A dynamic-kinematic 3D model for density-driven ocean circulation flows: Construction, global well-posedness and dynamics

Ori Saporta-Katz
Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 7610001, Israel

Edriss S. Titi
Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 7610001, Israel
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB30WA, UK and
Department of Mathematics, Texas A& M University, College Station, TX 77843, USA

Hezi Gildor
Institute of Earth Sciences, Hebrew University, Jerusalem 9190401, Israel

Vered Rom-Kedar
Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 7610001, Israel,
The Estrin Family Chair of Computer Science and Applied Mathematics

Differential buoyancy surface sources in the ocean may induce a density-driven flow that joins faster flow components to create a multi-scale, 3D flow. Potential temperature and salinity are active tracers that determine the ocean's potential density: their distribution strongly affects the density-driven component while the overall flow affects their distribution. We present a robust framework that allows one to study the effects of a general 3D flow on a density-driven velocity component, by constructing a modular observation-based 3D model of intermediate complexity. The model contains an incompressible velocity that couples two advection-diffusion equations, for temperature and salinity. Instead of solving the Navier-Stokes equations for the velocity, we consider a flow composed of several temporally separated, spatially predetermined modes. One of these modes models the density-driven flow: its spatial form describes the density-driven flow structure and its strength is determined dynamically by average density differences. The other modes are completely predetermined, consisting of any incompressible, possibly unsteady, 3D flow, e.g. as determined by kinematic models, observations, or simulations. The model is a non-linear, weakly coupled system of two non-local PDEs. We prove its well-posedness in the sense of Hadamard, and obtain rigorous bounds regarding analytical solutions. The model's relevance to oceanic systems is demonstrated by tuning the model to mimic the North Atlantic ocean's dynamics. In one limit the model recovers a simplified oceanic box model and in another limit a kinematic model of oceanic chaotic advection, suggesting it can be utilized to study spatially dependent feedback processes in the ocean.

I. INTRODUCTION

In various oceanic systems, significant horizontal density differences are induced by differential surface buoyancy sources, resulting from spatially varying heating, precipitation, evaporation, river run-off etc. These drive a sedate flow that leads eventually to a sinking of heavier water under lighter water. If the differential sources are transient, and no other forces are at play, the system will equilibrate through stratification, mixing and diffusion. However, if the differential forcing continues, a stable circulation may ensue; this phenomenon is commonly termed horizontal convection [29]. One well-known geophysical example is the Atlantic Meridional Overturning Circulation (AMOC), an important component of Earth's climate system [7, 31], induced in part by differential heating and freshwater sources between the equator and high latitudes [17].

In idealized horizontal convection studies, a rectangular basin forced only by an anti-symmetric differential surface heating exhibits for a range of Rayleigh numbers a robust circular flow structure that is independent of the forcing details. It includes a thin boundary layer close to the surface with a horizontal flow that contains a mixed layer, a narrow eddying plume of downwelling at the boundary of negative buoyancy forcing, and a broad compensating flow throughout the interior of the basin [29, 48]. While the intensity of the flow depends on the details of the problem, this robust flow structure appears when the aspect ratio is altered [47], a small rotation rate is added [5], salinity effects are considered [42], and the surface is also mechanically forced [52]. This flow structure is also a good schematic description of the AMOC [7], despite the complex geometry and various additional forcing features of the actual oceanic system, e.g. winds, tides, exchange flow with marginal seas, and open boundaries.

Density-driven flows have been studied qualitatively by 2D box models as presented in the seminal work Stommel [50], and extended in, e.g., Ashkenazy & Tziperman [4], Calmanti *et al.* [9], Cessi [11], Gildor & Tziperman [22], Griffies & Tziperman [25], Huang *et al.* [28], Pasquero & Tziperman [44], Tziperman *et al.* [56]. In these box models, an ocean basin is divided into instantaneously mixed boxes, and the transport between them scales like their average density difference, as determined by temperature and salinity. This simple dependence is derived from a rough finite-difference approximation to the zonally averaged Navier-Stokes (NS) equations [22, 56]. Temperature and salinity have an opposing effect on density, resulting in a bistable flow regime in most box models. This qualitative theoretical prediction has also been found as a robust feature of complex General Circulation Models (GCMs) [45], and has incited works regarding the stability of the AMOC (Weijer *et al.* [58] and references therein). An equivalent bistability of overturning in horizontal convection has been observed in laboratory experiments of temperature-induced horizontal convection with saltwater perturbations [42]. Using boundary-layer scaling, the same dependence between overturning intensity and a horizontal density difference as used in box models is derived in Mullarney *et al.* [41] for horizontal convection; furthermore, observational evidence for this scaling law for the AMOC appears in Butler *et al.* [8], Sijp *et al.* [49].

In reality, a geophysical flow contains additional fast-varying large scale velocity modes such as wind- and tidally-driven flows. Even if their underlying large-scale structure is laminar, their variability may induce chaotic advection, i.e. chaotic Lagrangian trajectories of passive scalars in a (generally non-turbulent) kinematic flow [2, 3]. A canonical example for employing kinematic modelling in the study of transport and mixing in the ocean are studies of a double-gyre surface flow such as occurs in the North Atlantic. In these studies, a laminar flow that models the oceanic double-gyre flow with a strong seasonal variability, leads to chaotic advection in the ocean [1, 32, 59–61]. Such simplified models are employed to provide a qualitative understanding of physical processes associated with fluid mixing and transport. However, aimed at studying passive scalars, they do not include any feedback mechanism.

Here we present, simulate and analyze a novel dynamic-kinematic model that serves as a missing link in modelling 3D density-driven flows, set between idealized ocean models such as horizontal convection and 2D box models, and fully coupled GCMs. The model is built with the goal of isolating and analyzing the interplay between a general 3D velocity mode and a density-driven velocity component, by allowing one to study the effect of the former on the strength, stability and variability of the latter. The construction starts with the coupled advection-diffusion equations for the temperature and the salinity in a closed 3D oceanic basin. The coupling velocity is a sum of two velocity modes: a general 3D flow that is completely predetermined, e.g. by kinematic models, previously deduced solutions to the NS equations, or observations; and a density-driven flow, with a predetermined form and a dynamically determined intensity that scales like the density difference between two regions of the basin. This dynamic-kinematic framework can easily be modified to a large variety of settings: different domains, spherical geometry, various types of boundary conditions, or more complicated scaling laws. We demonstrate the model's relevance to actual oceanic flows by tuning the model to the current-day North Atlantic ocean flow, and recover a realistic AMOC strength. We show, both rigorously and numerically, that our model recovers in a certain limit of the parameters the simplified 2X2 box model for the North Atlantic flow as appears in Huang *et al.* [28], Tziperman *et al.* [56]. Furthermore, we show how to reproduce different types of previously studied box models; thus, the model is the natural extension of 2D box models into a 3D setting that takes into account innerbox density variations. This model may help in understanding the effect of salinity and temperature fluctuations in the North Atlantic on the strength, variability and stability of the AMOC, by setting a convenient stage for a numerical study of the asymptotic behavior of the system.

The model is a non-local, non-linear coupled system of PDEs, for which basic properties such as the very existence of solutions are not obvious [33]. A desired property of a PDE model is global well-posedness in the sense of Hadamard, defined as the existence and uniqueness of solutions to the system for all times, along with a smooth dependence on the problem's data

[26]. This property has been shown to be satisfied in some oceanic and atmospheric models: see, e.g., the survey papers Li & Titi [35], Temam & Ziane [54] and references therein. Since explicit solutions to nonlinear evolution equations are generally inaccessible, and these models are probed by approximated numerical solutions, such a proof strengthens the physical viability of employed models. We address this issue rigorously, proving the global well-posedness of our model in the sense of Hadamard: if the boundary conditions satisfy a compatibility condition, we prove that the corresponding initial-boundary-value problem has a unique strong solution for all times, along with a smooth dependence of solutions on the problem data. The proof reveals several bounds on averages of solutions; furthermore, we calculate a bound for the Rayleigh number of the problem below which the system has a globally stable unique steady state solution, and above which the system may exhibit nontrivial long-time dynamics. For example, in our simulations, when using autonomous forcing, we observe bi-stability when the Rayleigh number is large enough; see section IV. Moreover, regardless of the size of the Rayleigh number, one can show that the infinite-dimensional dynamical system generated by this model is dissipative and possesses a finite-dimensional non-empty global attractor, a subject that we postpone to future study.

The structure of the work is as follows. In section II we present the model formulation. We describe the relevance of its numerical solutions to the North Atlantic ocean given an appropriate tuning of the parameters in section III. In section IV we show that our model is a natural generalization of the box model scheme, and in particular we illustrate that it is mathematically equivalent to the popular 2×2 box model for certain values of the parameters. Rigorous mathematical analysis of our model - a proof of the system's well-posedness, in the sense of Hadamard, as well as some useful bounds - is detailed in section V. We discuss the results and outline future directions in section VI.

II. MODEL CONSTRUCTION

The evolution of the temperature field $T(\boldsymbol{r},t)$ and of the salinity field $S(\boldsymbol{r},t)$ is modelled by the coupled advection-diffusion equations

$$\partial_t T(\boldsymbol{r},t) + \boldsymbol{u}(\boldsymbol{r},t;\rho(T,S)) \cdot \boldsymbol{\nabla} T(\boldsymbol{r},t) = \boldsymbol{\nabla} \cdot (\boldsymbol{\kappa} \boldsymbol{\nabla} T(\boldsymbol{r},t)) + f_T(\boldsymbol{r},t), \tag{1}$$

$$\partial_t S(\boldsymbol{r},t) + \boldsymbol{u}(\boldsymbol{r},t;\rho(T,S)) \cdot \boldsymbol{\nabla} S(\boldsymbol{r},t) = \boldsymbol{\nabla} \cdot (\boldsymbol{\kappa} \boldsymbol{\nabla} S(\boldsymbol{r},t)) + f_S(\boldsymbol{r},t), \tag{2}$$

inside a closed rectangular basin $\boldsymbol{r} \in \Omega = L_x \times L_y \times L_z \subset \mathbb{R}^3$. We define the aspect ratio $A = L_z/L_x$, and note that in general $L_x \sim L_y$. For large scale ocean phenomena, $A \sim 10^{-3}$, similar to a sheet of paper. This emphasizes how close to 2D are large-scale oceanic flows. Nonetheless, understanding the ocean requires a 3D approach; for our purposes, in density-driven flows, taking into account the vertical direction is imperative.

An equation of state relates temperature and salinity to the potential density, $\rho = \rho(T,S)$. Here we use the linearized equation of state,

$$\rho(T,S) = \rho_0(1 - \alpha T + \beta S), \tag{3}$$

where α is the thermal expansion coefficient and β is the haline contraction coefficient. We note that we define T and S as anomalies: $T = 0, S = 0$ are defined as the temperature and salinity for which $\rho = \rho_0$. For ambient ocean conditions the average density is of the order $\rho_0 \sim 1026 kg/m^3$, corresponding to a temperature of $15°C$ and a salinity of $35psu$. The respective coefficients are of the order $\alpha \sim 10^{-4} K^{-1}$ and $\beta \sim 10^{-4} psu^{-1}$.

The system of PDEs, equations (1) and (2), is coupled only via the advecting velocity field $\boldsymbol{u}(\boldsymbol{r},t;\rho(T,S))$. This coupling renders the tracers active: their distribution influences the flow field, which in turn changes the tracers' distributions. In the full set of equations describing oceanic flow, equations (1) and (2) are coupled to the incompressible NS through an equation of state that relates density to temperature, salinity, and pressure. Our approach employs a significant simplification: Assume we know the spatial form of a certain large-scale, basin-wide solution $\boldsymbol{u}$, e.g. based on observations of the actual velocity field in the ocean or by numerical simulations. This $\boldsymbol{u}$ is the advective field that appears in the advection-diffusion equations. Given such a solution, we assume that it can be written as a sum of several modes, and that only some of its modes' amplitudes are determined dynamically by the density distribution. The phenomenological motivation for this approach is that the velocity field in the ocean is built from several components; some are mainly driven by external sources, for example wind stress and tides, and some are dominantly driven by internal stresses deriving from density inhomogeneities, as discussed above. While the NS are nonlinear, one hopes that such a division can be justified by separation of temporal or spatial scales.

