Figure 15 for $k = 0$ and in Figure 16 for $k = 1, 2$. The methods are stable with $m = 1$ in each of these latter cases. The accuracy is much better for $k > 0$ than for $k = 0$. The results with $k = 1$ are slightly better than for $k = 2$, so $\Delta t = 2\, s$ is probably just larger than necessary to enter the asymptotic regime for the concurrent mode. Since the methods using the sequential mode were stable and had good accuracy with $\Delta t = 32\, s$, we conclude that the asymptotic regime for the concurrent mode requires $\Delta t$ at least one order of magnitude smaller.

7. Summary and conclusions

We have investigated methods for coupling two fluids across an interface with different time steps, motivated by atmosphere-ocean interaction codes that use separate code modules for each fluid. The focus is on the time representation of flux-form boundary conditions. We use least-squares polynomial flux reconstructions to couple the air and sea modules over a time interval. The high formal accuracy of these flux calculations is not shared by methods in application. The least-squares approach may reduce aliasing errors and numerical fluctuations that can occur with pointwise interpolants. Also, our flux computations exactly conserve moments up to the order of reconstruction, in a time-integrated sense, even without any iteration. The algorithms we investigate have two primary classifications: sequential and concurrent, in reference to different code configurations for the air and sea modules.

We consider using iteration on a coupling interval for stability, and possibly for accuracy, in the manner of the methods in [12, 11]. To analyze and test the methods
we introduced a model of two coupled fluids with natural heat convection. We use finite element methods for the spatial discretization. A numerical analysis proved that the algorithms remain stable if a certain condition for the fluxes is satisfied, which can be checked computationally on-the-fly as an \textit{a posteriori} stability indicator. Furthermore, it was proved that this condition must be satisfied with a finite number of iterations on each coupling interval, under a time step restriction that scales no worse than $O(h)$ in 3D. This is a significant improvement over previous scaling analyses in the context of finite elements.

Computational tests illustrated the stability and accuracy properties as they relate to the choices of coupling mode (sequential versus concurrent), the length of the coupling interval, the number of internal time steps on a coupling interval, iterations performed and order of least-squares data reconstruction. We did not observe a need to iterate for stability in the sequential mode, whereas in the concurrent mode it was necessary for a large enough coupling interval in one case using a high-order reconstruction. When iteration did help to enforce stability, it did not yield better accuracy than was achieved with the lowest order method. It was necessary to take a small enough coupling interval for accuracy of the high-order methods. In this case, iteration was not needed for stability and did not improve the accuracy much compared to the results from increasing the reconstruction order. Low-order coupling was more sensitive to the first few iterations, but often converged more slowly for tight tolerances.

In future work, variations of our algorithms will be explored to handle some outstanding issues. In the sequential mode, there is a question of what may happen at even larger coupling intervals, but the internal time step sizes are restricted by the dynamics, so this translates to more substeps. We expect some novel strategy would be needed to handle a very large number of substeps, since the computation of reconstruction coefficients may become susceptible to round-off errors. In the concurrent mode, the asymptotic regime for the coupling interval appears to be at least an order of magnitude smaller than for the sequential mode. It is not currently known how to improve this property without excessive iteration and still preserve our conservation properties. Adaptive methods should be considered as well.

References

Appendix A. Computational test parameters

Physical parameters are chosen for air and sea water at sea level. We use SI units. Gravitational acceleration is set to $g = 9.81 \text{ m} \cdot \text{s}^{-2}$. Since our computations are for two dimensions in space, units for the densities are converted accordingly. The Prandtl numbers are dimensionless, as are some other parameters later in this section.

\[
\begin{align*}
\text{Pr}_A & : 0.713 & \rho_A & : 1.2041 \left( \frac{\text{kg}}{\text{m}^2} \right) \\
\text{Pr}_S & : 7.2 & \rho_S & : 1025 \left( \frac{\text{kg}}{\text{m}^2} \right) \\
c_A & : 1004.9 \left( \frac{\text{J}}{\text{kg} \cdot \text{K}} \right) & \beta_A & : 3.43 \cdot 10^{-3} \left( \frac{1}{\text{K}} \right) \\
c_S & : 3993 \left( \frac{\text{J}}{\text{kg} \cdot \text{K}} \right) & \beta_S & : 2.07 \cdot 10^{-4} \left( \frac{1}{\text{K}} \right)
\end{align*}
\]

Other model parameters are based on the discussion in [14]; the top-of-air model heat flux proportionality constant $C_t$ is based on the longwave radiation parameter $C_{ir}$, but uses the temperature 285K corresponding to the initial conditions of the
test problem. These parameters are given below.

\[ C_{ir} : \theta_{REF}^3 \cdot 4 \cdot 0.985 \cdot \sigma_{St} \left( \frac{W}{m \cdot K} \right) \]
\[ \kappa : 0.5 \]
\[ \sigma_{St} : 5.6704 \cdot 10^{-8} \left( \frac{W}{m \cdot K^4} \right) \text{ (Stefan-Boltzmann)} \]
\[ C_i : (285)^3 \cdot 4 \cdot 0.985 \cdot \sigma_{St} \left( \frac{W}{m \cdot K} \right) \]
\[ \theta_{REF} : 300 \text{ (K)} \]

**APPENDIX B. PROOFS OF THEOREMS**

**B.1. Proof of Theorem 1**

**Proof.** We may drop the subindex \( m \)-notation for the iterations, since we only need to show data for the last iteration explicitly here. Set \( \hat{\theta} = c_{A_\rho A} \hat{\phi}_A^{l,j+1} \) in (40) to get

\[
\frac{c_{A_\rho A}}{2\Delta l_A} \left( 3\phi_A^{l,j+1} - 4\phi_A^{l,j} + \phi_A^{l,j-1}, \phi_A^{l,j+1} \right)_{\Gamma_z} + c_{A_\rho A} \left( E_{A_\rho A}^{l,j+1}, \phi_A^{l,j+1}, \phi_A^{l,j+1} \right)_{\Gamma_x} + \left( F_{A_\rho A}^{l,j+1}, \phi_A^{l,j+1} \right)_{\Gamma_x} + c_{A_\rho A} b_A \left( \phi_A^{l,j+1}, \phi_A^{l,j+1} \right)_{\Gamma_x} = 0.
\]

By Definition (7) the second term vanishes. Application of the polarization identity to the first and last terms yields

\[
\frac{c_{A_\rho A}}{4\Delta l_A} \left[ \left\| \phi_A^{l,j+1} \right\|^2 - \left\| \phi_A^{l,j} \right\|^2 - \left\| \phi_A^{l,j-1} \right\|^2 \right] + c_{A_\rho A} \left( \phi_A^{l,j+1} - 2\phi_A^{l,j} + \phi_A^{l,j-1} \right)_{\Gamma_x} + \left( F_{A_\rho A}^{l,j+1}, \phi_A^{l,j+1} \right)_{\Gamma_x} + c_{A_\rho A} b_A \left( \phi_A^{l,j+1}, \phi_A^{l,j+1} \right)_{\Gamma_x} + c_i \frac{\theta_i^2}{2U} \left\| \phi_A^{l,j+1} \right\|_{\Gamma_r}^2 + \left\| \phi_A^{l,j+1} - \theta_i \right\|_{\Gamma_r}^2 = \frac{c_i \theta_i^2}{2U} \left| \Gamma_r \right|.
\]

Insert \( \left| \Gamma_r \right| = 1 \) and sum over \( j \) for \( j = 0, 1, \ldots, N_A - 1 \),

\[
\frac{c_{A_\rho A}}{4\Delta l_A} \left[ \left\| \phi_A^{l,N_A} \right\|^2 - \left\| \phi_A^{l-1,N_A} \right\|^2 + \left\| \phi_A^{l,N_A-1} \right\|^2 - \left\| \phi_A^{l-1,N_A-1} \right\|^2 \right]
\]