Our model example will be a general oceanic velocity field that can be decomposed into a sum of two components: an externally-forced velocity mode that does not depend on density, and an internally-forced velocity mode that does. As described in the introduction, in a density-driven circulation flow it is plausible to assume that density inhomogeneities affect the strength, but not the form, of the internally forced velocity field components. Furthermore, we assume that the internal velocity's strength depends on spatial averages of the density, denoted in general $\langle \rho \rangle$, and not on pointwise density values. Thus, the internally-forced velocity field has the following structure: $\boldsymbol{U}_I(\boldsymbol{r},t;\langle\rho\rangle) = \sum_j a_I^j(\langle\rho\rangle)\boldsymbol{u}_I^j(\boldsymbol{r},t)$, where I signifies internal. Taking a first-order

approximation we consider one such internally forced mode, so that the velocity field has the form

$$\boldsymbol{u}(\boldsymbol{r},t;\langle\rho\rangle) = a_I(\langle\rho\rangle)\boldsymbol{u}_I(\boldsymbol{r},t) + a_E\boldsymbol{u}_E(\boldsymbol{r},t), \tag{4}$$

where E signifies external. We set up the problem such that a_I and a_E have units of velocity, while $\boldsymbol{u}_I$ and $\boldsymbol{u}_E$ are dimensionless. Each of the velocity modes is incompressible by construction:

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}_I = \boldsymbol{\nabla}\cdot\boldsymbol{u}_E = 0. \tag{5}$$

From the incompressibility, the horizontal velocity components are of order $1/A$ larger than the vertical components. Furthermore, we consider a flow enclosed in the domain, therefore each velocity mode $\boldsymbol{u}_I$ and $\boldsymbol{u}_E$ is constructed such that its normal component to the boundary of the domain is zero. Defining $\partial\Omega$ as the boundary of the domain, and $\hat{\boldsymbol{n}}(q)$ as the unit normal vector at $q \in \partial\Omega$ pointing outwards of the domain, this condition can be compactly written as

$$\boldsymbol{u}_I|_{\partial\Omega}\cdot\hat{\boldsymbol{n}} = \boldsymbol{u}_E|_{\partial\Omega}\cdot\hat{\boldsymbol{n}} \equiv 0. \tag{6}$$

Specifically, motivated by both models and observations [8, 22, 42, 49, 50, 56], we set the internal strength parameter $a_I(\langle\rho\rangle)$ to be proportional to the average density difference between two different regions of the basin, $D_1, D_2 \subset \Omega$ with $D_1 \cup D_2 = \Omega$. Defining for any function ϕ the averages $\langle\phi\rangle_i = \frac{1}{|D_i|}\int_{D_i}\phi dV$, $i = 1, 2$, the formula for a_I is

$$a_I(\langle\rho\rangle) = \Gamma\left(\langle\rho\rangle_2 - \langle\rho\rangle_1\right) = \Gamma\rho_0\left(-\alpha(\langle T\rangle_2 - \langle T\rangle_1) + \beta(\langle S\rangle_2 - \langle S\rangle_1)\right), \tag{7}$$

where Γ is a proportionality constant with units of velocity over density. This coupling has the advantage of retaining incompressibility of the composite flow assuming each of its components is individually incompressible. It also provides temporal separation between the two components - the changes in the internal flow amplitude are governed by spatial averages and thus have less time fluctuations than the external flow. Regarding the external flow, $a_E\boldsymbol{u}_E(\boldsymbol{r},t)$ can be any large-scale incompressible flow that does not depend on the tracers T and S, e.g. wind-driven surface flow, tidally driven flow, or any kinematic flow model.

The source terms f_T and f_S may be set to quantify sources of heat and salt in the defined oceanic basin, such as air-sea flux, exchange flow with marginal seas, river runoff, sea-ice formation and melting, or volcanic activity.

Motivated by oceanographic models, we consider a diffusion coefficient matrix $\boldsymbol{\kappa}$ that parameterizes background turbulence and small-scale eddy flow processes as an effective diffusion (eddy diffusivity, see Majda & Kramer [37]), rendering it orders of magnitude larger than molecular diffusivity. The diffusion constants are the same for both T and S, but differ between the horizontal $(\hat{x}, \hat{y})$ directions and the vertical $(\hat{z})$ direction, as isopycnal mixing is generally stronger than diapycnal mixing [19, 20] (isopycnals are taken here to be approximately horizontal):

$$\boldsymbol{\kappa} = \begin{pmatrix} \kappa_H & 0 & 0 \\ 0 & \kappa_H & 0 \\ 0 & 0 & \kappa_V \end{pmatrix}, \quad \kappa_V \ll \kappa_H. \tag{8}$$

The relevant orders of magnitude for the ocean are $\kappa_H \sim 10^3 m^2/sec$, $\kappa_V \sim 10^{-4} m^2/sec$ [19, 20, 37].

For a full formulation of the problem at hand, boundary conditions must be specified. In the oceanic basin, the air-sea interface is its most significant boundary in terms of heat and freshwater forcing. This interface has a significant impact on the temperature and salinity distribution in the ocean, and consequently its density-driven flows. The surface heat flux is related to the atmosphere-ocean temperature difference [27]; a warmer ocean surface will release heat to the atmosphere, cooling the ocean (and making it denser) while warming the atmosphere. The surface freshwater flux is related to the evaporation and precipitation rates such that net evaporation will result in higher surface salinity and thus in denser water. There is no direct feedback from ocean salinity on the atmosphere. Thus, it is common to use the so-called mixed boundary conditions for the surface [27, 56]: a Robin, also known as a relaxation, boundary condition for the temperature, in which the amount of heat flux depends on the air-sea temperature difference; and a Neumann boundary condition for the salinity. Neglecting (for now) geothermal heating and exchange flow with marginal seas, we consider a zero flux boundary condition for the sides and bottom of the box. Defining $z = L_z$ as the domain's surface boundary (air-sea interface), the appropriate boundary conditions are thus given by

$$(\boldsymbol{\kappa}\boldsymbol{\nabla}T)\cdot\hat{\boldsymbol{n}} = \begin{cases} g_A^T(T^*(x,y,t) - T) & \text{if } z = L_z \\ 0 & \text{else} \end{cases}, \tag{9}$$

$$(\boldsymbol{\kappa}\boldsymbol{\nabla}S)\cdot\hat{\boldsymbol{n}} = \begin{cases} g_A^S S^*(x,y,t) & \text{if } z = L_z \\ 0 & \text{else} \end{cases}. \tag{10}$$

g_A^T, g_A^S are the effective rates of convective heat and mass transfer at the boundary, respectively, with units of velocity. $T^*(x,y,t)$ and $S^*(x,y,t)$ are the temperature and salinity atmospheric sources (based, e.g. on observations), and are predetermined. Compatibility conditions on the box boundaries imply that at the box edges the normal derivatives of T^*, S^* must vanish (see appendix C), namely, that there is no flux to the shore. We note that the salinity forcing function $S^*(x,y,t)$ may be negative or positive, where a positive (negative) forcing signifies more (less) evaporation than precipitation. In order for the overall salinity to remain constant, we demand $\int_0^{L_x}\int_0^{L_y} S^*(x,y,t)dxdy = 0$ and $\int_\Omega f_S(\boldsymbol{r},t)dV = 0$.

Our model is given by system (1) - (10): a system of modified advection-diffusion equations for temperature and salinity with source terms, (1) and (2), that are weakly coupled by averages with a partially kinematic and partially dynamic incompressible velocity field (4). The system is subject to mixed Robin-Neumann boundary conditions for the temperature and salinity, (9) and (10). The diffusion matrix (8), parameterizing the eddy diffusivity, is diagonal, and the overall salinity is conserved in the basin throughout the evolution. The dynamic coefficient of the internal velocity, $a_I(\langle\rho\rangle)$, provides a natural observable of the dynamics. It reveals when the density-induced velocity component stabilizes to a steady state and when, as parameters are changed, the steady state bifurcates (and the solutions become bi-stable, oscillatory or, possibly, chaotic).

A. Non-dimensionalization and rescaling

System (1) - (10) has multiple natural timescales, associated with the horizontal and vertical eddy diffusion, the velocity field and the surface sources. Rescaling the problem to contain dimensionless parameters that demonstrate the ratios between these, we define the basic timescale $\tau \equiv L_z^2/\kappa_V$, describing the time it would take a tracer to cross the domain from surface to bottom via eddy diffusion only. We note that τ takes into account both molecular diffusion and eddy mixing on the small-scale (e.g. due to breaking of internal waves) that is parameterized by the vertical eddy diffusivity κ_V and is not included in $\boldsymbol{u}$, and therefore is an insightful timescale for the advection-diffusion equation. For typical oceanic values of $L_z \sim 4\times10^3 m$ and $\kappa_V \sim 10^{-4} m^2/sec$, the timescale is of the order $\tau \sim 5\times10^3 yrs$. We further define the temperature scale $T_\Delta^* = \max T^* - \min T^*$ and the salinity scale $S_\Delta^* = \max S^* - \min S^*$ as the maximal differences in the surface sources, with the additional conditions $T_\Delta^* \neq 0$, $S_\Delta^* \neq 0$, else the problem would not reflect the physical process we are modeling.

We perform a non-dimensionalization of the variables t, x, y, z, T, and S by $\tau, L_x, L_y, L_z, T_\Delta^*$, and S_Δ^*, respectively. Notice that the rescaling differs between the spatial dimensions, such that the rescaled domain is the symmetric cube $[0,1]^3$. Correspondingly, we rescale the dimensional functions T^*, S^*, f_T, f_S by $T_\Delta^*, S_\Delta^*, T_\Delta^*/\tau, S_\Delta^*/\tau$. Likewise, the dimensional parameters of $\boldsymbol{\kappa}$ are replaced by $\delta_x \equiv \frac{\kappa_H/L_x^2}{\kappa_V/L_z^2}$ in the $\hat{x}$ component, $\delta_y \equiv \frac{\kappa_H/L_y^2}{\kappa_V/L_z^2}$ in the $\hat{y}$ component, and 1 in the $\hat{z}$ component; we further define $\delta \equiv \min\{\delta_x, \delta_y\}$. Note that due to stratification strongly limiting cross-isopycnal flow, and due to the almost-2D nature of oceanic domains, it may occur that $\delta_x \sim \delta_y \sim 1$. g_A^T and g_A^S are replaced by $Nu \equiv \frac{g_A^T}{\kappa_V/L_z}$, the Nusselt number for heat transfer at the boundary, and $Sh \equiv \frac{g_A^S}{\kappa_V/L_z}$, the Sherwood number for mass transfer at the boundary, respectively. The $\hat{x}$, $\hat{y}$ and $\hat{z}$ components of the velocity field $\boldsymbol{u}$, as defined in equation (4), are rescaled, respectively, by L_x/τ, L_y/τ, and L_z/τ. Thus, by defining the Péclet number related to the external velocity field as $Pe \equiv \frac{a_E}{\kappa_V/L_z}$; the thermal Rayleigh number as $Ra^T \equiv \frac{\Gamma\rho_0\alpha T_\Delta^*}{\kappa_V/L_z}$; and the density stability ratio as $R_\rho \equiv \frac{\beta S_\Delta^*}{\alpha T_\Delta^*}$; the dimensionless, rescaled velocity field is given by

$$\begin{aligned} \boldsymbol{u}_{rs} &= Pe\boldsymbol{u}_{E,rs} + Ra^T a_{I,rs}\boldsymbol{u}_{I,rs}, \\ a_{I,rs}(\langle T\rangle, \langle S\rangle) &= \frac{a_I(\langle\rho\rangle)}{\Gamma\rho_0\alpha T_\Delta^*} = -\frac{\langle T\rangle_2 - \langle T\rangle_1}{T_\Delta^*} + R_\rho\frac{\langle S\rangle_2 - \langle S\rangle_1}{S_\Delta^*}, \end{aligned} \tag{11}$$

where rs signifies rescaled. Due to the rescaling, the three components of the rescaled velocity functions $\boldsymbol{u}_{E,rs}$ and $\boldsymbol{u}_{I,rs}$ are all of order 1. The difference in timescales between the externally-driven velocity $\boldsymbol{u}_{E,rs}$ and the density-driven velocity $\boldsymbol{u}_{I,rs}$ is controlled now by the two non-dimensional parameters Pe and Ra^T. Thus, in the rescaled time units, 1 is the timescale of eddy diffusion from surface to bottom, $1/Ra^T$ is the timescale of a vertical circulation around the basin, and $1/Pe$ is the timescale of externally-driven velocity contributions; typically, $1 \gg 1/Ra^T \gg 1/Pe$. Summarizing, Ra^T is the tuning parameter of the rescaled problem, and the rescaled variable $a_{I,rs}(\langle T\rangle, \langle S\rangle)$ is the dynamic observable of interest that results from the model. Hereafter we consider the nondimensional problem (and drop the rs subscript), unless indicated otherwise.

III. THE NORTH ATLANTIC OCEAN'S LARGE-SCALE FLOW

The North Atlantic's large-scale flow is built mainly from two velocity modes with a clear time-scale separation between them: a wind-driven double-gyre flow that decays with depth [57], and the AMOC [21, 31]. The wind-driven flow has an average period of the order of 5 years, while the AMOC has an average period of the order of 1000 years. The AMOC, unique to the

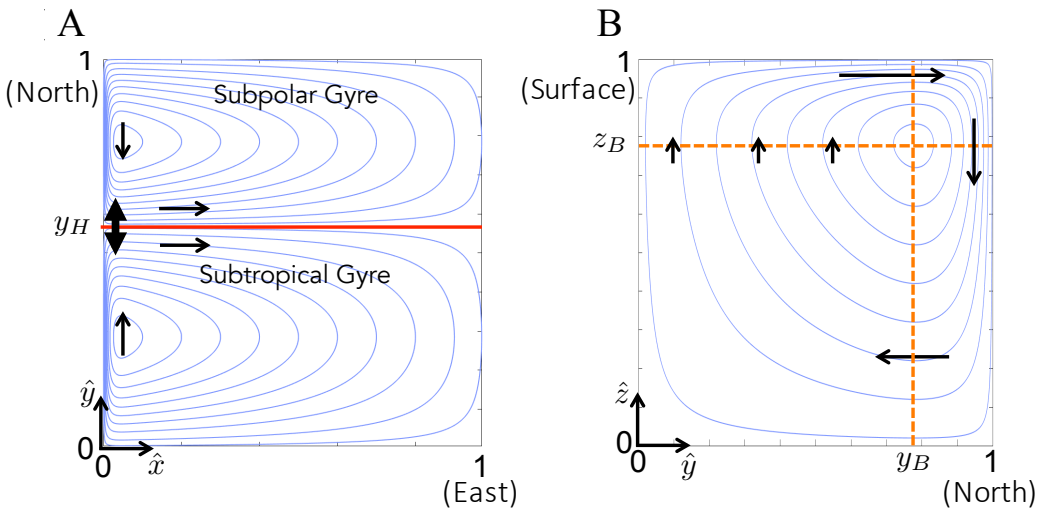

FIG. 1: Velocity field structure for the North Atlantic model in dimensionless form. (A) Surface flow $\boldsymbol{u}_E$, with a penetration depth $H_{TC} \sim 600m$. y_H is the inter-gyre demarcation line, and may be time-varying to model the seasonal north-south oscillations of the eastward branch of the Gulf Stream [18]. (B) Zonally averaged overturning flow $\boldsymbol{u}_I$. The flow has a northbound branch above z_B, and a southbound branch below. z_B is set to 0.78, corresponding to 890 meters. Downwelling occurs north of y_B, set to 0.78 for the North Atlantic and corresponding to 55°N.

Atlantic Ocean under present-day climate [17], is an important component of the global climate that transports approximately 18 Sv (Sverdrup, equals $10^6 \mathrm{m}^3$/sec) of warm surface waters from the tropics towards the North Pole, sinks in the northern latitudes and flows southwards in the abyss. The circulation is driven and maintained by a combination of wind-driven upwelling in the Antarctic region, upwelling throughout the ocean, and an anomalously high salinity in the northern latitudes causing a sinking of salty cold surface waters upon winter cooling [17]. The high salinity is commonly attributed to a combination of factors, including an abundance of salty sources, atmospheric forcing, and the salt-advection feedback loop: the AMOC sustains itself by advecting high-salinity surface waters from equatorial latitudes to the north. Evaluating the relative contribution of each of these factors to the structure, strength and variability of the AMOC is important in order to understand its (somewhat debated) stability in current, past and projected climates (see Weijer *et al.* [58], and references therein).

The extent of the effect of the feedback loop is related to the amount of salty waters that actually reach the northern latitudes. This depends on the time-dependent pathway statistics, affected both by the overturning circulation itself, and by chaotic advection resulting from the large-scale double-gyre surface flow of the basin, along with the other large-scale flow components in the basin, since the actual transport pathways differ from the mean flow [2]. Thus, the AMOC should be sensitive to density variations resulting from inhomogeneous external fluxes of salinity and temperature, that are transported in an interplay between diffusion, resulting in local mixing, and 3D chaotic advection, resulting in stirring [2, 6]. The extent of this sensitivity is unknown [58].

A. Applying the general model to the North Atlantic

In order to apply the general model, equations (1) - (10), to the North Atlantic, consider an idealized rectangular basin with a constant depth of $4 \times 10^3 m$ and straight edges, and a horizontal extent of order $L \sim O(10^6 m)$ fitting the longitudinal region $[0°W,\ 60°W]$ and the latitudinal region $[0°N,\ 70°N]$. Thus, the aspect ratio is $A \sim 10^{-3}$, similar to that of a sheet of paper's thickness vs. length. See appendix B for a comprehensive list of the parameters we used for the North Atlantic simulations. We note that in this work, we neglect the spherical geometry of the North Atlantic, however it is easy to see that the same framework can be naturally extended to a spherical basin.

The two major contributions to the large-scale flow in the basin are the horizontal wind-driven surface flow in the horizontal (x, y) direction, representing the North Atlantic sub-tropical and sub-polar gyres, contained in the external velocity field $\boldsymbol{u}_E(\boldsymbol{r}, t)$; and the vertical density-driven overturning flow in the latitudinal-abyssal $(\hat{y}, \hat{z})$ direction, contained in $\boldsymbol{u}_I(\boldsymbol{r})$. The surface double-gyre flow $\boldsymbol{u}_E$ is a solution to the Sverdrup balance with a boundary as used in Yang & Liu [60], see figure 1A. It exhibits a westward-biased asymmetry that models the western boundary flow, and a finite penetration depth of a few hundred meters, corresponding to the approximate measured depth of the oceanic thermocline. The time-scale associated with the surface flow is of the order of a decade. In some of the simulations, we also consider the effect of a seasonal north-south variation of the demarcation line between the gyres with an amplitude of approximately 1° in latitude, with the north-most position obtained in

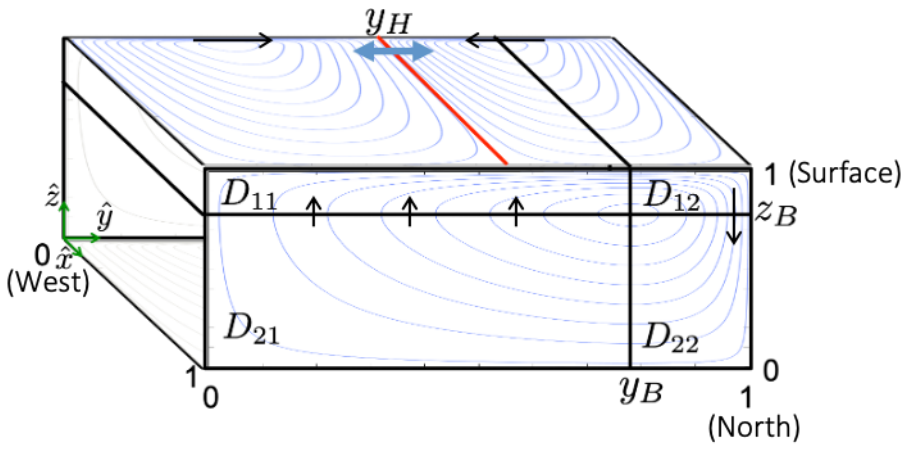

FIG. 2: 3D flow and division of basin into boxes after rescaling. The division is chosen such that the line y_B, z_B is the overturning gyre's center; thus, between D_{11} and D_{12} (D_{21} and D_{22}) there is only northbound (southbound) flow, and between D_{12} and D_{22} (D_{11} and D_{21}) there is only downwelling (upwelling).

the fall season [18]. The overturning flow $\boldsymbol{u}_I$ is set as a zonally ($\hat{x}$) independent gyre in the $\hat{y}-\hat{z}$ plane, with a northward branch above the thermocline, a downwelling branch around the northern basin border, a southward branch below the thermocline and semi-uniformly upwelling south of the downwelling region. To model this flow, we use a similar functional form as the surface flow, see figure 1B. Such a loop is completed in approximately 1000 years [31]. The separation in time scales between the two velocity modes further justifies the decomposition of the full velocity field into two decoupled contributions. The exact functions we use for these flows are presented in appendix A, and the parameter values that mimic the North Atlantic flow are presented in appendix B.

We note that the AMOC does not actually have a southern border at the equator; although the transport loop does indeed attain partial closure to the north of the equator by upwelling, most of the southbound transport branch continues all the way south to Antarctica, where it participates in a complicated interplay with the Antarctic Circumpolar Current and eventually returns as a northbound branch [31]. We follow previous studies that consider a similar configuration: Cessi [11], Stommel [50], Tziperman *et al.* [56], and others. Yet, our model could take this into account by inserting a source term at the southern border that parameterizes the appropriate inputs and outputs of temperature and salinity through the equatorial basin wall; alternatively, the velocity field $\boldsymbol{u}_I$ itself could have an open southern border. These options are left for exploration in later works.

By construction, both flow vector field components $\boldsymbol{u}_E$ and $\boldsymbol{u}_I$ are of order 1 and non-dimensional, as described in section II A. In the composite flow, they are multiplied by their (dimensionless) respective strength coefficients Pe and $Ra^T a_I(\langle T\rangle, \langle S\rangle)$. The parameters are set such that $Pe \sim 10^6$ corresponds to realistic sub-tropical gyre velocities with a circulation time of approximately 5 years, and $Ra^T \sim 10^3$ corresponds to realistic AMOC velocities with a circulation time of approximately 1000 years and an AMOC strength of approximately 18 Sv. The vertical strength parameter $a_I(\langle T\rangle, \langle S\rangle)$, of order 1 due to the rescaling, is determined dynamically by the average density difference between the north part and the south part of the basin. We first divide the rescaled basin into four boxes, $\{D_{ij}\}_{i,j=1,2}$ as depicted in figure 2, such that $D_{11} = [0,1]\times[0,y_B]\times[z_B,1]$ (south-top), $D_{12} = [0,1]\times[y_B,1]\times[z_B,1]$ (north-top), $D_{21} = [0,1]\times[0,y_B]\times[0,z_B]$ (south-bottom), $D_{22} = [0,1]\times[y_B,1]\times[0,z_B]$ (north-bottom); the division is according to the overturning gyre's center, where the flow switches from upwelling to downwelling and from northbound to southbound (figure 2). The average north-bound velocity between the two surface boxes D_{11} and D_{12} is then proportional to the overturning strength parameter $a_I(\langle T\rangle, \langle S\rangle)$.

The realistic atmospheric sources for the North Atlantic differ between the southern and the northern sections of the basin. While the northern latitudes in the ocean are colder and fresher, the southern latitudes are warmer and saltier. This induces a competition between the average densities in each section, with the temperature forcing pulling towards a denser north and the salinity forcing pulling towards a denser south. In most box models, this competition allows for two stable steady states of the overturning flow given the same atmospheric forcing with one directed clockwise (sinking branch in the north) and the other directed counter-clockwise (sinking branch in the south); this bi-stability is of great interest in the context of the actual North Atlantic flow, as its existence and sensitivity have a direct impact on the AMOC stability and on our climate system [45, 50, 55]. In section IV we explore such possibilities in our model.