\[
- \left\| \phi_A^{l-1,N_A-1} \right\|^2 + \sum_{j=0}^{N_A-1} \left( \phi_A^{l,j+1} - 2\phi_A^{l,j} + \phi_A^{l,j-1} \right)_{\Gamma_x} + \left( F_{A_\rho A}^{l,j+1}, \phi_A^{l,j+1} \right)_{\Gamma_x} + c_{A_\rho A} \sum_{j=0}^{N_A-1} b_A \left( \phi_A^{l,j+1}, \phi_A^{l,j+1} \right)_{\Gamma_x} + \frac{c_i}{2U} \sum_{j=0}^{N_A-1} \left( \left\| \phi_A^{l,j+1} \right\|_{\Gamma_r}^2 + \left\| \phi_A^{l,j+1} - \theta_i \right\|_{\Gamma_r}^2 \right) = N_A \frac{c_i \theta_i^2}{2U}.
\]
Set \( \hat{\theta} = c_S \rho_S \phi_S^{l,j+1} \) in (43). Perform the analogous steps to find that
\[
\frac{c_S \rho_S}{4 \Delta t_S} \left[ \| \phi_S^{l,N_S} \| - \| \phi_S^{l-1,N_S} \| + \| 2 \phi_S^{l,N_S} - \phi_S^{l-1,N_S} - 2 \| \right] \\
(73) \\
- \left[ \| 2 \phi_S^{l-1,N_S} - \phi_S^{l-1,N_S} \| + \left\| \sum_{j=0}^{N_S-1} \phi_S^{l,j+1} - 2 \phi_S^{l,j} + \phi_S^{l,j-1} \right\| \right] \\
- \sum_{j=0}^{N_S-1} (F_{S(m)}^{l,j+1}, \phi_S^{l,j+1})_{\Gamma_x} + c_S \rho_S \sum_{j=0}^{N_S-1} b_S (\phi_S^{l,j+1}, \phi_S^{l,j+1}) = 0.
\]
Multiply through (72) by \( 4 \Delta t_A \) and (73) by \( 4 \Delta t_S \), then sum the results and drop unnecessary terms to get
\[
\sum_{i \in \{A, S\}} c_i \rho_i \left[ \| \phi_i^{l,N_i} \|^2 - \| \phi_i^{l-1,N_i} \|^2 + \| 2 \phi_i^{l,N_i} - \phi_i^{l-1,N_i} \|^2 - \| 2 \phi_i^{l-1,N_i} - \phi_i^{l-1,N_i-1} \|^2 \right] \\
+ 4 \sum_{i \in \{A, S\}} c_i \rho_i \Delta t_i \sum_{j=0}^{N_i-1} b_i (\phi_i^{l,j+1}, \phi_i^{l,j+1}) + 4 \Delta t_A \sum_{j=0}^{N_A-1} (F_{A(m)}^{l,j+1}, \phi_A^{l,j+1})_{\Gamma_x} \\
- 4 \Delta t_S \sum_{j=0}^{N_S-1} (F_{S(m)}^{l,j+1}, \phi_S^{l,j+1})_{\Gamma_x} \leq \Delta t \frac{2C_i \theta_i^2}{U}.
\]
On both sides of this inequality, add the term \( \Delta t \frac{2C_i \theta_i^2}{U \theta_{\text{REF}}^2 C_{ir}} \). Since (44) is equivalent to
\[
\Delta t \frac{2C_i \theta_i^2}{U \theta_{\text{REF}}^2 C_{ir}} + 4 \Delta t_A \sum_{j=0}^{N_A-1} (F_{A(m)}^{l,j+1}, \phi_A^{l,j+1})_{\Gamma_x} \\
- 4 \Delta t_S \sum_{j=0}^{N_S-1} (F_{S(m)}^{l,j+1}, \phi_S^{l,j+1})_{\Gamma_x} \geq 0,
\]
it follows that
\[
\sum_{i \in \{A, S\}} c_i \rho_i \left[ \| \phi_i^{l,N_i} \|^2 - \| \phi_i^{l-1,N_i} \|^2 + \| 2 \phi_i^{l,N_i} - \phi_i^{l-1,N_i} \|^2 - \| 2 \phi_i^{l-1,N_i} - \phi_i^{l-1,N_i-1} \|^2 \right] \\
+ 4 \sum_{i \in \{A, S\}} c_i \rho_i \Delta t_i \sum_{j=0}^{N_i-1} b_i (\phi_i^{l,j+1}, \phi_i^{l,j+1}) \leq \Delta t \left( \frac{2C_i \theta_i^2}{U \theta_{\text{REF}}^2 C_{ir}} + \frac{2C_i \theta_i^2}{U \theta_{\text{REF}}^2 C_{ir}} \right).
\]
Bound the diffusion terms below using the coercivity constants \( \delta_i \) (Definition 6) and sum over \( l = 0, \ldots, n \) to get (56).

To derive a bound for the velocity, choose \( \tilde{u} = \rho_A w_A^{l,j+1} \) and \( \tilde{q} = q_A^{l,j+1} \). We again drop the \( m \)-notation on the computed states, with the result
\[
\frac{\rho_A}{2 \Delta t_A} \left( 3w_A^{l+1,j} - 4w_A^{l,j} + w_A^{l,j-1}, w_A^{l,j+1} \right) + \rho_A \Lambda_A \left( E_A w_A^{l+1,j}, w_A^{l,j+1}, w_A^{l,j+1} \right) \\
+ \rho_A \gamma_A \| \nabla \cdot w_A^{l+1,j} \| + \rho_A a_A \left( w_A^{l+1,j}, w_A^{l,j+1} \right) - \rho_A \left( q_A^{l+1,j}, \nabla \cdot w_A^{l,j+1} \right)_A \\
+ \rho_A \left( G_A^{l+1,j}, w_A^{l,j+1} \right)_{\Gamma_x} = \rho_A \left( E_A h_A^{l+1,j}, w_A^{l,j+1} \right)_A,
\]
and 
\[(\nabla \cdot \mathbf{w}_A^{l,j+1}, q_A^{l,j+1})_A = 0.\]
As before, the second term vanishes. Also, the pressure term vanishes by the second equation. Applying the polarization identity and Lemma 1 yields
\[
\frac{\rho_A}{4\Delta t_A} \left[ \| \mathbf{w}_A^{l,j+1} \|^2 + 2\| \mathbf{w}_A^{l,j+1} - \mathbf{w}_A^{l,j} \|^2 - \| \mathbf{w}_A^{l,j} \|^2 - 2\| \mathbf{w}_A^{l,j} - \mathbf{w}_A^{l,j-1} \|^2 \right]
\]
\[
+ \| \mathbf{w}_A^{l,j+1} - 2\mathbf{w}_A^{l,j} + \mathbf{w}_A^{l,j-1} \|^2 \right] + \rho_A \gamma_A \| \nabla \cdot \mathbf{w}_A^{l,j+1} \| + \frac{\rho_A^2}{\text{Re}_A} \| (\nabla \mathbf{w}_A^{l,j+1})^T \|^2
\]
\[
+ \frac{\rho_A}{\text{Re}_A} \| (\nabla \mathbf{w}_A^{l,j+1})^T \|^2 + \rho_A \left( G_A^{l,j+1}, \mathbf{w}_A^{l,j+1} \right)_A = \rho_A \left( E_A \mathbf{f}_A^{l,j+1}(m), \mathbf{w}_A^{l,j+1} \right)_A.
\]
Since the data terms on the right side only multiply the vertical component of the velocity, we can bound these terms using Lemma 4 and Young’s inequality:
\[
\left| \rho_A \left( E_A \mathbf{f}_A^{l,j+1}(m), \mathbf{w}_A^{l,j+1} \right)_A \right|
\leq 2\rho_A \left( \mathbf{f}_A^{l,j}(m), \mathbf{w}_A^{l,j+1} \right)_A + \left| \rho_A \left( \mathbf{f}_A^{l,j-1}(m), \mathbf{w}_A^{l,j+1} \right)_A \right|
\leq 2\rho_A \left\| \tilde{g} \left( 1 - \tilde{\beta}_A \left( \phi_A^{l,j} - \tilde{\phi}_A^{l,j} \right) \right) \right\| \left\| (\mathbf{w}_A^{l,j+1})^T \right\|_d
\]
\[
+ \rho_A \left( \tilde{g} \left( 1 - \tilde{\beta}_A \left( \phi_A^{l,j-1} - \tilde{\phi}_A^{l,j-1} \right) \right) \right\| \left\| (\mathbf{w}_A^{l,j+1})^T \right\|_d
\]
\[
\leq 2\mathcal{C}_p \rho_A \left\| \mathbf{w}_A^{l,j+1} \right\| + \mathcal{C}_p \rho_A \| \mathbf{w}_A^{l,j+1} \|
\]
\[
\leq \frac{4(C_p^0 g)^2}{\mathcal{C}_p} \rho_A \left[ \alpha(\phi_A^{l,j}) \right]^2 + \frac{(C_p^0 g)^2}{\mathcal{C}_p} \rho_A \left[ \alpha(\phi_A^{l,j-1}) \right]^2 + \frac{\rho_A^2}{2} \| \nabla \mathbf{w}_A^{l,j+1} \|^2,
\]
where we define \( \alpha(\phi_A^{l,j}) \equiv \left( 1 + \mathcal{C}_p^0 \tilde{\beta}_A \right) \| \nabla \phi_A^{l,j} \| \).

Insert this result into (74) and use \( \varepsilon_A \) to bound the viscosity terms below (Definition 6). Subsume the gradient terms that appear on the right, then multiply through by \( 4\Delta t_A \) and sum over \( j = 0, 1, \ldots, N_A - 1 \). After dropping unnecessary terms, the result is:
\[
\rho_A \left[ \| \mathbf{w}_A^{l,N_A} \|^2 + 2\| \mathbf{w}_A^{l,N_A} - \mathbf{w}_A^{l,N_A-1} \|^2 \right]
\]
\[
+ 2\Delta t_A \rho_A \varepsilon_A \sum_{j=0}^{N_A-1} \| \nabla \mathbf{w}_A^{l,j+1} \|^2 + 4\Delta t_A \rho_A \sum_{j=0}^{N_A-1} \left( G_A^{l,j+1}, \mathbf{w}_A^{l,j+1} \right)_A
\]
\[
\leq \rho_A \left[ \| \mathbf{w}_A^{l-1,N_A} \|^2 + 2\| \mathbf{w}_A^{l-1,N_A} - \mathbf{w}_A^{l-1,N_A-1} \|^2 \right]
\]
\[
+ 16\Delta t_A \sum_{j=0}^{N_A-1} \left( \frac{(C_p^0 g)^2}{\mathcal{C}_p} \rho_A \left[ \alpha(\phi_A^{l,j}) \right]^2 + 4\Delta t_A \sum_{j=0}^{N_A-1} \left( \frac{(C_p^0 g)^2}{\mathcal{C}_p} \rho_A \left[ \alpha(\phi_A^{l,j-1}) \right]^2 \right) \right).
\]
A convenient bound for the last two terms on the right is
\[
16\Delta t_A \sum_{j=0}^{N_A-1} \left( \frac{(C_p^0 g)^2}{\mathcal{C}_p} \rho_A \left[ \alpha(\phi_A^{l,j}) \right]^2 + 4\Delta t_A \sum_{j=0}^{N_A-1} \left( \frac{(C_p^0 g)^2}{\mathcal{C}_p} \rho_A \left[ \alpha(\phi_A^{l,j-1}) \right]^2 \right) \right)
\]
\[
\leq 40\Delta t_A \sum_{j=-1}^{N_A-1} \left( \frac{(C_p^0 g)^2}{\mathcal{C}_p} \rho_A \left( 1 + \left( \frac{\mathcal{C}_p^0 \tilde{\beta}_A}{2} \right)^2 \| \nabla \phi_A^{l,j} \|^2 \right) \right).\]
Next, set \( \tilde{u} = \rho_S \mathbf{w}_S^{l,j+1} \) and \( \tilde{q} = \mathbf{q}_S^{l,j+1} \) in (41). Drop the \( m \)-notation and follow the analysis as shown above to achieve the analogous result

\[
\rho_S \left[ \left\| \mathbf{w}_S^{l,N_S} \right\|^2 + \left\| 2 \mathbf{w}_S^{l,N_S} - \mathbf{w}_S^{l,N_S-1} \right\|^2 \right]
+ 2 \Delta t_S \rho_S \epsilon_S \sum_{j=0}^{N_S-1} \left\| \nabla \mathbf{w}_S^{l,j+1} \right\|^2 - 4 \Delta t_S \rho_A \sum_{j=0}^{N_S-1} \left( \mathbf{g}_S^{l,j+1}, \mathbf{w}_S^{l,j+1} \right)_{\Gamma_x} \\
\leq \rho_S \left\| \mathbf{w}_S^{l-1,N_S} \right\|^2 + \rho_S \left\| 2 \mathbf{w}_S^{l-1,N_S} - \mathbf{w}_S^{l-1,N_S-1} \right\|^2 \\
+ 40 \Delta t_S \sum_{j=1}^{N_S-1} \left( \mathbf{C}_p^0 \mathbf{g} \right)^2 \rho_S \left( 1 + \left( \mathbf{C}_p^0 \mathbf{g} \right)^2 \left\| \nabla \phi_S^{l,j} \right\|^2 \right).
\]

(77)

Insert (76) in (75), then add the resulting inequality to (77). We then bound the flux terms below by zero. Indeed, by (44) it holds that

\[
0 \leq 4 \Delta t_A \rho_A \sum_{j=0}^{N_A-1} \left( \mathbf{g}_A^{l,j+1}, \mathbf{w}_A^{l,j+1} \right)_{\Gamma_x} - 4 \Delta t_S \rho_A \sum_{j=0}^{N_S-1} \left( \mathbf{g}_S^{l,j+1}, \mathbf{w}_S^{l,j+1} \right)_{\Gamma_x}.
\]

Sum the result over \( l = 0, 1, \ldots, n \). This gives

\[
0 \leq \sum_{i \in \{A, S\}} \rho_i \left[ \left\| \mathbf{w}_i^{n,N_i} \right\|^2 + \left\| 2 \mathbf{w}_i^{n,N_i} - \mathbf{w}_i^{n,N_i-1} \right\|^2 \right] + 2 \sum_{i \in \{A, S\}} \Delta t_i \rho_i \epsilon_i \sum_{l=0}^{n} \sum_{j=0}^{N_i-1} \left\| \nabla \mathbf{w}_i^{l,j+1} \right\|^2 \\
\leq \sum_{i \in \{A, S\}} \rho_i \left[ \left\| \mathbf{w}_i^{-1,N_i} \right\|^2 + \left\| 2 \mathbf{w}_i^{-1,N_i} - \mathbf{w}_i^{-1,N_i-1} \right\|^2 \right] \\
+ 40 \sum_{i \in \{A, S\}} \Delta t_i \sum_{l=0}^{n} \sum_{j=0}^{N_i-1} \left( \mathbf{C}_p^0 \mathbf{g} \right)^2 \rho_i \left( 1 + \left( \mathbf{C}_p^0 \mathbf{g} \right)^2 \left\| \nabla \phi_i^{l,j} \right\|^2 \right).
\]

The remaining terms on the right are bounded as follows. We have

\[
40 \sum_{i \in \{A, S\}} \Delta t_i \sum_{l=0}^{n} \sum_{j=0}^{N_i-1} \left( \mathbf{C}_p^0 \mathbf{g} \right)^2 \rho_i \left( 1 + \left( \mathbf{C}_p^0 \mathbf{g} \right)^2 \left\| \nabla \phi_i^{l,j} \right\|^2 \right) \\
\leq 80 t_f (\mathbf{C}_p^0 \mathbf{g})^2 \sum_{i \in \{A, S\}} \frac{\rho_i}{\epsilon_i} + 40 L (\mathbf{C}_p^0 \mathbf{C}_p^0 \mathbf{g})^2 \sum_{i \in \{A, S\}} \Delta t_i \sum_{l=0}^{n} \sum_{j=-1}^{N_i-1} c_i \rho_i \delta_i \left\| \nabla \phi_i^{l,j} \right\|^2 \\
\leq 80 t_f (\mathbf{C}_p^0 \mathbf{g})^2 \sum_{i \in \{A, S\}} \frac{\rho_i}{\epsilon_i} \\
+ 40 L (\mathbf{C}_p^0 \mathbf{C}_p^0 \mathbf{g})^2 \sum_{i \in \{A, S\}} \Delta t_i \sum_{l=0}^{n} \sum_{j=-1}^{N_i-1} c_i \rho_i \delta_i \left\| \nabla \phi_i^{-1,N_i} \right\|^2 + \left\| \nabla \phi_i^{-1,N_i-1} \right\|^2 \\
+ 80 L (\mathbf{C}_p^0 \mathbf{C}_p^0 \mathbf{g})^2 \sum_{i \in \{A, S\}} \Delta t_i \sum_{l=0}^{n} \sum_{j=0}^{N_i-1} c_i \rho_i \delta_i \left\| \nabla \phi_i^{l,j+1} \right\|^2.
\]

where \( L = \max_{i \in \{A, S\}} \left( \frac{\mathbf{g}^2}{\epsilon^2} \right) \). The last term shown on the right of (79) may be bounded using (56), so (57) follows. \( \square \)
B.2. Proof of Lemma 9

Proof. In this proof, the superscript \( n + 1 \) will be left off of all data associated with time \( t^{n+1} \), to make the notation more compact. Set \( \tilde{\mathbf{u}} = \rho_i \mathbf{v}_{i(m)} , \ i \in \{A, S\} \), in the weak momentum equations for the monolithic formulation and sum the equations. Due to the property (66) for the monolithic flux, the result is

\[
\begin{align*}
&\sum_{i \in \{A, S\}} \frac{\rho_i}{4\Delta t} \left( \|\mathbf{v}_{i(m)}\|^2 + 2\|\mathbf{v}_{i(m)} - \mathbf{w}_i^n\|^2 - \|\mathbf{w}_i^n\|^2 - 2\|\mathbf{w}_i^n - \mathbf{w}_i^{n-1}\|^2 \right) \\
&\quad + \sum_{i \in \{A, S\}} \frac{\rho_i}{4\Delta t} \left( \cdot \right) \\
&\quad + \sum_{i \in \{A, S\}} \left( \cdot \right) \\
&\quad + \rho_A \kappa \int_I \left| \mathbf{v}_{A(m-1)} - \mathbf{v}_{S(m-1)} \right| \left| \mathbf{v}_{A(m)} - \mathbf{v}_{S(m)} \right|^2 d\sigma \\
&\quad = \sum_{i \in \{A, S\}} \rho_i \left( 2\mathbf{f}_i^n - \mathbf{f}_i^{n-1}, \mathbf{v}_{i(m)} \right).
\end{align*}
\]

(80)

To bound the right side of the equation, apply Holder, Lemma 4, and Young’s inequality:

\[
\left| \rho_i \left( 2\mathbf{f}_i^n - \mathbf{f}_i^{n-1}, \mathbf{v}_{i(m)} \right) \right| \leq \rho_i \left( 2\mathbf{f}_i^n - \mathbf{f}_i^{n-1} \right) ||\mathbf{v}_{i(m)}||_i \\
\leq 2\Delta t \rho_i \left( 2\mathbf{f}_i^n - \mathbf{f}_i^{n-1} \right) ||\mathbf{v}_{i(m)}||_i + \rho_i \frac{8\Delta t}{8\Delta t} ||\mathbf{v}_{i(m)}||_i^2.
\]

Apply this bound to (80) and subsume the \( \mathbf{v}_{i(m)} \) terms. Multiply through by \( 8\Delta t \). After dropping unneeded terms and moving data terms to the right, the result is:

\[
\begin{align*}
&\sum_{i \in \{A, S\}} \rho_i \left( \|\mathbf{v}_{i(m)}\|^2 + 2\|\mathbf{v}_{i(m)} - \mathbf{w}_i^n\|^2 \right) \\
&\leq 16(\Delta t)^2 \sum_{i \in \{A, S\}} \rho_i \left( \|\mathbf{w}_i^n\|^2 + \|2\mathbf{w}_i^n - \mathbf{w}_i^{n-1}\|^2 \right).
\end{align*}
\]

(81)

By assumption, all of the terms on the right of (81) may be bounded by Theorem 1.

The first bound we apply is:

\[
2 \sum_{i \in \{A, S\}} \rho_i \left( \|\mathbf{w}_i^n\|^2 + \|2\mathbf{w}_i^n - \mathbf{w}_i^{n-1}\|^2 \right) \leq 2Cw.
\]
The bound for the remaining term is:

\[
16 (\Delta t)^2 \sum_{i \in \{A,S\}} \rho_i \|2\phi_i^n - \phi_i^{n-1}\|^2 \\
\leq 16\hat{g}^2 (\Delta t)^2 \sum_{i \in \{A,S\}} \rho_i \left(1 - \hat{\beta}_i (2\phi_i^n - \phi_i^{n-1}) + \hat{\beta}_i (2\phi_i^n - \phi_i^{n-1})\right)^2 \\
\leq 16\hat{g}^2 (\Delta t)^2 \sum_{i \in \{A,S\}} \rho_i \left(\|1\| + \hat{\beta}_i \|2\phi_i^n - \phi_i^{n-1}\| + \hat{\beta}_i \|2\phi_i^n - \phi_i^{n-1}\|\right)^2 \\
\leq 16\hat{g}^2 (\Delta t)^2 \sum_{i \in \{A,S\}} \rho_i \left(\|\Omega_i\| + 2\hat{\beta}_i \|2\phi_i^n - \phi_i^{n-1}\|\right)^2 \\
\leq 32\hat{g}^2 (\Delta t)^2 \sum_{i \in \{A,S\}} \rho_i \left(\|\Omega_i\| + 4\hat{\beta} (\hat{\beta}_i)^2 \|2\phi_i^n - \phi_i^{n-1}\|\right)^2 \\
\leq 32\hat{g}^2 (\Delta t)^2 \sum_{i \in \{A,S\}} \left(\rho_i + \frac{4\hat{\beta}_i}{c_i} C_\phi\right).
\]

This proves (68).