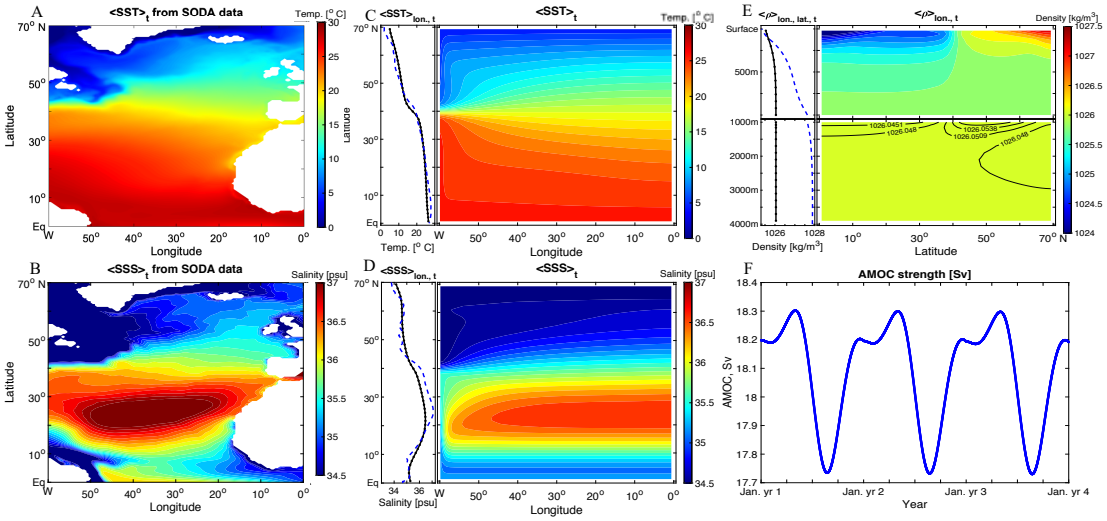

FIG. 3: Simulation results of Atlantic Ocean model, in comparison with realistic reanalysis experimental data. (A-B) SODA3.4.2 reanalysis experiment data [10], averaged over 1980-2017: (A) depicts sea-surface temperature (SST), and (B) depicts sea-surface salinity (SSS). (C-F) Simulation results: (C) SST [$^\circ C$], (D) SSS [psu], (E) zonally-averaged density [kg/m^3], and (F) simulation result of the AMOC strength [Sverdrup (Sv)], proportional to $a_I(\langle\rho\rangle)$ of the model. The black curves on the left of (C) and (D) are longitudinal averages of the colored data. The black curve on the left of (E) is a zonal average of the colored data. The blue dashed curves are the corresponding averages from the SODA3.4.2 data.

B. Numerical results for the North Atlantic example

Numerical solutions of the model are obtained step-wise: the equations for temperature and salinity (1) and (2) are solved using the previous step's velocity field, then the velocity field is updated given the new T, S functions. This 3D, optionally time-varying, non-linear, coupled PDE system, with multiple time-scales, requires a customized numerical scheme, which we have implemented in Matlab. The scheme employs a finite volume integration scheme, which allows for a differential box size and has good conservation properties. The diffusion and source terms are straightforward to simulate using standard discretization methods. However for the advection term one needs to retain numerical stability on the one hand, while reducing numerical diffusion on the other hand. For example, the simple first-order upwind scheme has excellent stability properties, but large numerical diffusion. On the other hand, straight-forward higher order schemes introduce spurious oscillations that may cause the numerical scheme to diverge. To circumvent these issues, we use the stable flux-limiter advection scheme, combining the low-resolution upwind scheme with the high-resolution Lax-Wendroff scheme, weighted with a Sweby flux limiter with $\beta = 1.5$ [51]. The spatial and temporal grid sizes were determined to be below the Courant-Friedrichs-Lewy (CFL) stability condition [13], and fine enough to resolve the small structure elements of the velocity field. The final grids we used were checked for sensitivity of results to grid size changes. The simulations shown here are run with a given parameter set until a quasi-stable steady state/periodic state is reached; results shown are of the stable solution.

To model the North Atlantic basin, we set the parameters to include an oscillating demarcation line between the gyres and a seasonally oscillating surface forcing; see table I in appendix B for a list of the parameter values we use, and figure 3 for the numerical results. In order to tune the surface forcing, we used the SODA3.4.2 reanalysis experiment [10], extracting the climatological average from 1980 to 2017 and deriving the zonally averaged data of the sea-surface temperature (SST) and salinity (SSS). This realistic data was used as a restoring (Robin) boundary condition for both the temperature and the salinity. After approximately 1000 simulation years, the system reached a quasi stable time-periodic state. Then, we switched the salinity restoring force with a corresponding constant flux forcing (Neumann) boundary condition. This is a common procedure for tuning realistic systems with mixed Robin-Neumann boundary conditions, e.g. Tziperman *et al.* [56]. Tuning the overturning strength parameter Γ, we obtained a reasonable distribution of sea-surface temperature (SST), sea-surface salinity (SSS), zonally averaged density (except at the northern-most latitudes, where we get density inversions), and AMOC overturning strength (figure 3). The temperature and salinity fronts observed between the two gyres are due to the velocity field, and are clearly apparent also in the realistic SST and SSS observations (see e.g. Mercator ocean analysis 2008 [14]); see discussion in section VI regarding the similarities and differences between the model results and the realistic data.

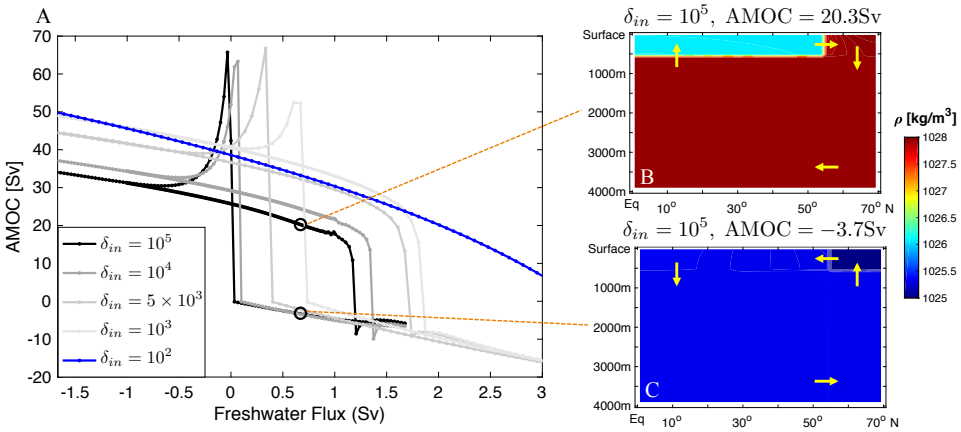

FIG. 4: Simulation results of the 2D ($Pe = 0$) autonomous model. (A) Bi-stability of the steady state box-like solutions as a function of freshwater flux (see main text for its definition). The inter-box diffusion varies between the curves, whereas the inner-box diffusion is constant. Thus their ratio δ_{in} varies between the curves; see main text for details. A hysteresis loop is clearly discernible for $\delta_{in} \geq 10^3$, corresponding to a thermal Rayleigh number of $Ra^T \geq 2 \times 10^3$. As δ_{in} (and, correspondingly, Ra^T) increases, the hysteresis loop moves toward higher freshwater flux values, and the positive overturning strength given a freshwater flux value grows. At $\delta_{in} = 100$ ($Ra^T = 200$) bistability is lost and a single steady state appears for a given forcing value, as expected from Theorem 2(iii), presented and proven in the main text. (B, C) Density distribution in two steady states in the bistable regime, for freshwater forcing of 0.67 Sv and Rayleigh number 2×10^5. The overturning strength of (B) is 20.3 Sv, the overturning strength of (C) is −3.7 Sv.

IV. RELATION TO SIMPLIFIED BOX MODELS

Our system, equation set (1) - (10), is a natural extension of the 2×2 box model presented in Huang *et al.* [28], Tziperman *et al.* [56], extended to take into account density perturbations resulting from 3D advection, diffusion and forcing. Under the assumptions of inner-box fast mixing, and inter-box interactions resulting only from advection, and using $\boldsymbol{u}_I$ from the motivating example (figure 1B) as the internal velocity in equation (4), the system reduces to the four-box model. To show this, we use the basin division presented in figure 2, and define the instantaneous box averages $T_{ij}(t) \equiv \frac{1}{|D_{ij}|}\int_{D_{ij}} T(\boldsymbol{r},t)dV$, and a function describing the density perturbations from the box averages, $T'(\boldsymbol{r},t)$:

$$T(\boldsymbol{r},t) = \begin{cases} T_{11}(t) + T'(\boldsymbol{r},t) & \text{if } \boldsymbol{r} \in D_{11} \\ T_{12}(t) + T'(\boldsymbol{r},t) & \text{if } \boldsymbol{r} \in D_{12} \\ T_{21}(t) + T'(\boldsymbol{r},t) & \text{if } \boldsymbol{r} \in D_{21} \\ T_{22}(t) + T'(\boldsymbol{r},t) & \text{if } \boldsymbol{r} \in D_{22} \end{cases},$$

and the same for the salinity function. While $T(\boldsymbol{r},t)$ and $S(\boldsymbol{r},t)$ are smooth, $T'(\boldsymbol{r},t)$ and $S'(\boldsymbol{r},t)$ are only piecewise-smooth. The fast mixing assumption is equivalent to assuming $T'(\boldsymbol{r},t)$ and $S'(\boldsymbol{r},t)$ are small at all times; the idealistic assumption of immediate inner-box mixing would correspond to $T'(\boldsymbol{r},t) = S'(\boldsymbol{r},t) \equiv 0$. Employing these definitions, we integrate equation (1) with $f_T = 0$ over D_{ij} and divide by its volume. Using Gauss theorem and applying the boundary conditions on the dimensionless problem, and using the upwind advection scheme, we obtain the following equations:

$$\begin{aligned} \dot{T}_{11} &= \frac{Nu}{1-z_B}(T^*_{South} - T_{11}) + \frac{1}{y_B}(T_{21} - T_{11})\,V + f_{11}(\delta_y, T') \\ \dot{T}_{12} &= \frac{Nu}{1-z_B}(T^*_{North} - T_{12}) + \frac{1}{1-y_B}(T_{11} - T_{12})\,V + f_{12}(\delta_y, T') \\ \dot{T}_{21} &= \frac{1}{y_B}(T_{22} - T_{21})\,V + f_{21}(\delta_y, T') \\ \dot{T}_{22} &= \frac{1}{1-y_B}(T_{12} - T_{22})\,V + f_{22}(\delta_y, T') \end{aligned} \tag{12}$$

where $T^*_{South} = \frac{1}{y_B}\int_0^1 dx \int_0^{y_B} dy\; T^*(x,y)$, $T^*_{North} = \frac{1}{1-y_B}\int_0^1 dx \int_{y_B}^1 dy\; T^*(x,y)$ are the average surface fluxes in the south and north, correspondingly; and $V = \frac{1}{1-z_B}\int_0^1 dx \int_{z_B}^1 dz\; v|_{y=y_B}$ is the average velocity between the two top boxes. The averaged

velocity V is itself a closed function of the averaged functions T_{ij} and S_{ij},

$$V = Ra^T a_I(T_{ij}, S_{ij}) V_I; \quad V_I = \frac{1}{1-z_B}\int_0^1 dx \int_{z_B}^1 dz\, \boldsymbol{u}_I \cdot \hat{y}|_{y=y_B},$$
$$a_I(T_{ij}, S_{ij}) = -\frac{1}{T^*_\Delta}\left(\frac{\sum_{j=1}^2 D_{j2}T_{j2}}{\sum_{j=1}^2 D_{j2}} - \frac{\sum_{j=1}^2 D_{j1}T_{j1}}{\sum_{j=1}^2 D_{j1}}\right) + \frac{R_\rho}{S^*_\Delta}\left(\frac{\sum_{j=1}^2 D_{j2}S_{j2}}{\sum_{j=1}^2 D_{j2}} - \frac{\sum_{j=1}^2 D_{j1}S_{j1}}{\sum_{j=1}^2 D_{j1}}\right).$$

Note that V_I is a constant determined by the box division and the chosen overturning velocity field form $\boldsymbol{u}_I$; it is basically the average velocity between boxes D_{11} and D_{12} if $Ra^T a_I = 1$. In equations (12), we assume $a_I > 0$, and use the upwind scheme: since the transport between the boxes is unidirectional by construction of the boxes, we assume V transports temperature and salinity from T_{11} to box D_{12} and so on for transport between the other boxes. Thus, if $a_I < 0$ the overturning flow switches direction and, just like in the 4-box model, the velocity terms switch signs and the advected values change accordingly.