For the temperature, set \(\hat{\vartheta} = c_i \rho_i \xi_i(m), \ i \in \{A,S\}\), in the monolithic formulation and sum to obtain

\[
\sum_{i \in \{A,S\}} \frac{c_i \rho_i}{4\Delta t} \left[\|\xi_i(m)\|^2 + 2\|\xi_i(m) - \phi_i^n\|^2 - \|\phi_i^n\|^2 - \|2\phi_i^n - \phi_i^{n-1}\|^2\right] \\
+ \left[\|\xi_i(m) - 2\phi_i^n + \phi_i^{n-1}\|^2\right] \\
+ \frac{C_{ir}}{U} \left[\|\xi_i(m)\|^2 + \|\xi_i(m) - \theta_i\|^2\right] = \frac{C_i \theta_i^2}{2U}.
\]

It follows from (67) that

\[
\sum_{i \in \{A,S\}} \frac{c_i \rho_i}{4\Delta t} \left[\|\xi_i(m)\|^2 + 2\|\xi_i(m) - \phi_i^n\|^2 - \|\phi_i^n\|^2 - \|2\phi_i^n - \phi_i^{n-1}\|^2\right] \\
+ \left[\|\xi_i(m) - 2\phi_i^n + \phi_i^{n-1}\|^2\right] \\
+ \frac{C_i}{2U} \left[\|\xi_i(m)\|^2 + \|\xi_i(m) - \theta_i\|^2\right] \leq \frac{C_i \theta_i^2}{2U} + \frac{C_{sol}^2}{2U C_{ir}^2}.
\]

Multiply through by \(4\Delta t\), drop unnecessary terms and move the data terms to the right. The result is

\[
\sum_{i \in \{A,S\}} c_i \rho_i \left[\|\xi_i(m)\|^2 + 2\|\xi_i(m) - \phi_i^n\|^2\right] \\
\leq \frac{2\Delta t}{U} \left(C_i \theta_i^2 + \frac{C_{sol}^2}{2 U C_{ir}^2}\right) + \sum_{i \in \{A,S\}} c_i \rho_i \left[\|\phi_i^n\|^2 - \|2\phi_i^n - \phi_i^{n-1}\|^2\right].
\]

Now (69) follows from Theorem 1. \qed
B.3. Proof of Theorem 2

Proof. We simplify our notation by dropping all superscripts \( n + 1 \) for data associated with time \( t^{n+1} \). It is necessary to retain the superscripts for other time levels. Also, unless otherwise stated, we will take \( \| \cdot \|_{L^2} \equiv \| \cdot \|_{L^2(\Omega)} \). The overall method of proof we use is to show that the monolithic iterations and those of Algorithm 1 are converging to a common limit as \( m \to \infty \).

Set \( \Delta w_i(m) \equiv w_i(m) - v_i(m) \) and \( \Delta q_i(m) \equiv q_i(m) - r_i(m) \). Subtract (58) and (60) from the corresponding monolithic formulations to get

\[
\frac{3}{2\Delta t} \left( \Delta w_{A(m)}, \tilde{u} \right)_{A} + A_{A} \left( 2w_{A}^{n} - w_{A}^{n-1}, \Delta w_{A(m)} \right)_{\tilde{u}} + \gamma_{A} \left( \nabla \cdot \Delta w_{A(m)}, \nabla \cdot \tilde{u} \right)_{A} + a_{A} \left( \Delta w_{A(m)}, \tilde{u} \right) - \left( \Delta q_{A(m)}, \nabla \cdot \tilde{u} \right)_{A} \\
+ \kappa \int_{\Gamma_{I}} \left| w_{A(m-s)} - w_{S(m-s-1)} \right| \left( w_{A(m-s)} - w_{S(m-s-1)} \right) \cdot \tilde{u} \, d\sigma \\
- \kappa \int_{\Gamma_{I}} \left| v_{A(m-1)} - v_{S(m-1)} \right| \left( v_{A(m)} - v_{S(m)} \right) \cdot \tilde{u} \, d\sigma = 0, \forall \tilde{u} \in X_{A}^{h}. 
\]

Here \( s = 0 \) for the sequential mode and \( s = 1 \) for the concurrent mode. Without loss of generality, if \( s = 1 \) then we assume \( m \geq 2 \) so that \( m - s - 1 \geq 0 \). Similarly,

\[
\frac{3}{2\Delta t} \left( \Delta w_{S(m)}, \tilde{u} \right)_{S} + A_{S} \left( 2w_{S}^{n} - w_{S}^{n-1}, \Delta w_{S(m)} \right)_{\tilde{u}} + \gamma_{S} \left( \nabla \cdot \Delta w_{S(m)}, \nabla \cdot \tilde{u} \right)_{S} + a_{S} \left( \Delta w_{S(m)}, \tilde{u} \right) - \left( \Delta q_{S(m)}, \nabla \cdot \tilde{u} \right)_{S} \\
- \frac{\rho_{A}}{\rho_{S}} \kappa \int_{\Gamma_{I}} \left| w_{A(m-s)} - w_{S(m-s-1)} \right| \left( w_{A(m-s)} - w_{S(m-s-1)} \right) \cdot \tilde{u} \, d\sigma \\
+ \frac{\rho_{A}}{\rho_{S}} \kappa \int_{\Gamma_{I}} \left| v_{A(m-1)} - v_{S(m-1)} \right| \left( v_{A(m)} - v_{S(m)} \right) \cdot \tilde{u} \, d\sigma = 0, \forall \tilde{u} \in X_{S}^{h}. 
\]

Define new notation

\[
L(m) \equiv \left| w_{A(m-s)} - w_{S(m-s-1)} \right| \left( w_{A(m-s)} - w_{S(m-s-1)} \right) - \left| v_{A(m-1)} - v_{S(m-1)} \right| \left( v_{A(m)} - v_{S(m)} \right). 
\]

Set \( \tilde{u} = \rho_{i} \Delta w_{i(m)}, i \in \{ A, S \} \), respectively. Note that the pressure terms vanish. Apply skew symmetry and sum over \( i \) to get

\[
\sum_{i \in \{ A, S \}} \rho_{i} \left[ \frac{3}{2\Delta t} \left\| \Delta w_{i(m)} \right\|^{2} + \gamma_{i} \left\| \nabla \cdot \Delta w_{i(m)} \right\|^{2} + a_{i} \left( \Delta w_{i(m)}, \Delta w_{i(m)} \right) \right] \\
+ \rho_{A} \kappa \int_{\Gamma_{I}} L(m) \cdot (\Delta w_{A(m)} - \Delta w_{S(m)}) \, d\sigma = 0. 
\]

The left side is bounded below by using the coercivity constants \( \epsilon_{i} \) and dropping the unneeded grad-div terms:

\[
\sum_{i \in \{ A, S \}} \rho_{i} \left[ \frac{3}{2\Delta t} \left\| \Delta w_{i(m)} \right\|^{2} + \epsilon_{i} \left\| \nabla \Delta w_{i(m)} \right\|^{2} \right] \\
+ \rho_{A} \kappa \int_{\Gamma_{I}} L(m) \cdot (\Delta w_{A(m)} - \Delta w_{S(m)}) \, d\sigma \leq 0. 
\]
Move the interface terms to the right side. These are bounded by first adding and subtracting \(|w_A(m-s) - w_S(m-s-1)| (v_A(m) - v_S(m))\) in \(\mathcal{L}(m)\) to get
\[
|\mathcal{L}(m)| \leq |w_A(m-s) - w_S(m-s-1)| \left| (w_A(m-s) - w_S(m-s-1)) - (v_A(m) - v_S(m)) \right|
+ |w_A(m-s) - w_S(m-s-1)| - |v_A(m-1) - v_S(m)| \left| v_A(m) - v_S(m) \right|.
\]
We bound each of these terms using the following inequalities:
\[
|w_A(m-s) - w_S(m-s-1)| \leq |\Delta w_A(m-s)| + |\Delta w_S(m-s-1)| + |v_A(m-s)| + |v_S(m-s-1)|
\]
(84)
\[
|w_A(m-s) - v_A(m)| + |w_S(m-s-1) - v_S(m)| \leq |\Delta w_A(m-s)| + |v_A(m-s) - v_A(m)| + |\Delta w_S(m-s-1)| + |v_S(m-s-1) - v_S(m)|
\]
(85)
\[
|\Delta w_A(m-s)| + |v_A(m-s) - v_A(m)| + |\Delta w_S(m-s-1)| + |v_S(m-s-1) - v_S(m)|
\]
(86)
\[
|v_A(m) - v_S(m)| \leq |v_A(m)| + |v_S(m)|
\]
(87)
Apply the inequalities (84)-(87) along with Hölder's inequality to bound:
\[
\rho_A \kappa \int_{\mathcal{I}} |\mathcal{L}(m)| \left| \Delta w_A(m) - \Delta w_S(m) \right| d\sigma
\]
\[
\leq \rho_A \kappa \left[ \left( \|\Delta w_A(m-s)\|_{L^3} + \|\Delta w_S(m-s-1)\|_{L^3} + \|v_A(m-s)\|_{L^3} \right)
+ \|v_S(m-s-1)\|_{L^3} \right] \times \left( \left( \|\Delta w_A(m-s)\|_{L^3} + \|v_A(m-s) - v_A(m)\|_{L^3} \right)
+ \|\Delta w_S(m-s-1)\|_{L^3} + \|v_S(m-s-1) - v_S(m)\|_{L^3} \right) + \left( \|\Delta w_A(m-s)\|_{L^3} \right)
\times \left( \left( \|v_A(m)\|_{L^3} + \|v_S(m)\|_{L^3} \right) \right) \sum_{i \in \{A,S\}} \|\Delta w_i(m)\|_{L^3}.
\]
In case \(s = 0\), there is a term which is cubic in \(\|\Delta w_A(m)\|_{L^3(\mathcal{I})}\), requiring a special bound. To derive a bound for \(\|\Delta w_A(m)\|_{L^3(\mathcal{I})}\), set \(\bar{u} = \rho_A \Delta w_A(m)\) in (82) to get
\[
\frac{3}{2 \Delta t} \rho_A \left\| \Delta w_A(m) \right\|^2 + \rho_A \gamma_A \left\| \nabla \cdot \Delta w_A(m) \right\|^2 + \rho_A \theta_A \left( \Delta w_A(m), \Delta w_A(m) \right)
\]
\[
+ \kappa \int_{\mathcal{I}} |w_A(m-s) - w_S(m-s-1)| (w_A(m) - w_S(m-1)) \cdot \Delta w_A(m) d\sigma
\]
\[
- \kappa \int_{\mathcal{I}} |v_A(m-s-1) - v_S(m-s-1)| (v_A(m) - v_S(m)) \cdot \Delta w_A(m) d\sigma = 0, \forall \bar{u} \in X_A^h.
\]
Let \( F(x) = |x|x \). Add and subtract \( |v_{A(m)} - w_{S(m-1)}| (v_{A(m)} - w_{S(m-1)}) \) to the interface terms in (89) to get
\[
\rho_{A^k} \int_I [F(w_{A(m)} - w_{S(m-1)}) - F(v_{A(m)} - w_{S(m-1)})] \cdot \Delta w_{A(m)} d\sigma
\]
\[
+ \rho_{A^k} \int_I (v_{A(m)} - w_{S(m-1)}) (v_{A(m)} - w_{S(m-1)}) \cdot \Delta w_{A(m)} d\sigma
\]
\[
- \rho_{A^k} \int_I (v_{A(m-1)} - v_{S(m-1)}) (v_{A(m)} - v_{S(m)}) \cdot \Delta w_{A(m)} d\sigma.
\]

By applying the monotonicity property stated in Lemma 3 and dropping unnecessary terms on the left side, we can then bound (89) below by
\[
\frac{3}{2\Delta t} \rho_A \| \Delta w_{A(m)} \|^2 + \frac{\rho_{A^k}}{4} \| \Delta w_{A(m)} \|_{L^3}^3
\]
\[
+ \rho_{A^k} \left( \| v_{A(m-1)} - v_{S(m-1)} \|_{L^3}^3 + \| v_{A(m)} - v_{S(m-1)} \|_{L^3}^3
\]
\[
+ \| v_{A(m)} - v_{S(m)} \|_{L^3}^3 + \| \Delta w_{S(m-1)} \|_{L^3}^3 \right).
\]

After subsuming the first term on the right, the following bound follows:
(90)
\[
\| \Delta w_{A(m)} \|_{L^3} \leq C \left( \sum_{i \in \{A, S\}} \| v_i(m) \|_{L^3} + \sum_{i \in \{A, S\}} \| v_i(m-1) \|_{L^3} + \| \Delta w_{S(m-1)} \|_{L^3} \right).
\]

Moreover, as an induction hypothesis, assume that there exists a fixed constant \( M \) such that
(91)
\[
\sqrt{\beta_l} \| \Delta w_{i(l)} \| \leq M, \ i \in \{A, S\},
\]
holds for \( m - s - 1 \leq l < m \). Let \( \beta > 0 \) be arbitrary; the value will be specified later. Apply (52), (68) and (91), then use \( \rho_A < \rho_S \) to bound (90) as follows:
\[
\| \Delta w_{A(m)} \|_{L^3(\Gamma^I')} \leq \frac{C}{\sqrt{\beta \Delta t^{1/3}}} \left( \sum_{i \in \{A, S\}} \| v_i(m) \| + \sum_{i \in \{A, S\}} \| v_i(m-1) \| + \| \Delta w_{S(m-1)} \| \right)
\]
\[
\leq \frac{C}{\sqrt{\beta \Delta t^{1/3}}} C_w,
\]
where we define the constant \( C_w \equiv \frac{\beta^{1/3} + M}{\sqrt{\beta S}} \).

We now insert this into (88) and then bound the remaining terms of the form \( \| v_i(m) \|_{L^3(\Gamma^I')} \) as was just shown. After applying (52) again, the result may be
bounded for both $s = 0$ and $s = 1$ as

\begin{equation}
(92) \\
\rho_A \kappa \int_{\Gamma_z} |L(m)| \left| \Delta w_{A(m)} - \Delta w_{S(m)} \right| \, ds \leq \frac{C \rho_A k C_w}{(\beta)^{3/2} \Delta t} \left( \| \Delta w_{A(m-s)} \| + \| \Delta v_{A(m)} \| \\
+ \| \Delta w_{S(m-s-1)} \| + \| \Delta v_{S(m-1)} \| \right) \sum_{i \in \{A,S\}} \| \Delta w_{i(m)} \|
\end{equation}

where we set $\Delta v_{i(m)} \equiv v_{i(m)} - v_{i(m-1)}$.

The interface terms in (83) are bounded now by using (92). Multiply through by $\Delta t$, then use Young’s inequality on the right side and $\rho_A < \rho_S$ to get the bound

\[ \sum_{i \in \{A,S\}} \rho_i \left[ 3 \frac{2}{3} \| | \Delta w_{i(m)} \| + \epsilon_i \Delta t \| \nabla \Delta w_{i(m)} \| \right] \leq \frac{C \rho_A k C_w}{(\beta)^{3/2}} \left( \sum_{i \in \{A,S\}} \rho_i \| \Delta w_{i(m)} \|^2 \right) + \rho_A \| \Delta w_{A(m-s)} \|^2 + \rho_S \| \Delta w_{S(m-s-1)} \|^2 + \rho_A \| \Delta v_{A(m)} \|^2 + \rho_S \| \Delta v_{S(m-1)} \|^2. \]

For all $m$, define

\[ \eta_m \equiv \sum_{i \in \{A,S\}} \rho_i \left[ 3 \frac{2}{3} \| \Delta w_{i(m)} \|^2 + \epsilon_i \Delta t \| \nabla \Delta w_{i(m)} \|^2 \right]. \]

It follows that

\begin{equation}
(93) \\
\eta_m \leq \frac{C \rho_A k C_w}{(\beta)^{3/2}} \left( \frac{2}{3} \eta_m + \frac{2}{3} \eta_{m-s} + \frac{2}{3} \eta_{m-s-1} + \rho_A \| \Delta v_{A(m)} \|^2 + \rho_S \| \Delta v_{S(m-1)} \|^2 \right).
\end{equation}

It is also necessary to show that the $\Delta v_{i(m)}$ -terms are correspondingly small. For this purpose, subtract the monolithic equations at iteration level $m$ and $m-1$ to get

\[ \frac{3}{2 \Delta t} \left( \Delta v_{A(m), \tilde{u}} \right)_A + \Lambda_A \left( 2 w_{A,n} - w_{A:n-1}, \Delta v_{A(m), \tilde{u}} \right) \]

\[ + \gamma_A \left( \nabla \cdot \Delta v_{A(m), \tilde{u}} \right)_A + a_A \left( \Delta v_{A(m), \tilde{u}} \right) - \left( \Delta q_{A(m), \tilde{u}} \right)_A \]

\[ + \kappa \int_{\Gamma_z} \left| v_{A(m-1)} - v_{S(m-1)} \right| \left( v_{A(m)} - v_{S(m)} \right) \cdot \tilde{u} \, d\sigma \]

\[ - \kappa \int_{\Gamma_z} \left| v_{A(m-2)} - v_{S(m-2)} \right| \left( v_{A(m-1)} - v_{S(m-1)} \right) \cdot \tilde{u} \, d\sigma = 0, \forall \tilde{u} \in X_A^h. \]

and

\[ \frac{3}{2 \Delta t} \left( \Delta v_{S(m), \tilde{u}} \right)_S + \Lambda_S \left( 2 w_{S,n} - w_{S:n-1}, \Delta v_{S(m), \tilde{u}} \right) \]

\[ + \gamma_S \left( \nabla \cdot \Delta v_{S(m), \tilde{u}} \right)_S + a_S \left( \Delta v_{S(m), \tilde{u}} \right) - \left( \Delta q_{S(m), \tilde{u}} \right)_S \]

\[ - \frac{\rho_A}{\rho_S} \kappa \int_{\Gamma_z} \left| v_{A(m-1)} - v_{S(m-1)} \right| \left( v_{A(m)} - v_{S(m)} \right) \cdot \tilde{u} \, d\sigma \]

\[ + \frac{\rho_A}{\rho_S} \kappa \int_{\Gamma_z} \left| v_{A(m-2)} - v_{S(m-2)} \right| \left( v_{A(m-1)} - v_{S(m-1)} \right) \cdot \tilde{u} \, d\sigma = 0, \forall \tilde{u} \in X_S^h. \]
Set \( \hat{u} = \rho_i \Delta v_{i(m)} \), \( i \in \{A, S\} \), respectively. Note that the pressure terms vanish. Apply skew symmetry to the convective terms and sum over \( i \) to get

\[
\sum_{i \in \{A, S\}} \rho_i \left[ \frac{3}{2 \Delta t} \| \Delta v_{i(m)} \|_2^2 + \gamma_i \| \nabla \cdot \Delta v_{i(m)} \|_2^2 + a_i (\Delta v_{i(m)}, \Delta v_{i(m)}) \right]
\]

and drop the \( \text{grad-div} \) terms:

\[
+ \rho A \kappa \int_{\Gamma_t} \left| v_{A(m-1)} - v_{S(m-1)} \right| \left| v_{A(m)} - v_{S(m)} \right| \cdot (\Delta v_{A(m)} - \Delta v_{S(m)}) \, d\sigma
\]

\[
- \rho A \kappa \int_{\Gamma_t} \left| v_{A(m-2)} - v_{S(m-2)} \right| \left| v_{A(m-1)} - v_{S(m-1)} \right| \cdot (\Delta v_{A(m)} - \Delta v_{S(m)}) \, d\sigma = 0.
\]

Add and subtract \( (v_{A(m-1)} - v_{S(m-1)})(v_{A(m-1)} - v_{S(m-1)}) \) to the interface terms and multiply through by \( \Delta t \). Bound the gradient terms below using coercivity and drop the \( \text{grad-div} \) terms:

\[
\sum_{i \in \{A, S\}} \rho_i \left[ \frac{3}{2} \| \Delta v_{A(m)} \|_2^2 + \epsilon_i \| \nabla \Delta v_{i(m)} \|_2^2 \right] + \Delta t \rho A \kappa \int_{\Gamma_t} \left| v_{A(m-1)} - v_{S(m-1)} \right| \left| \Delta v_{A(m)} - \Delta v_{S(m)} \right| \, d\sigma
\]

\[
\leq \Delta t \rho A \kappa \int_{\Gamma_t} \left| (v_{A(m-2)} - v_{S(m-2)}) - (v_{A(m-1)} - v_{S(m-1)}) \right| \, d\sigma
\]

\[
\times \left( v_{A(m-1)} - v_{S(m-1)} \right) \cdot (\Delta v_{A(m)} - \Delta v_{S(m)}) \, d\sigma.
\]