The remaining terms $f_{ij}(\delta_y, T')$ include the diffusion between the boxes and the advection of T', which is the deviation in each box from the mean value inside the box. For example, for T_{11},

$$f_{11}(\delta_y, T') = \frac{1}{|D_{11}|}\int_0^1 \left(\int_0^{y_B} \left(w_I T' + \frac{\partial T}{\partial z}\right)_{z=z_B} dy + \int_{z_B}^1 \left(-v_I T' + \delta_y \frac{\partial T}{\partial y}\right)_{y=y_B} dz - \int_0^{y_B} (NuT')_{z=1} dy\right) dx.$$

The equations for S_{ij} are similar, except for the surface boundary conditions which have a constant flux for the salinity. If T', S' and inter-box diffusion are neglected, $f_{ij} \equiv 0$ and the equation set (12) attains closure; along with the corresponding salinity equations, these are exactly the 2×2-box equations, with known steady-state solutions [28, 56].

To investigate this parameter limit, we performed a 2D autonomous numerical experiment; its results are summarized in figure 4, and the parameters used are summarized in appendix B. Specifically, we set $Pe = 0$, thus shutting off the external velocity field, to obtain a 2D system. Furthermore, we set high diffusion values inside each box, and small diffusion values between the boxes. The surface temperature function $T^*(x, y)$ was autonomous and piecewise constant, set as $+4°C$ over the south box $y < y_B$ and $-14°C$ over the north box $y > y_B$. The salinity temperature function $S^*(x, y)$ was also autonomous and piecewise constant such that $S^*(x, y < y_B) = -\frac{y_B}{1-y_B} S^*(x, y > y_B)$ (condition to conserve salinity). Its values were varied throughout the experiment, as shown in figure 4, where we define the northern freshwater flux with dimensional parameters as $FWF[\mathrm{Sv}] = L_x L_y g^S_A \frac{\langle S^*\rangle_{y>y_B}}{35.5\mathrm{psu}}/10^6$. In the experiment, for each value of FWF we let the system reach a (seemingly) steady state, and denoted the value of the AMOC strength (in units of Sv, proportional to $a_I(\langle\rho\rangle)$), then used this steady state as the initial condition for the next adjacent FWF value. Note that in this experiment, since we use a spatially-varying diffusion matrix, the temporal rescaling of section II A is not well-determined. Hence, we set the rescaling timescale to correspond to the minimal vertical eddy diffusivity: $\tau = L_z^2/\kappa^z_{out}$, where $\kappa^z_{out} = \min_{\boldsymbol{r}}\{\kappa_V(\boldsymbol{r})\}$ is the inter-box diffusion in the $\hat{z}$ direction. We further define $\delta^i_{in} = \frac{\kappa^i_{in}/L_i^2}{1/\tau}$, $i = x, y, z$ in the same spirit as δ_x, δ_y as defined in section II A. We set up the simulation parameters such that these dimensionless diffusion ratios are equal: $\delta^x_{in} = \delta^y_{in} = \delta^z_{in} \equiv \delta_{in}$; and we compare values of δ_{in} ranging between $10^2 - 10^5$. We note that the thermal Rayleigh number Ra^T scales like τ, and therefore scales like δ_{in}.

At the largest value of δ_{in}, where the Rayleigh number is maximal and mixing between boxes is negligible with respect to inner-box diffusion, we obtain equivalent solutions to the 2×2 box model of Huang *et al.* [28], Tziperman *et al.* [56], including a regime of bi-stability that leads to a hysteresis loop upon a continuous variation of the FWF parameter, as seen in figure 4. In the bi-stable regime, the two locally stable steady state solutions are a positive AMOC and a small, negative AMOC, similar to figure 2 in Rahmstorf [45]. As δ_{in} decreases from 10^5 to 10^2 and, correspondingly, the Rayleigh number decreases from 2×10^5 to 2×10^2, the hysteresis curve shifts until bi-stability is lost when the Rayleigh number is low enough, as expected from Theorem 2(iii) presented in the next section. Using similar methods, with this framework we can reproduce various previously-studied box models with any number of boxes, with a controllable amount of inter-box diffusion and inner-box inhomogeneities: all that is required is a tailoring of a suitable internal velocity field $\boldsymbol{u}_I$ with transport in the desired direction between the various boxes.

As stated previously, even when the diffusion is homogenous in the full domain, there may be cases where neglecting T' and S' is justified. For example, ignoring the inter-box diffusion can be justified if there are some barriers for mixing between the boxes. Such barriers may be induced by coherent structures localized in the boxes, rendering the eddy diffusivity across such coherent structures effectively negligible. To demonstrate this idea, we performed an autonomous 3D experiment where we set homogenous eddy diffusivity throughout the basin, and aligned the inter-gyre demarcation line of the horizontal velocity field with the horizontal barrier between the boxes y_B. Thus, each box is relatively mixed due to the combination of velocity fields and

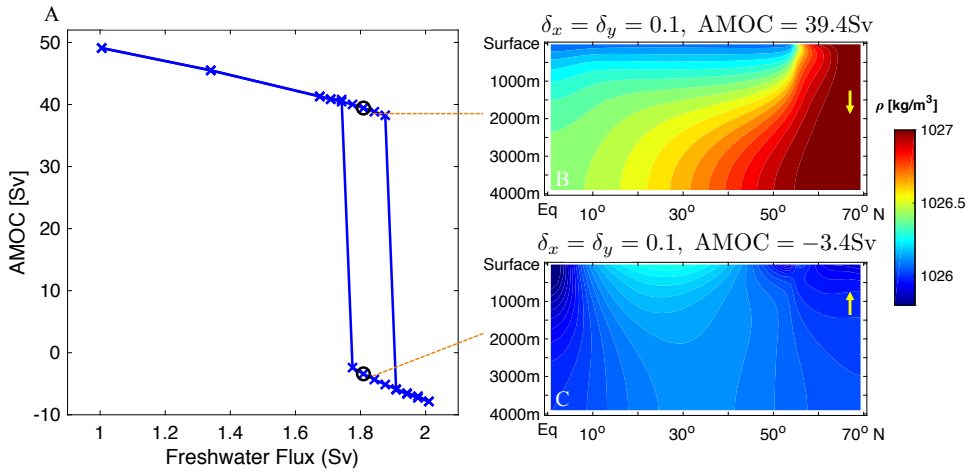

FIG. 5: Simulation results of the 3D autonomous model. (A) Bi-stability of steady state solutions as a function of freshwater flux with realistic eddy diffusivity. The mixing between the north and south boxes is reduced by the horizontal gyre flow. In all data points, the external parameters are the same, as is the diffusion, set such that $\delta_x = \delta_y = 0.1$. For a narrow range of freshwater forcing, the system displays bistability. (B, C) Density distribution in two steady states in the bistable regime, for freshwater forcing of 1.8 Sv. The overturning strength of (B) is 39.4 Sv, the overturning strength of (C) is −3.4 Sv.

vertical diffusion, while interbox mixing is minimized due to the velocity field structure. The results are summarized in figure 5, and the parameters used are summarized in appendix B. The temperature and salinity surface functions were taken as in the 2D autonomous experiment described above in this section. As in the 2D experiment, we varied the FWF parameter continuously and plotted the resulting steady-state AMOC strength. The resulting hysteresis loop and steady states bear a qualitative resemblance to those obtained in the 2D experiment; see figure 5.

As in previous studies of hysteresis AMOC loops in GCM models [45], finding these hysteresis loops requires some tweaking of parameters. Thus, some of the parameters of figures 4 and 5 differ from those of figure 3 (see table I). In particular, notice that the spatial dependence of the velocity field changes considerably the bi-stable range. A more comprehensive study of the range of realistic geophysical settings in which bi-stability is possible in our model (possibly including unsteady forcing) and its relation to corresponding parameter ranges in box models and in full GCM models is left for future research.

V. WELL-POSEDNESS OF THE MODEL

It is well known that the linear evolution problem for the inhomogeneous 3D advection-diffusion equation subject to Neumann or Robin boundary conditions, has a unique solution if the initial values, boundary values and coeffiecients satisfy some smoothness properties (see, e.g., Nittka [43], and references therein). On the other hand, the 3D Boussinesq system describing ocean flow is still lacking a proof of existence of global, smooth solutions. Our system (1) - (10), is a non-trivial extension of the regular advection-diffusion equation due to its non-local coupling. Due to this addition, although the velocity field itself (4) is smooth (by construction), there is a potential problem in the formulation coming from the dynamic coupling term a_I: *a priori*, there is no guarantee, for example, that the gradient of the temperature and salinity functions $T(\boldsymbol{r}, t)$ and $S(\boldsymbol{r}, t)$, and the corresponding coupling strength function a_I, remain finite and do not exhibit blow up in finite times.

In this section we present the theorem that guarantees the global, for all times and initial data, existence and uniqueness of well-behaving solutions. Furthermore, we find several bounds on the temperature and salinity functions, their derivatives, and the coupling term a_I itself, that may be useful for numerical purposes, allowing some distinction between numerical errors and real features. We prove that there is a bound on the thermal and salinity Rayleigh numbers with respect to a function of the other dimensionless parameters of the problem, such that if the former is large enough with respect to the latter, all initial conditions will eventually converge to a single, stable steady-state solution. Finally, relatively simple generalizations of the system are subject to similar proofs, and are quite wide-spread in various applications (see discussion in section VI).

In the next part we use the following notation: For some function $Q(\boldsymbol{r}) \in L^2(\Omega)$, we denote

- $\|Q\| \equiv \left(\int_\Omega Q(\boldsymbol{r})^2 dV\right)^{1/2}$, the L^2 norm of Q on the domain Ω;
- $\tilde{\|}Q\tilde{\|} \equiv \left(\iint Q\,(x, y, z = 1)^2\, dxdy\right)^{1/2}$, the L^2 norm of Q on the surface part of the boundary of Ω, defined as $z = 1$ in the

rescaled problem.

The main claim we prove regards the well-posedness in the sense of Hadamard of system (1)-(10), in the case of constant forcing and a constant external velocity. We note that the proof can be easily extended to time-dependent forcing and external velocity, as long as the relevant spatial norms are bounded in time: see remark 2. The exact mathematical formulation of the problem and the theorems appears in appendix C; here we present the main ideas.

Theorem 1 *Given initial conditions T_0, S_0 that satisfy assumptions (C1):*

(i.) The evolution problem described by equations (1) - (10) has a unique, global-in-time strong solution, denoted (T, S).

(ii.) This solution depends continuously on initial conditions, boundary conditions and sources in the sense described by equation (C26).

(iii.) For every $\bar{\mathcal{T}} > 0$, the rescaled, non-dimensional solution satisfies the following bounds:

$$\begin{aligned} \|T\|^2(t) &\le C_1^T; \quad \int_0^{\bar{\mathcal{T}}} \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}T\right\|^2(t)dt \le C_2^T + \bar{\mathcal{T}}C_3^T; \\ \|S\|^2(t) &\le C_1^S; \quad \int_0^{\bar{\mathcal{T}}} \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S\right\|^2(t)dt \le C_2^S + \bar{\mathcal{T}}C_3^S; \end{aligned} \tag{13}$$

where

$$\begin{aligned} C_1^T &\equiv \|T_0\|^2 + \frac{2Nu}{\min\{Nu,1\}(1-a)}\tilde{\|}T^*\tilde{\|}^2 + \frac{4}{\min\{Nu^2,1\}a(1-a)}\|f_T\|^2; \\ C_2^T &\equiv \frac{1}{2}\|T_0\|^2; \\ C_3^T &\equiv \frac{\min\{Nu,1\}}{8}\|T_0\|^2 + Nu\tilde{\|}T^*\tilde{\|}^2 + \frac{4}{\min\{Nu,1\}}\|f_T\|^2; \\ C_1^S &\equiv \|S_0\|^2 + \frac{0.27}{b}\frac{Sh^2}{\min\{1,\delta^2\}}\tilde{\|}S^*\tilde{\|}^2 + \frac{0.06}{\left(\frac{1}{2}-b\right)}\frac{1}{\min\{1,\delta^2\}}\|f_S\|^2; \\ C_2^S &\equiv 1.6\min\{\delta,1\}\|S_0\|^2; \\ C_3^S &\equiv 1.6\min\{\delta,1\}\|S_0\|^2 + \frac{6\ Sh^2}{\min\{\delta,1\}}\tilde{\|}S^*\tilde{\|}^2 + \frac{3.9\ Sh}{\min\{\delta^2,1\}}\|f_S\|^2, \end{aligned} \tag{14}$$

for $a = \left(1+\sqrt{1+2Nu\frac{\tilde{\|}T^*\tilde{\|}^2}{\|f_T\|^2}}\right)^{-1}$ *and* $b = \frac{1}{2}\left(1+\frac{1}{Sh\sqrt{2+\pi}}\frac{\|f_S\|}{\tilde{\|}S^*\tilde{\|}}\right)^{-1}$.