To bound the integral on the right, apply the reverse triangle inequality, Hölder’s inequality, and Lemma 8 with a new arbitrary parameter \( \alpha \) in place of \( \beta \):

\[
\Delta t \rho A \kappa \int_{\Gamma_t} \left| (v_{A(m-2)} - v_{S(m-2)}) - (v_{A(m-1)} - v_{S(m-1)}) \right|
\]

\[
\times \left( v_{A(m-1)} - v_{S(m-1)} \right) \cdot (\Delta v_{A(m)} - \Delta v_{S(m)}) \, d\sigma
\]

\[
\leq \rho A \kappa \frac{\alpha^{3/2}}{\sqrt{\rho A}} \left( \| v_{A(m-1)} \| + \| v_{S(m-1)} \| \right) \left( \| \Delta v_{A(m-1)} \| \| \Delta v_{A(m)} \| + \| \Delta v_{A(m-1)} \| \| \Delta v_{S(m)} \| \right)
\]

\[
+ \left( \| \Delta v_{S(m-1)} \| \| \Delta v_{A(m)} \| + \| \Delta v_{S(m-1)} \| \| \Delta v_{S(m)} \| \right).
\]

Use Lemma 8 to bound the \( \| v_{i(m-1)} \| \) terms and then apply Young’s inequality. Insert the result in (94) and drop the gradient and interface terms that still appear on the left side to get

\[
\sum_{i \in \{A, S\}} \frac{3}{2} \rho_i \| \Delta v_{i(m)} \|_2^2 \leq \rho A \kappa \frac{\alpha^{3/2}}{\sqrt{\rho A}} \left( \sum_{i \in \{A, S\}} \| \Delta v_{i(m)} \|_2^2 + \sum_{i \in \{A, S\}} \| \Delta v_{i(m-1)} \|_2^2 \right).
\]

Now choose \( \alpha^{3/2} \) to satisfy

\[
\frac{\kappa}{\alpha^{3/2}} \frac{2 \sqrt{C_v}}{\sqrt{\rho A}} = \frac{1}{2} \Rightarrow \alpha = \frac{(4\kappa)^{2/3} C_v^{1/3}}{\rho A^{1/3}}.
\]
Since $\rho_A < \rho_S$, it follows from (95) that
\[
\sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(m)} \|^2 \leq \frac{1}{2} \left( \sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(m-1)} \|^2 \right) \\ \leq \left( \frac{1}{2} \right)^{m-1} \left( \sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(1)} \|^2 \right).
\]
(97)

Insert this bound in (93):
\[
\eta_m \leq C \kappa C_w \left( \frac{2}{3} \eta_m + \frac{2}{3} \eta_{m-s} + \frac{2}{3} \eta_{m-s-1} + \left( \frac{3}{2^{m-1}} \right) \sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(1)} \|^2 \right).
\]
(98)

Choose $\beta$ to satisfy these conditions, depending on $s$:
\[
(s = 0) \quad \frac{C \kappa C_w 4}{(\beta)^{3/2}} = \frac{1}{3},
\]
(99)
\[
(s = 1) \quad \frac{C \kappa C_w 2}{(\beta)^{3/2}} = \frac{1}{12}.
\]
(100)

In case $s = 0$, it follows from (98) that
\[
\eta_m \leq \frac{1}{3} \eta_m + \frac{1}{6} \eta_{m-1} + \frac{1}{4} \left( \frac{3}{2^{m-1}} \right) \sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(1)} \|^2 \\ \Rightarrow \eta_m \leq \frac{1}{3} \eta_{m-1} + \left( \frac{9}{2^{m+2}} \right) \sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(1)} \|^2.
\]

For $m \geq 1$, this implies the bound
\[
\eta_m \leq \frac{1}{4m} \eta_0 + \left( \sum_{j=0}^{m-1} \frac{1}{2^j} \right) \left( \frac{9}{2^{m+2}} \right) \sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(1)} \|^2 \\ \Rightarrow \eta_m \leq \frac{1}{4m} \eta_0 + \left( \frac{9}{2^{m+1}} \right) \sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(1)} \|^2.
\]
(101)

In case $s = 1$, it follows from (98) and (100) that
\[
\eta_m \leq \frac{1}{12} \eta_m + \frac{1}{12} \eta_{m-1} + \frac{1}{12} \eta_{m-2} + \frac{1}{8} \left( \frac{3}{2^{m-1}} \right) \sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(1)} \|^2 \\ \Rightarrow 11 \eta_m \leq \eta_{m-1} + \eta_{m-2} + \frac{1}{2^m} \left( \frac{9}{2^{m+1}} \right) \sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(1)} \|^2 \\ \Rightarrow 11 \eta_m + 3 \eta_{m-1} \leq 4 \eta_{m-1} + \eta_{m-2} + \frac{1}{2^m} \left( \frac{9}{2^{m+1}} \right) \sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(1)} \|^2 \\ \Rightarrow (4 \eta_m + \eta_{m-1}) \leq \frac{4}{11} (4 \eta_{m-1} + \eta_{m-2}) + \frac{8}{11} \frac{1}{2^{m+1}} \left( \frac{9}{2^{m+1}} \right) \sum_{i \in \{A, S\}} \rho_i \| \Delta \mathbf{v}_{i(1)} \|^2.
\]
from which one can show that for \( m \geq 2 \),

\[(102)\]

\[
4\eta_m + \eta_{m-1} \leq \left( \frac{4}{11} \right)^{m-1} (4\eta_1 + \eta_0) + \left( \frac{1}{2^{m+1}} \right) \left( \frac{1}{2} \sum_{i \in \{A\}} \rho_i \|\Delta v_i(1)\|^2 \right) \
\]

\[
= \left( \frac{4}{11} \right)^{m-1} (4\eta_1 + \eta_0) + \left( \frac{1}{2^{m+1}} \right) \left( 24 \sum_{i \in \{A\}} \rho_i \|\Delta v_i(1)\|^2 \right). \
\]

Set \( m = 1 \) on the right side of \([101]\) and \( m = 2 \) on the right side of \([102]\) to obtain the bounds

\[
\eta_m \leq \frac{1}{4} \eta_0 + \frac{9}{4} \sum_{i \in \{A\}} \rho_i \|\Delta v_i(1)\|^2, \\
4\eta_m + \eta_{m-1} \leq \frac{4}{11} (4\eta_1 + \eta_0) + 3 \sum_{i \in \{A\}} \rho_i \|\Delta v_i(1)\|^2. \
\]

Set \( \frac{3}{2}M^2 \) equal to the right side of the previous inequalities for \( s = 0 \) and \( s = 1 \), respectively. This choice of \( M \) ensures that the induction hypothesis holds, since for both \( s = 0 \) and \( s = 1 \) we have

\[(103)\]

\[
\frac{3}{2} \rho_i \|\Delta w_i(m)\|^2 \leq \eta_m \leq \frac{3}{2} M^2, \
\]

and \([91]\) will be satisfied for \( l = m \). By the definition of \( \eta_m \), the sequences \( w_i(m) \) are Cauchy in \((H^1(\Omega_i))^d\).

For the temperature, set \( \Delta \phi_i(m) \equiv \phi_i(m) - \xi_i(m) \). Subtract \([59]\) and \([61]\) from the corresponding monolith formulas to get

\[
\frac{3}{2\Delta t} \left( \Delta \phi_i(m), \tilde{\theta} \right) + \Lambda A \left( 2w_iA^2 - w_iA^1, \Delta \phi_i, \tilde{\theta} \right) + b_A \left( \Delta \phi_i(m), \tilde{\theta} \right) \\
+ \frac{C_{ir}}{c_A \rho_A U} \int_{\Gamma_x} \left( \phi_iA(m-s) - \phi_iS(m-s-1) \right) \cdot \tilde{\phi} dS - \frac{C_{ir}}{c_A \rho_A U} \int_{\Gamma_x} \left( \xi_iA(m) - \xi_iS(m) \right) \cdot \tilde{\phi} dS \\
+ C_{sen} \int_{\Gamma_x} \left| w_iA(m-s) - w_iS(m-s-1) \right| \left( \phi_iA(m-s) - \phi_iS(m-s-1) \right) \cdot \tilde{\phi} dS \\
- C_{sen} \int_{\Gamma_x} \left| v_iA(m-1) - v_iS(m-1) \right| \left( \xi_iA(m) - \xi_iS(m) \right) \cdot \tilde{\phi} dS \\
+ C_t(U c_A \rho_A)^{-1} \left( \Delta \phi_i(m), \tilde{\theta} \right) \bigg|_{\Gamma_x} = 0, \forall \tilde{\theta} \in \Gamma_A, \\
\]

and

\[
\frac{3}{2\Delta t} \left( \Delta \phi_i(m), \tilde{\theta} \right) + \Lambda S \left( 2w_iS^2 - w_iS^1, \Delta \phi_i, \tilde{\theta} \right) + b_S \left( \Delta \phi_i(m), \tilde{\theta} \right) \\
- \frac{C_{ir}}{c_{S \rho S U}} \int_\Omega \left( \phi_i(m-s) - \phi_iS(m-s-1) \right) \cdot \tilde{\phi} dS + \frac{C_{ir}}{c_{S \rho S U}} \int_\Omega \left( \xi_iA(m) - \xi_iS(m) \right) \cdot \tilde{\phi} dS \\
- C_{sen} \frac{c_A \rho_A}{c_{S \rho S}} \int_\Omega \left[ w_iA(m-s) - w_iS(m-s-1) \right] \left( \phi_iA(m-s) - \phi_iS(m-s-1) \right) \cdot \tilde{\phi} dS \\
+ C_{sen} \frac{c_A \rho_A}{c_{S \rho S}} \int_\Omega \left[ v_iA(m-1) - v_iS(m-1) \right] \left( \xi_iA(m) - \xi_iS(m) \right) \cdot \tilde{\phi} dS = 0, \forall \tilde{\theta} \in \Gamma_S. \\
\]
Define new notation

\[ T(m) \equiv |\mathbf{w}_{A(m-s)} - \mathbf{w}_{S(m-s-1)}| \left( \phi_{A(m-s)} - \phi_{S(m-s-1)} \right) \]
\[ - |\mathbf{v}_{A(m-1)} - \mathbf{v}_{S(m-1)}| \left( \xi_{A(m)} - \xi_{S(m)} \right). \]