(iv.) If, in addition, the initial conditions satisfy (C4), namely, if they satisfy the boundary conditions and are sufficiently regular, then the problem has a unique, global-in-time classical L^2-solution.

The definitions of strong and classical L^2 solutions, an exact list of the regularity requirements on the system parameters and functions, and the full proof of the theorem appear in appendix C. Here we present the proof outline. The proof of existence relies on the concept of iteration. As a first step, given initial conditions $(T_0(\boldsymbol{r}), S_0(\boldsymbol{r}))$, we construct a sequence of solutions $\{(T_n(\boldsymbol{r},t), S_n(\boldsymbol{r},t))\}_{n=1}^{\infty}$ to the iteratively-defined, linear set of advection-diffusion equations

$$\partial_t T_n - \boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}T_n) + (\boldsymbol{u}_{n-1}\cdot\boldsymbol{\nabla})T_n = f_T; \quad \partial_t S_n - \boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}S_n) + (\boldsymbol{u}_{n-1}\cdot\boldsymbol{\nabla})S_n = f_S, \tag{15}$$

where $\boldsymbol{u}_{n-1} \equiv \boldsymbol{u}(\boldsymbol{r}; S_{n-1}, T_{n-1})$ is determined from the previous step, according to equation (4). The above system is subject to the boundary conditions of the original system, (9) and (10), with T_n and S_n in place of T and S, respectively. Note that these are linear PDEs subject to mixed Robin and Neumann boundary conditions, and the existence and uniqueness of a solution (T_n, S_n) is guaranteed in Nittka [43]. This solution is then proven to satisfy certain energy estimates from which we deduce bounds on norms of the solutions, that are independent of n. Using these bounds, we show that this iterative sequence is a Cauchy sequence converging strongly to functions that solve the PDEs and satisfy the boundary conditions for short times. We can show that the relevant norms of the solutions remain finite on the maximal interval of existence, and this guarantees the global existence of solutions. Next, the idea of a maximal interval of existence is used in order to show that these solutions are in fact global in time. Uniqueness of these solutions is shown along with smooth dependence on initial and boundary conditions by bounding the

difference between two solutions with a bound that goes to zero when the difference in their initial and boundary conditions goes to zero.

We emphasize that the bounds in Theorem 1 are generally true for temperature and salinity functions that satisfy equations (1) - (10), with any velocity field that is incompressible and non-penetrating, irrespective of the method through which the velocity field is achieved.

A straightforward conclusion from the proof of Theorem 1 is that the dynamical overturning strength parameter a_I is bounded:

Corollary 1 *For a given rescaled, non-dimensional solution (T, S) of problem (1) - (10), the dynamic weight of the velocity function, $a_I(\langle T\rangle, \langle S\rangle)$, is bounded at all times:*

$$|a_I(\langle T\rangle, \langle S\rangle)| \leq \frac{\sqrt{C_1^T} + R_\rho\sqrt{C_1^S}}{\min\{|D_1|, |D_2|\}}, \tag{16}$$

where the values of C_1^T and C_1^S are given in (C6).

Finally, we discuss steady-state solutions of the problem. We show that for any choice of time-independent sources, parameters and external velocity functions, there exists a steady state solution to the system. Note that for a general choice of parameters, multiple steady-state solutions to our nonlinear problem are expected to coexist, as we show numerically in Section IV. However, as is expected in advection-diffusion-type problems, when the system is not vigorously forced with respect to its dissipation, the steady-state solution *is* unique; furthermore, solutions to the time-dependent problem converge to this unique steady state solution as $t \to \infty$. These results are summarized in the following theorem:

Theorem 2 *i. The nonlinear steady-state problem corresponding to (1) - (10) has a weak solution, denoted (T, S).*

ii. The rescaled, non-dimensional solution satisfies the following bounds:

$$\|T\|^2 \leq C_4^T, \quad \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}T\right\|^2 \leq C_5^T, \quad \|S\|^2 \leq C_4^S, \quad \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S\right\|^2 \leq C_5^S; \tag{17}$$

where

$$\begin{aligned} C_4^T &= \frac{1}{16\min\{Nu,1\}^2}\|f_T\|^2 + \frac{Nu}{4\min\{Nu,1\}}\tilde{\|}T^*\tilde{\|}^2, \quad C_5^T = \frac{1}{4\min\{Nu,1\}}\|f_T\|^2 + \frac{2Nu}{3}\tilde{\|}T^*\tilde{\|}^2, \\ C_4^S &= \frac{1}{\epsilon_1}\tilde{\|}S^*\tilde{\|}^2 + \frac{1}{\epsilon_2}\|f_S\|^2, \quad C_5^S = \frac{1}{\epsilon_1}\tilde{\|}S^*\tilde{\|}^2 + \frac{1}{\epsilon_3}\|f_S\|^2. \end{aligned} \tag{18}$$

The constants are given by $\epsilon_1 = 2\min\{\delta,1\}/Sh^2$, $\epsilon_2 = 4\min\{\delta,1\}(3\pi/8-1)$, $\epsilon_3 = 2(\min\{\delta,1\}(\pi-2)+1)$.

iii. Let T_0, S_0 be initial conditions that satisfy assumptions (C1), and let $(\tilde{T}, \tilde{S})$ be a global strong solution to equations (1) - (10), as established in Theorem 1. If the following condition is satisfied by the problem parameters:

$$Ra^T\frac{\max\{R_\rho,1\}}{\min\{\delta,1\}} \leq \min\{|D_1|,|D_2|\}\frac{\min\{4\pi\delta, Nu, 1\}}{8\max\left\{C_5^T, C_5^S\right\}}, \tag{19}$$

where C_5^T, C_5^S are given in equation (18), then $(\tilde{T}, \tilde{S})$ converges to a unique steady-state solution as $t \to \infty$.

Theorem 2(iii) proves that the bistability of the system, a robust feature of known models of density-driven flows with both salinity and temperature forcing, is lost if the minimal diffusion in the domain is large enough with respect to the advection. This ratio is represented by $Ra^T/\min\{\delta,1\}$. Since the salinity Rayleigh number is $Ra^S = Ra^T R_\rho$, Theorem 2(iii) provides a bound on both the thermal and the salinity Rayleigh numbers. In the special case of no external sources, $f_T = f_S = 0$, equation (19) simplifies; for example, we can immediately conclude that the bound on $Ra^T\max\{R_\rho,1\}$ is smaller than both $1/\tilde{\|}T^*\tilde{\|}^2$ and $1/Sh^2\tilde{\|}S^*\tilde{\|}^2$.

We note that Theorems 1 and 2 and Corollary 1 provide rigorous bounds that are not in general tight, since in their formulation, the "worst case scenario" of the most extreme functions allowed by the problem formulation is considered. This is illustrated by Corollary 1, where the bound for $a_I(\langle T\rangle, \langle S\rangle)$ does not depend on gradients or differences of T^* and S^*, although it is obvious that if these are constant, there will be no overturning circulation after sufficient time. Also, one may calculate the numerical values of the bounds for the parameters used in the simulations. For example, Corollary 1 promises that the overturning strength of the steady state solution in figure 5 for freshwater forcing 1.8 Sv is smaller than 900 Sv - this is indeed the case. However, this bound is not very useful, as the asymptotic overturning strength is found numerically to be smaller than 25 Sv. On the other hand, qualitatively the statements are sound. For example, the transition promised by Theorem 2(iii) between a unique steady state and non-trivial dynamics is indeed observed in the simulations: While for a large Rayleigh number, bistable solutions are found, when the minimal diffusion in the basin is large enough with respect to the advection coefficient as encapsulated by a small enough Rayleigh number, bistability is lost and all solutions seem to converge to a unique steady state; see figures 4 and 5.

Remark 1 *Generalizations.*
The proof presented here can be generalized to the following cases:

- *Any number of boxes $n \in \mathbb{N}$, and any linear dependence of the coupling on the tracer functions T and S, which can be different in each box. Thus, the internal velocity mode's coupling strength parameter a_I can take the general form:*

$$a_I(\langle T\rangle, \langle S\rangle) = \sum_{j=1}^{n} \alpha_j \frac{\int_{D_j^T} T dV / T_\Delta^*}{D_j^T} + \beta_j \frac{\int_{D_j^S} S dV / S_\Delta^*}{D_j^S}, \tag{20}$$

where $D_j^T, D_j^S \subset \Omega$; $\alpha_j, \beta_j \in \mathbb{R}$, $j = 1, ..., n$. One could further allow α_j, β_j to be spatially dependent, as long as they are bounded. The variable values of α_j and β_j can be used to better approximate the nonlinear equation of state [23].

- *Any number of density-driven velocity modes $m \in \mathbb{N}$, resulting in a composite velocity field of the form*

$$\boldsymbol{u}(\boldsymbol{r}, t; T, S) = Pe\boldsymbol{u}_E(\boldsymbol{r}, t) + Ra^T \sum_{k=1}^{m} a_I^k(\langle T\rangle, \langle S\rangle)\boldsymbol{u}_I^k(\boldsymbol{r}), \tag{21}$$

where a_I^k, $k = 1, ..., m$, are as in equation (20). Such a generalization may be used to examine the balance between several competing effects, possibly leading to nontrivial temporal competition between the different modes.

- *Different geometries, domain shapes, etc., as long as the velocity modes do not exit the boundaries, i.e. their normal component to the boundaries of the domain is 0. Also, the work can be generalized to spherical geometry by inserting the appropriate curvature parameters.*

- *Spatially dependent values of $\boldsymbol{\kappa}$, provided they are bounded from below by a positive constant.*

- *A different combination of boundary conditions - Robin-Robin or Neumann-Neumann.*

- *Non-autonomous systems - we expand on these in the following remarks.*

Remark 2 *Non-autonomous systems.*
In the case of a non-autonomous system, with the time-dependence arising from the external velocity field $\boldsymbol{u}_E$, the source terms f_T, f_S and/or the boundary forcing terms T^*, S^*, Theorem 1 is still valid using the same proof. The only modifications would be in the bounds, which would need to always take into account a global-in-time bound for each time-dependent quantity. Similarly, Corollary 1 is still valid, with the same modifications to the bounds.

Remark 3 *Periodically non-autonomous systems.*
In the case of non-autonomous sources with a period $\mathcal{T}$, it is possible to follow similar arguments as in Theorem 1 to show that a time-periodic set of solutions exists for any parameters. The structure of this proof would be to create a time-$\mathcal{T}$ Poincaré map of the original PDEs. Since the equations are parabolic, they have a smoothing effect that would allow one to show that the map's embedding is compact by the Rellich lemma. The Schauder-Tychonoff fixed point theorem would then guarantee the existence of a fixed point for the time-$\mathcal{T}$ map. Since the same steps can be followed for any starting point, this would prove the existence of a time-$\mathcal{T}$-periodic solution. Furthermore, if the forcing is weak with respect to the diffusion, it should be possible to show in this case that a time-periodic solution is stable in much the same way as we prove Theorem 2(iii) in this work.

Indeed, we see numerical evidence of remark 3. For example, in figure 3, a time-periodic forcing with a yearly period in the velocity field and surface forcing results in a stable time-periodic solution, with the same period.