Set \( \tilde{\theta} = c_i \rho_i \Delta \phi_i(m) \), apply skew symmetry, and sum over \( i \) to get

\[ \sum_{i \in \{A,S\}} c_i \rho_i \left[ \frac{3}{2\Delta t} \left\| \Delta \phi_i(m) \right\|^2 + b_i \left( \Delta \phi_i(m), \Delta \phi_i(m) \right) \right] \]
\[ + \frac{C_{ir}}{U} \int_I \left\{ \left( \phi_{A(m-s)} - \phi_{S(m-s-1)} \right) - \left( \xi_{A(m)} - \xi_{S(m)} \right) \right\} \cdot \left( \Delta \phi_A(m) - \Delta \phi_S(m) \right) dS \]
\[ + c_A \rho_A C_{sen} \int_I T(m) \cdot \left( \Delta \phi_A(m) - \Delta \phi_S(m) \right) dS + \frac{C_r}{U} \left\| \Delta \phi_A(m) \right\|^2_{\Gamma_x} = 0. \]

Bound the diffusion terms below by using the coercivity constants \( \delta_i \) and drop the unneeded boundary term to get

(104)
\[ \sum_{i \in \{A,S\}} c_i \rho_i \left[ \frac{3}{2\Delta t} \left\| \Delta \phi_i(m) \right\|^2 + \delta_i \left\| \nabla \left( \Delta \phi_i(m) \right) \right\|^2 \right] \]
\[ + \frac{C_{ir}}{U} \int_I \left\{ \left( \phi_{A(m-s)} - \phi_{S(m-s-1)} \right) - \left( \xi_{A(m)} - \xi_{S(m)} \right) \right\} \cdot \left( \Delta \phi_A(m) - \Delta \phi_S(m) \right) dS \]
\[ + c_A \rho_A C_{sen} \int_I T(m) \cdot \left( \Delta \phi_A(m) - \Delta \phi_S(m) \right) dS \leq 0. \]

Move the interface terms to the right side. To bound the first term, add and subtract \( \xi_{A(m-s)} \) and \( \xi_{S(m-s-1)} \) to the integrand to get

(105)
\[ \left| \left( \phi_{A(m-s)} - \phi_{S(m-s-1)} \right) - \left( \xi_{A(m)} - \xi_{S(m)} \right) \right| \left| \Delta \phi_A(m) - \Delta \phi_S(m) \right| \]
\[ \leq \left| \Delta \phi_A(m-s) \right| + \left| \Delta \phi_S(m-s) \right| + \left| \xi_{A(m-s)} - \xi_{A(m)} \right| + \left| \xi_{S(m-s-1)} - \xi_{S(m)} \right| \]
\[ \times \left| \Delta \phi_A(m) - \Delta \phi_S(m) \right| . \]

To bound the second term, add and subtract \( |\mathbf{w}_{A(m-s)} - \mathbf{w}_{S(m-s-1)}| (\xi_{A(m)} - \xi_{S(m)}) \) in \( T(m) \) to get

\[ |T(m)| \leq |\mathbf{w}_{A(m-s)} - \mathbf{w}_{S(m-s-1)}| \left( \phi_{A(m-s)} - \phi_{S(m-s-1)} \right) - \left( \xi_{A(m)} - \xi_{S(m)} \right) \]
\[ + |\mathbf{w}_{A(m-s)} - \mathbf{w}_{S(m-s-1)} - \mathbf{v}_{A(m-1)} + \mathbf{v}_{S(m-1)}| \left| \xi_{A(m)} - \xi_{S(m)} \right|. \]

Each of these terms are bounded using analogous arguments to those shown in (91)-(97):
\[ \frac{C_{ir}}{U_T} \int \left( \phi_A(m-s) - \phi_S(m-s-1) \right) - \left( \xi_A(m) - \xi_S(m) \right) \left| \Delta \phi_A(m) - \Delta \phi_S(m) \right| dS \\
+ c_{A \rho_A} C_{sen} \int \left| T(m) \right| \left| \Delta \phi_A(m) - \Delta \phi_S(m) \right| dS \leq \frac{C_{ir}}{U_T} \left( \left\| \phi_A(m-s) \right\|_{L^3} + \left\| \phi_S(m-s-1) \right\|_{L^3} \right) + \left\| \phi_A(m-s) - \phi_S(m-s-1) \right\|_{L^3} + \left\| \Delta \phi_A(m-s) - \Delta \phi_S(m-s-1) \right\|_{L^3} \\
+ \left\| \phi_A(m-s) - \phi_S(m-s-1) \right\|_{L^3} \sum_{i \in \{A,S\}} \left\| \Delta \phi_i(m) \right\|_{L^3} + c_{A \rho_A} C_{sen} \left( \left\| \Delta \phi_A(m-s) \right\|_{L^3} \\
+ \left\| \Delta \phi_S(m-s-1) \right\|_{L^3} \right) \times \left( \left\| \Delta \phi_A(m-s) \right\|_{L^3} + \left\| \Delta \phi_S(m-s-1) \right\|_{L^3} \\
+ \left\| \phi_A(m-s) - \phi_S(m-s-1) \right\|_{L^3} \right) \times \left( \left\| \Delta \phi_A(m-s) \right\|_{L^3} + \left\| \Delta \phi_S(m-s-1) \right\|_{L^3} \right) \sum_{i \in \{A,S\}} \left\| \Delta \phi_i(m) \right\|_{L^3}. \]

Let \( \beta > 0 \) be arbitrary. Apply (52) and bound terms of the form \( \|v_i(m)\| \) and \( \|\xi_i(m)\| \) using (68) and (69). For terms of the form \( \|\Delta v_i(m)\| \), apply (91), which was already shown to hold for all \( m \). Set \( \Delta \xi_i(m) \equiv \xi_i(m) - \xi_i(m-1) \). The result may be bounded for both \( s = 0 \) and \( s = 1 \) as

\[ \frac{C_{ir}}{U_T} \int \left( \phi_A(m-s) - \phi_S(m-s-1) \right) - \left( \xi_A(m) - \xi_S(m) \right) \left| \Delta \phi_A(m) - \Delta \phi_S(m) \right| dS \\
+ c_{A \rho_A} C_{sen} \int \left| T(m) \right| \left| \Delta \phi_A(m) - \Delta \phi_S(m) \right| dS \leq \frac{C_{ir \rho_A} C_{sen} \sqrt{\phi}}{\beta \Delta t} \left( \left\| \Delta \phi_A(m-s) \right\| + \left\| \Delta \phi_S(m-s-1) \right\| \right) + \left\| \Delta \xi_A(m) \right\| + \left\| \Delta \xi_S(m-1) \right\| \\
+ \left\| \Delta v_A(m-s) \right\| + \left\| \Delta v_S(m-s-1) \right\| \sum_{i \in \{A,S\}} \left\| \Delta \phi_i(m) \right\|, \]

where we define the constant \( C_{\phi} \equiv \frac{(M+\sqrt{\phi_A})\sqrt{\phi_A} + \sqrt{\phi_S}}{\sqrt{\phi_A \rho_A} \sqrt{\phi_S} + c_{ir \rho_A} C_{sen}} \).

The interface terms in (104) are bounded now by using (110). Multiply through by \( \Delta t \), then use Young's inequality on the right side as well as \( \rho_A < \rho_S \) and
\( c_A < c_S \) to get the bound

\[
\sum_{i \in \{A,S\}} c_i \rho_i \left[ \frac{3}{2} \left\| \Delta \phi_{i(m)} \right\|^2 + \delta_i \Delta t \left\| \nabla \left( \Delta \phi_{i(m)} \right) \right\|^2 \right] \\
\leq \frac{CC_{sen} C_\phi}{(\beta)^{3/2}} \left( \sum_{i \in \{A,S\}} c_i \rho_i \left\| \Delta \phi_{i(m)} \right\|^2 + c_A \rho_A \left\| \Delta \phi_{A(m-s)} \right\|^2 + c_S \rho_S \left\| \Delta \phi_{S(m-s-1)} \right\|^2 \right. \\
+ c_A \rho_A \left\| \Delta \xi_{A(m)} \right\|^2 + c_S \rho_S \left\| \Delta \xi_{S(m-1)} \right\|^2 + c_A \left( \rho_A \left\| \Delta v_{A(m)} \right\|^2 + \rho_S \left\| \Delta v_{S(m-1)} \right\|^2 \right) \\
+ c_A \left( \rho_A \left\| \Delta w_{A(m-s)} \right\|^2 + \rho_S \left\| \Delta w_{S(m-s-1)} \right\|^2 \right) \bigg).
\]

Note that \( c_A < c_S \) has only been applied to some terms on the right, which helps for some technical bounds later. For all \( m \), define

\[
\lambda_m = \sum_{i \in \{A,S\}} c_i \rho_i \left[ \frac{3}{2} \left\| \Delta \phi_{i(m)} \right\|^2 + \delta_i \Delta t \left\| \nabla \left( \Delta \phi_{i(m)} \right) \right\|^2 \right].
\]

It follows that

\[
\lambda_m \leq \frac{CC_{sen} C_\phi}{(\beta)^{3/2}} \left( \frac{2}{3} \lambda_m + \frac{2}{3} \lambda_{m-s} + \frac{2}{3} \lambda_{m-s-1} \right) \\
+ c_A \rho_A \left\| \Delta \xi_{A(m)} \right\|^2 + c_S \rho_S \left\| \Delta \xi_{S(m-1)} \right\|^2 + c_A \left( \rho_A \left\| \Delta v_{A(m)} \right\|^2 + \rho_S \left\| \Delta v_{S(m-1)} \right\|^2 \right) \\
+ c_A \left( \rho_A \left\| \Delta w_{A(m-s)} \right\|^2 + \rho_S \left\| \Delta w_{S(m-s-1)} \right\|^2 \right) \right). \tag{111}
\]

It is also necessary to show that the \( \Delta \xi_{i(m)} \)-terms are correspondingly small. To this end, subtract the monolithic equations at iteration level \( m \) and \( m-1 \) to get

\[
\frac{3}{2 \Delta t} \left( \Delta \xi_{A(m),\theta} \right)_A + A \left( 2w_{A} - w_{A}^{-1}, \Delta \xi_{A,\theta} \right) + b_A \left( \Delta \xi_{A(m),\theta} \right) \\
+ \frac{C_{ir}}{c_A \rho_A U} \int_I \left( \Delta \xi_{A(m)} - \Delta \xi_{S(m)} \right) \cdot \hat{\theta} dS + C_t(U c_A \rho_A)^{-1} \left( \Delta \xi_{A(m),\theta} \right) \Gamma_{\theta} \\
+ C_{sen} \int_I \left( v_{A(m-1)} - v_{S(m-1)} \right) \left( \xi_{A(m)} - \xi_{S(m)} \right) \cdot \hat{\theta} dS \\
- C_{sen} \int_I \left( v_{A(m-2)} - v_{S(m-2)} \right) \left( \xi_{A(m-1)} - \xi_{S(m-1)} \right) \cdot \hat{\theta} dS = 0, \forall \hat{\theta} \in \Theta_A^h,
\]

and

\[
\frac{3}{2 \Delta t} \left( \Delta \xi_{S(m),\theta} \right)_S + A \left( 2w_{S} - w_{S}^{-1}, \Delta \xi_{S,\theta} \right) + b_S \left( \Delta \xi_{S(m),\theta} \right) \\
- \frac{C_{ir}}{c_S \rho_S U} \int_I \left( \Delta \xi_{A(m)} - \Delta \xi_{S(m)} \right) \cdot \hat{\theta} dS \\
- C_{sen} \frac{c_A \rho_A}{c_S \rho_S} \int_I \left( v_{A(m-1)} - v_{S(m-1)} \right) \left( \xi_{A(m)} - \xi_{S(m)} \right) \cdot \hat{\theta} dS \\
+ C_{sen} \frac{c_A \rho_A}{c_S \rho_S} \int_I \left( v_{A(m-2)} - v_{S(m-2)} \right) \left( \xi_{A(m-1)} - \xi_{S(m-1)} \right) \cdot \hat{\theta} dS = 0, \forall \hat{\theta} \in \Theta_S^h.
\]
Set $\tilde{\theta} = c_i \rho_i \Delta \xi_i(m)$, $i \in \{A, S\}$, respectively. Apply skew symmetry and sum over $i$:

$$\frac{3}{2 \Delta t} \left( \sum_{i \in \{A,S\}} c_i \rho_i \left\| \Delta \xi_i(m) \right\|^2 \right) + \frac{C_U}{U} \int_I \left| \Delta \xi_A(m) - \Delta \xi_S(m) \right|^2 dS$$

$$+ c_A \rho_A C_{sen} \int_I \left| \mathbf{v}_A(m-1) - \mathbf{v}_S(m-1) \right| \left( \xi_A(m) - \xi_S(m) \right) \cdot \left( \Delta \xi_A(m) - \Delta \xi_S(m) \right) dS$$

$$- c_A \rho_A C_{sen} \int_I \left| \mathbf{v}_A(m-2) - \mathbf{v}_S(m-2) \right| \left( \xi_A(m-1) - \xi_S(m-1) \right) \cdot \left( \Delta \xi_A(m) - \Delta \xi_S(m) \right) dS$$

$$+ \sum_{i \in \{A,S\}} c_i \rho_i b_i \left( \Delta \xi_i(m), \Delta \xi_i(m) \right) + \frac{C_U}{U} \left\| \Delta \xi_A(m) \right\|^2 \Gamma_i = 0.$$ 

Bound the diffusion terms below by $\delta_i$ and add and subtract $\left| \mathbf{v}_A(m-1) - \mathbf{v}_S(m-1) \right| \left( \xi_A(m-1) - \xi_S(m-1) \right)$:

$$\frac{3}{2 \Delta t} \left( \sum_{i \in \{A,S\}} c_i \rho_i \left\| \Delta \xi_i(m) \right\|^2 + \delta_i \Delta t \left\| \nabla (\Delta \xi_i(m)) \right\|^2 \right) + \frac{C_U}{U} \left\| \Delta \xi_A(m) \right\|^2 \Gamma_i$$

$$+ c_A \rho_A C_{sen} \int_I \left| \mathbf{v}_A(m-1) - \mathbf{v}_S(m-1) \right| \left| \Delta \xi_A(m) - \Delta \xi_S(m) \right|^2 dS$$

$$= c_A \rho_A C_{sen} \int_I \left( \left| \mathbf{v}_A(m-2) - \mathbf{v}_S(m-2) \right| - \left| \mathbf{v}_A(m-1) - \mathbf{v}_S(m-1) \right| \right)$$

$$\times \left( \xi_A(m-1) - \xi_S(m-1) \right) \cdot \left( \Delta \xi_A(m) - \Delta \xi_S(m) \right) dS.$$ 

Bound the right side analogously to the proof of showing $(\mathbf{v}_{i(m)})_m$ was Cauchy and drop the unnecessary terms that still appear on the left side to obtain

$$\frac{3}{2} \left( \sum_{i \in \{A,S\}} c_i \rho_i \left\| \Delta \xi_i(m) \right\|^2 \right) \leq \frac{c_A \rho_A C_{sen}}{\alpha^{3/2}} \sqrt{\frac{C_\xi}{C_A \rho_A}} \left( \sum_{i \in \{A,S\}} \left\| \Delta \xi_i(m) \right\|^2 + \sum_{i \in \{A,S\}} \left\| \Delta \mathbf{v}_{i(m-1)} \right\|^2 \right).$$

Now choose $\alpha^{3/2}$ to satisfy

$$\frac{C_{sen}}{\alpha^{3/2}} \sqrt{\frac{C_\xi}{C_A \rho_A}} = \frac{1}{2} \Rightarrow \alpha = \left( \frac{4C_{sen}}{C_\xi} \right)^{2/3} \frac{1}{(C_A \rho_A)^{1/3}}.$$ 

Since $c_A \rho_A < c_S \rho_S$, it follows from (112) that

$$\frac{3}{2} \sum_{i \in \{A,S\}} c_i \rho_i \left\| \Delta \xi_i(m) \right\|^2 \leq \frac{1}{2} \sum_{i \in \{A,S\}} c_i \rho_i \left\| \Delta \xi_i(m) \right\|^2 + \frac{1}{2} \sum_{i \in \{A,S\}} c_i \rho_i \left\| \Delta \mathbf{v}_{i(m-1)} \right\|^2.$$
Subsume the first term on the right side and apply the previously derived bound (97) to obtain
\[
\sum_{i \in \{A, S\}} c_i \rho_i \|\Delta \xi_i(m)\|^2 \leq \frac{1}{2} \left( \sum_{i \in \{A, S\}} c_i \rho_i \|\Delta v_i(1)\|^2 \right)
\]
\[
\leq \left( \frac{1}{2} \right)^{m-1} \left( \sum_{i \in \{A, S\}} c_i \rho_i \|\Delta v_i(1)\|^2 \right).
\]
Insert this bound in (111). Also, bound the \(\|\Delta v_i(m)\|^2\) terms by again applying the bound (97). For the remaining \(\|\Delta w_i(m)\|^2\) terms, insert (101) for \(s = 0\) and (102) for \(s = 1\). We then obtain for \(s = 0\)
\[
\lambda_m \leq \frac{CC_{\text{sen}} C_\phi}{(\beta)^{3/2}} \left( \frac{4}{3} \lambda_m + \frac{2}{3} \lambda_{m-1} + \left( \frac{3}{2m-1} \right) \left( \sum_{i \in \{A, S\}} c_i \rho_i \|\Delta v_i(1)\|^2 \right) \right)
\]
\[
+ \frac{1}{4m} \frac{10c_A}{3} \eta_0 + \frac{15c_A}{2m} \sum_{i \in \{A, S\}} \rho_i \|\Delta v_i(1)\|^2 \right) ,
\]
which holds for \(m \geq 1\), and for \(s = 1\) with \(m \geq 2\)
\[
\lambda_m \leq \frac{CC_{\text{sen}} C_\phi}{(\beta)^{3/2}} \left( \frac{2}{3} \lambda_m + \frac{2}{3} \lambda_{m-1} + \frac{2}{3} \lambda_{m-2} + \left( \frac{3}{2m-1} \right) \left( \sum_{i \in \{A, S\}} c_i \rho_i \|\Delta v_i(1)\|^2 \right) \right)
\]
\[
+ \frac{4c_A}{3} \left( \frac{4}{11} \right)^{m-2} (4\eta_1 + \eta_0) + \frac{19c_A}{2m-1} \sum_{i \in \{A, S\}} \rho_i \|\Delta v_i(1)\|^2 \right) \right).
\]
Choose \(\beta\) to satisfy these conditions, depending on \(s\):
\[
\begin{align*}
(s = 0) & \quad \frac{CC_{\text{sen}} C_\phi}{(\beta)^{3/2}} \frac{4}{3} = \frac{1}{3}, \\
(s = 1) & \quad \frac{CC_{\text{sen}} C_\phi}{(\beta)^{3/2}} \frac{2}{3} = \frac{1}{12}.
\end{align*}
\]
By using analogous arguments to those shown for the velocity, it follows from the definition of \(\lambda_m\) that the sequences \(\phi_i(m)\) are Cauchy in \(H^1(\Omega_i)^d\).

We have proved that the monolithic velocity and temperature states must converge as \(m \to \infty\), and to the same limits as the corresponding states computed from Algorithm [1]. Since the trace operator is continuous as a map from \(H^1(\Omega_i)\) into \(L^p(\partial \Omega_i)\) for the necessary range \(1 \leq p \leq 3\) (and for both dimensions \(d = 2, 3\)), the stability condition (44) must hold for some finite value \(m = m_0\) because of the positivity assumptions (70)-(71). □