VI. DISCUSSION AND CONCLUSIONS

We have presented a novel intermediate-complexity phenomenological dynamic model of the interaction between a 3D time-dependent kinematic flow with a density-driven component, and the density function in a closed basin. The model is an extended version of the advection-diffusion equations for temperature and salinity, with the velocity field serving as a non-local (integral) coupling term. Additionally, the model allows to include realistic sources for temperature and salinity. The coupling renders the equation set non-linear and non-local, thus the very existence and uniqueness of solutions is not a straightforward property of the model. We have proven here, using an iteration scheme and energy estimates, that the model is well-posed in the sense of Hadamard. We have shown that, for a small enough Rayleigh number, solutions will converge to a unique steady state. Though the analytical bound we obtain is not tight, this qualitative behavior does appear in numerical simulations of the system. Furthermore, the theorems hold for a larger class of velocities, coupling and boundary conditions, as described in remark 1. In addition, our

model bears resemblance to non-local continuum models used to model swarm dynamics and cell migration [34, 40]. Assuming the advection field is incompressible, proving well-posedness of such models should follow the same framework as presented here.

The presented model is an intermediate model between the fully coupled NS equations and the uncoupled advection-diffusion equation which is used in kinematic models. Thus, it may allow, in the future, to produce additional simplifications by which Lagrangian trajectories produced by chaotic advection can serve as the active force that changes the internal velocity field component.

The model is relevant for any incompressible flow with a density-driven component. Our motivating example has been the North Atlantic large scale flow, built from rapid externally driven modes and a slow density-driven overturning mode. Indeed, we have shown that, given realistic parameters, the model can produce a stable periodic state with an overturning strength, and temperature and salinity functions, that resemble realistic and more complex simulations of the North Atlantic ocean in current climate conditions (figure 3). The major discrepancy between our results and realistic distributions of the tracer functions is the density inversions we obtain in the northern latitudes. These density inversions are to be expected, as we do not have a mechanism that mixes inverted water columns in the model, such as convective adjustment [24]. In the actual North Atlantic ocean, density inversions appear only in the wintertime, as a result of intense cooling events, upon which convection ensues [38]. Introducing a mechanism for resolving the density inversions in our model will solve this discrepancy. There are a few options for such solutions, including convective-adjustment correction schemes [24], a local increase of the vertical diffusion coefficient, or a transient, localized kinematic flow that mixes the column through chaotic advection. We are currently exploring the different options; in later works we will present a model that includes density-inversion corrections.

Another realistic effect that is left for a future study, is the Mediterranean outflow water's (MOW) role in the variability of the AMOC. The Mediterranean, through the Strait of Gibraltar (36°N, 5.7°W), inserts approximately 1 Sv of warm, salty and dense water that sinks to a depth of 1000 meters, and spreads westwards, all the way to the eastern coasts of America. The various paths of these salty dense waters to the abyss and their dominance in driving the AMOC can be examined by our model. The effect of a modelled MOW on the asymptotic solutions is interesting in the oceanographic context, as it is still under debate to what extent the MOW influences the AMOC strength in current climate conditions [30, 36, 39, 46].

VII. ACKNOWLEDGEMENTS

OSK acknowledges the support of a research grant from the Yotam Project and the Weizmann Institute Sustainability and Energy Research Initiative; and the support of the Séphora Berrebi Scholarship in Mathematics. VRK and OSK acknowledge the support of the Israel Science Foundation, Grant 1208/16. EST is grateful to support in part by the Einstein Stiftung/Foundation - Berlin, through the Einstein Visiting Fellow Program. HG acknowledges the support of the Vigevani Research Project Prize.

Appendix A: Velocity field of the North Atlantic ocean - kinematic model

Here we present the rescaled version of the velocity fields $\boldsymbol{u}_{E,rs}$ and $\boldsymbol{u}_{I,rs}$ that we use in this work to mimic the North Atlantic ocean's large-scale flow, where $\boldsymbol{u}_{E,rs}$ describes the wind-driven surface flow, and $\boldsymbol{u}_{I,rs}$ describes the density-driven overturning flow. Thus, the domain we consider is $\Omega = [0,1] \times [0,1] \times [0,1] \in \mathbb{R}^3$, and the velocity fields' components are of the order 1. In the following, we neglect the rs notation for simplicity of notation.

The flow is incompressible, and is derived from a vector streamfunction. To model the wind-driven surface flow, we build a streamfunction component with a "hill" and "valley" along the latitudinal component $\hat{y}$, and an asymmetric westward-biased "hill" along the longitudinal component $\hat{x}$. The function we use along the $\hat{y}$ component is a simple combination of sine functions, $\psi_y^H = \sin \pi y \sin \pi (y - y_H(t))$, where $y_H(t)$ is the demarcation line between the gyres. Considering a yearly periodicity, we set $y_H(t) = \bar{y}_H + \tilde{y}_H \sin \pi t \tau / year$, where $t = 0$ corresponds to January, and $\tau = L_z^2/\kappa_V$ is the typical timescale of the rescaling (see subsection II A). The time dependency of the demarcation line causes chaotic mixing between the two gyres, as shown in [1, 60].

The function we use along the $\hat{x}$ component is an approximated solution to the Sverdrup equation with a narrow boundary layer in a wide basin [60]:

$$\psi_x^H = \frac{l_x}{m_x}(1 - e^{-x/l_x})(x - 1),$$

where m_x is a normalizing factor, $m_x = \max_x \frac{l_x}{m_x}(1 - e^{-x/l_x})(x - 1)$, so that the peak of the "hill" is at the value 1. l_x is the westward asymmetry parameter, and determines the width of the western boundary current and the Gulf Stream. At its limits, $l_x \to 0$ corresponds to a width of 0 and an infinite velocity at $x = 0$, while $l_x \to \infty$ corresponds to a symmetric function around $x = 0.5$.

This construction is similar to that used in [60], in which they considered a two-dimensional flow built from two asymmetric gyres of a similar functional form with an oscillating demarcation line between gyres. For a given time t, the Eulerian streamlines are closed contours with a westward bias. The time dependancy of ψ_y causes the Lagrangian trajectories to be chaotic, resulting in transport between the two gyres.

To these wind-driven features we add the influence of the penetration depth, set as the depth of the thermocline H_{TC}. Thus the z-dependence of the flow is given by an exponential decay of the horizontal velocity with depth, $\psi_z^H = 1-(1+\exp-(1-z)/H_{TC})^{-1}$. The overall streamfunction component ψ_H is the product of these contributions:

$$\psi_H(\boldsymbol{r},t) = \psi_x^H(x)\psi_y^H(y,t)\psi_z^H(z),$$

and the horizontal component of the velocity field is derived in the common method:

$$\boldsymbol{u}_E \equiv (u_E, v_E, w_E)\ ;\ \ u_E = \partial_y\psi_H,\ v_E = -\partial_x\psi_H,\ w_E = 0.$$

In the full, 3D velocity field, $\boldsymbol{u}_E$ is multiplied by the strength parameter Pe. Present day measurements evaluate the Gulf Stream width at around 150km, and its maximal velocity at around 2.5m/s. The average demarcation line position is around $40°N$, and its oscillation is seasonal with an amplitude of approximately $2°$ latitude. The thermocline depth is around 500 meters. In order to reproduce these numbers, the relevant parameters are set to be of the order $Pe \sim 10^6$,$\bar{y}_H = 0.57$, $\tilde{y}_H = 0.05$, $H_{TC} = 0.1$, $l_x = 0.01$.

To model the vertical overturning flow, we build a zonally symmetric streamfunction component with an asymmetric "hill" along the depth component $\hat{z}$, and an asymmetric "hill" along the latitudinal component $\hat{y}$. This formulation creates an overturning flow with north-bound transport above the thermocline, a steep downwelling regime along the northern border of the basin, south-bound flow below the thermocline and semi-uniform upwelling south of the downwelling regime. The asymmetric functions we use are of the same form as ψ_x^H:

$$\psi_z^V = \frac{l_z}{m_z}\left(1-e^{\frac{z-1}{l_z}}\right)z,\ \ \psi_y^V = \frac{l_y}{m_y}\left(1-e^{-\frac{y}{l_y}}\right)(1-y),$$

where m_z, m_y are rescaling parameters as in ψ_H. The overall streamfunction is given by

$$\psi_V(\boldsymbol{r}) = \psi_z^V(z)\psi_y^V(y),$$

and the velocity is derived in the common form:

$$\boldsymbol{u}_I \equiv (u_I, v_I, w_I)\ ;\ \ u_I = 0,\ v_I = \partial_z\psi_V,\ w_I = -\partial_y\psi_V\,.$$

For a downwelling regime between $55°N$ and $70°N$, the $\hat{y}$-direction asymmetry parameter is set to $l_y = 0.1$. For a northbound transport extending to a depth of 1000 meters, the $\hat{z}$-direction asymmetry parameter is set to $l_z = 0.1$. In this case, the maximal velocity is obtained at the downwelling branch. The full velocity field is given by a weighted sum of the individual contributions:

$$\boldsymbol{u}(\boldsymbol{r},t) = Pe\boldsymbol{u}_E(\boldsymbol{r},t) + Ra^T a_I(\langle\rho\rangle)\boldsymbol{u}_I(\boldsymbol{r}).$$

Appendix B: Model parameters

In table I, we summarize the model parameters used in the numerical experiments shown in figures 3 (realistic experiment), 4 (2D autonomous experiment) and 5 (3D autonomous experiment). To obtain the bi-stability seen in figures 4 and 5, the dimensions of the top box z_B and the effective rates of convective heat and mass transfer at the boundary, g_A^T and g_A^S respectively, had to be tweaked, as is common in box model experiments. The rest of the changes in parameters between the experiments are motivated in section IV. Overall, the physical control parameters we changed between the three experiments are the internal strength parameter a_I's proportionality constant Γ, the eddy diffusivity $\boldsymbol{\kappa}$, the temperature and salinity surface functions T^* and S^*, the strength of the external velocity field a_E, the intergyre mean demarcation line $\bar{y}_H$ and oscillation amplitude $\tilde{y}_H$, the depth box boundary z_B and the depth asymmetry parameter l_z. Correspondingly, the rescaling parameters $\tau = L_z^2/\kappa_V$, $T_\Delta^* = \max T^* - \min T^*$, and $S_\Delta^* = \max S^* - \min S^*$ change between the experiments, as do the non-dimensional parameters $\delta_x = \frac{\kappa_H/L_x^2}{\kappa_V/L_z^2}$, $\delta_y = \frac{\kappa_H/L_y^2}{\kappa_V/L_z^2}$, $Sh = \frac{g_A^S}{\kappa_V/L_z}$, $Pe = \frac{a_E}{\kappa_V/L_z}$, $Ra^T = \frac{\Gamma\rho_0\alpha T_\Delta^*}{\kappa_V/L_z}$, and $R_\rho = \frac{\alpha T_\Delta^*}{\beta S_\Delta^*}$.

Parameters [units]	**Parameter description**	**Realistic** figure 3	**2D aut.** figure 4	**3D aut.** figure 5
Rescaling				
L_x [m]	Domain longitudinal length ($0°W - 60°W$)	4×10^6 m	4×10^6 m	4×10^6 m
L_y [m]	Domain latitudinal length ($Eq - 70°N$)	7.7×10^6 m	7.7×10^6 m	7.7×10^6 m
L_z [m]	Domain depth (approximate AMOC depth)	4×10^3 m	4×10^3 m	4×10^3 m
τ [yrs]	Vertical eddy diffusivity timescale	2×10^3 yr	$10^3 - 10^6$ yr	100 yr
T^*_Δ [°C]	Temperature scale ($\max T^* - \min T^*$)	30°C	18°C	18°C
S^*_Δ [psu]	Salinity scale ($\max S^* - \min S^*$)	17 psu	0.1 – 5 psu	0.1 – 5 psu
Diffusion				
δ_x	Rescaled longitudinal diffusion	10	$10^2 - 10^5$	0.1
δ_y	Rescaled latitudinal diffusion	10	$10^2 - 10^5$	0.1
Sources				
Nu	Nusselt number (heat transfer at boundary)	140	140	140
Sh	Sherwood number (mass transfer at boundary)	20	$10 - 10^4$	1
f_T	Bulk temperature source	0	0	0
f_S	Bulk salinity source	0	0	0
External velocity	**(wind-driven)**			
Pe	Péclet number	1.6×10^6	0	5×10^4
$\bar{y}_H$	Mean intergyre demarcation line (see figure 1)	0.57	-	0.78
$\tilde{y}_H$	Intergyre oscillation amplitude	0.02	-	0
l_x	Longitudinal asymmetry parameter	0.01	-	0.01
H_{TC} [m]	Penetration depth	600 m	-	600 m
Internal velocity	**(density-driven)**			
Ra^T	Thermal Rayleigh number	1.8×10^3	$2 \times 10^2 - 2 \times 10^5$	20
R_ρ	Density stability ratio	5.8	1 – 50	1 – 50
y_B	Latitudinal box boundary (see figure 1)	0.78	0.78	0.78
z_B	Depth box boundary (see figure 1)	0.75	0.85	0.85
l_y	Latitudinal asymmetry parameter	0.1	0.1	0.1
l_z	Depth asymmetry parameter	0.1	0.04	0.04

TABLE I: Parameters used in the simulations presented in this work. Aut. = Autonomous.

Appendix C: Proof of global well-posedness of the model and corollaries

1. Problem formulation

Let $\Omega = [0,1]^3$ be a box domain describing a (rescaled) oceanic basin, $\boldsymbol{r} = (x, y, z) \in \Omega$ a general point in the domain, $q \in \partial\Omega$ a general point on the boundary of the domain, and define $\hat{\boldsymbol{n}}(q)$ as the unit vector pointing outwards from the boundary of the domain. We further define $\sigma_1 \equiv (x, y, z = 1) \subset \partial\Omega$ as the upper surface of the ocean, and $\sigma_0 \equiv (x, y, z = 0) \subset \partial\Omega$ as the ocean bottom. We divide the domain into two subdomains $D_1, D_2 \subset \Omega$, $D_1 \cup D_2 = \Omega$, and define for any function $\phi \in L^2(\Omega)$ the averages $\langle\phi\rangle_i = \frac{1}{D_i}\int_{D_i} \phi dV$, for $i = 1, 2$. As a shorthand, we define the vector $\langle\phi\rangle \equiv (\langle\phi\rangle_1, \langle\phi\rangle_2) \in \mathbb{R}^2$.

Let $\boldsymbol{\kappa}$, Γ, α, β, a_E, $\boldsymbol{u}_E$, $\boldsymbol{u}_I$, f_T, f_S, g_A^T, g_A^S, T^*, S^*, T_0, and S_0 be given, and satisfy the following assumptions:

$$
\begin{cases}
(a1)\ \ \boldsymbol{\kappa} = \begin{pmatrix} \kappa_x & 0 & 0 \\ 0 & \kappa_y & 0 \\ 0 & 0 & \kappa_z \end{pmatrix};\ \ \kappa_x, \kappa_y, \kappa_z \ \ \textit{positive constants}, \\
(a2)\ \ \Gamma, \alpha, \beta, a_E, g_A^T, g_A^S \ \ \textit{positive constants}, \\
(a3)\ \ \boldsymbol{u}_E, \boldsymbol{u}_I \in (L^\infty(\boldsymbol{r}))^3, \\
(a4)\ \ \boldsymbol{\nabla}\cdot\boldsymbol{u}_E(\boldsymbol{r}) = \boldsymbol{\nabla}\cdot\boldsymbol{u}_I(\boldsymbol{r}) = 0, \\
(a5)\ \ \boldsymbol{u}_E\cdot\hat{\boldsymbol{n}}(q) = \boldsymbol{u}_I\cdot\hat{\boldsymbol{n}}(q) = 0, \\
(a6)\ \ f_T, f_S \in L^2(\Omega); \int_\Omega f_S dV = 0, \\
(a7)\ \ T^*, S^* \in H^1(\sigma_1); \int_{\sigma_1} S^* dxdy = 0, \\
(a8)\ \ \partial_x T^*(x,y)|_{x=0,1} = \partial_y T^*(x,y)|_{y=0,1} = 0, \\
(a9)\ \ \partial_x S^*(x,y)|_{x=0,1} = \partial_y S^*(x,y)|_{y=0,1} = 0, \\
(a10)\ \ T_0, S_0 \in L^2(\Omega). \\
(a11)\ \ \int_\Omega S_0 dV = 0.
\end{cases}
\tag{C1}
$$

Then we consider the following nonlinear evolution problem for a temperature T and a salinity S:

$$
\begin{cases}
(P1)\ \ \partial_t T(t,\boldsymbol{r}) - \boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}T(t,\boldsymbol{r})) + (\boldsymbol{u}(\boldsymbol{r};\langle T\rangle,\langle S\rangle)\cdot\boldsymbol{\nabla})T(t,\boldsymbol{r}) = f_T(\boldsymbol{r}), & t>0\,,\ \ \boldsymbol{r}\in\Omega \\
(P2)\ \ \partial_t S(t,\boldsymbol{r}) - \boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}S(t,\boldsymbol{r})) + (\boldsymbol{u}(\boldsymbol{r};\langle T\rangle,\langle S\rangle)\cdot\boldsymbol{\nabla})S(t,\boldsymbol{r}) = f_S(\boldsymbol{r}), & t>0\,,\ \ \boldsymbol{r}\in\Omega \\
\\
(P3)\ \ a_I(\langle T\rangle,\langle S\rangle) = \Gamma\left(-\alpha(\langle T\rangle_2(t) - \langle T\rangle_1(t)) + \beta(\langle S\rangle_2(t) - \langle S\rangle_1(t))\right), & t>0 \\
(P4)\ \ \boldsymbol{u}(\boldsymbol{r};\langle T\rangle,\langle S\rangle) = a_E\boldsymbol{u}_E(\boldsymbol{r}) + a_I(\langle T\rangle,\langle S\rangle)\boldsymbol{u}_I(\boldsymbol{r}), & t>0\,,\ \ \boldsymbol{r}\in\Omega \\
\\
(P5)\ \ (\boldsymbol{\kappa}\boldsymbol{\nabla}T(t,q))\cdot\hat{\boldsymbol{n}}(q) = \begin{cases} g_A^T(T^*(x,y) - T(t,q)) & \text{if } q\in\sigma_1 \\ 0 & \text{else} \end{cases}, & t>0\,,\ \ q\in\partial\Omega \\
(P6)\ \ (\boldsymbol{\kappa}\boldsymbol{\nabla}S(t,q))\cdot\hat{\boldsymbol{n}}(q) = \begin{cases} g_A^S S^*(x,y) & \text{if } q\in\sigma_1 \\ 0 & \text{else} \end{cases}, & t>0\,,\ \ q\in\partial\Omega \\
\\
(P7)\ \ T(t=0,\boldsymbol{r}) = T_0(\boldsymbol{r}), & \boldsymbol{r}\in\Omega \\
(P8)\ \ S(t=0,\boldsymbol{r}) = S_0(\boldsymbol{r}). & \boldsymbol{r}\in\Omega
\end{cases}
\tag{C2}
$$

Note that assumptions $(a8)$ and $(a9)$ are compatibility conditions for the boundary. The coupled system $(P1)-(P8)$ equipped with assumptions (C1) is denoted by $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$. Let $\lambda = \frac{1}{4}\min\{g_A^T,\kappa_z\}$ and $\kappa_{min} = \min\{\kappa_x,\kappa_y,\kappa_z\}$; λ and κ_{min} are strictly positive from assumption (a2). Throughout the text, $\boldsymbol{\kappa}$, Γ, α, β, a_E, $\boldsymbol{u}_E$, $\boldsymbol{u}_I$ g_A^T, g_A^S, λ and κ_{min} are called the problem parameters; f_T, f_S are called the source functions; T^*, S^* are called the boundary functions, and T_0, S_0 are the initial conditions.

Remark 4 *Integrating equation* $(P2)$ *over the domain, using the relevant boundary conditions, one can show that* $\int_\Omega S(t,\boldsymbol{r})dV = Constant$.

Hence, without loss of generality and for simplicity we consider solutions with $\int_\Omega S(t,\boldsymbol{r})dV = 0$. Suppose $W \in L^1(\Omega)$, then we denote $\dot{W} = \{f\in W : \int_\Omega f dV = 0\}$.

Here we follow closely the presentation in Nittka [43], and correspondingly we define three notions of solutions to $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$. For $\mathcal{T} > 0$, we denote the Banach spaces: $W_T = C([0,\mathcal{T}];L^2(\Omega)) \cap L^2(0,\mathcal{T};H^1(\Omega))$, $W_S = C([0,\mathcal{T}];\dot{L}^2(\Omega)) \cap L^2(0,\mathcal{T};\dot{H}^1(\Omega))$.

Definition 1 *Let* $\mathcal{T} > 0$. $(T,S) \in W_T \times W_S$ *is a weak solution on* $[0,\mathcal{T}]$ *of* $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ *if, for all test functions*

$\psi \in H^1([0,\mathcal{T}];H^1(\Omega))$ *with* $\psi(\mathcal{T})=0$*, the following holds:*

$$\begin{aligned}
&-\int_0^{\mathcal{T}}\int_\Omega T(t)\partial_t\psi(t)dVdt+\int_0^{\mathcal{T}}\left(\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}T(t)-\boldsymbol{u}T(t))\cdot\boldsymbol{\nabla}\psi(t)dV+\int_{\sigma_1}g_A^T T(t)\psi(t)dxdy\right)dt\\
&=\int_\Omega T_0\psi(0)dV+\int_0^{\mathcal{T}}\int_\Omega f_T\psi(t)dVdt+\int_0^{\mathcal{T}}\int_{\sigma_1}g_A^T T^*(x,y)\psi(t)dxdydt;
\end{aligned}\tag{C3}$$

$$\begin{aligned}
&-\int_0^{\mathcal{T}}\int_\Omega S(t)\partial_t\psi(t)dVdt+\int_0^{\mathcal{T}}\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}S(t)-\boldsymbol{u}S(t))\cdot\boldsymbol{\nabla}\psi(t)dVdt\\
&=\int_\Omega S_0\psi(0)dV+\int_0^{\mathcal{T}}\int_\Omega f_S\psi(t)dVdt+\int_0^{\mathcal{T}}\int_{\sigma_1}g_A^S S^*(x,y)\psi(t)dxdydt,
\end{aligned}$$

where $\boldsymbol{u}=\boldsymbol{u}(\boldsymbol{r};\langle T\rangle,\langle S\rangle)$ *is given by equation* $(P4)$ *in* (C2)*.*

Remark 5 *If a weak solution exists, then from* (C3) *one can prove that* $(T(t=0),S(t=0))=(T_0,S_0)$*.*

Definition 2 *Let* $\mathcal{T}>0$*.* (T,S) *is a strong solution on* $[0,\mathcal{T}]$ *of* $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ *if it is a weak solution, and* $(T,S)\in H^1([0,\mathcal{T}];L^2(\Omega))\times H^1([0,\mathcal{T}];\dot{L}^2(\Omega))$*.*

Definition 3 *Let* $\mathcal{T}>0$*.* (T,S) *is a classical* L^2*-solution on* $[0,\mathcal{T}]$ *of* $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ *if:*

i. $(T,S)\in(C^1([0,\mathcal{T}];L^2(\Omega))\cap C([0,\mathcal{T}];H^1(\Omega)))\ \times\ (C^1([0,\mathcal{T}];\dot{L}^2(\Omega))\cap C([0,\mathcal{T}];\dot{H}^1(\Omega)))$

ii. $\boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}T),\boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}S)\in C([0,\mathcal{T}];L^2(\Omega))$

iii. $(\boldsymbol{\kappa}\boldsymbol{\nabla}T)\cdot\hat{\boldsymbol{n}}|_{\partial\Omega},(\boldsymbol{\kappa}\boldsymbol{\nabla}S)\cdot\hat{\boldsymbol{n}}|_{\partial\Omega}\in C([0,\mathcal{T}];L^2(\partial\Omega))$

iv. $(T(t=0),S(t=0))=(T_0,S_0)$

v. (T,S) *satisfies equations* $(P1)$*,* $(P2)$ *in the sense of* $C([0,\mathcal{T}];L^2(\Omega))$ *and equations* $(P5)$*,* $(P6)$ *in the sense of* $C([0,\mathcal{T}];L^2(\partial\Omega))$*.*

As we shall see in Theorem 1, for classical L^2-solutions the initial conditions will need to satisfy, in addition to the relevant assumptions in (C1), the following compatibility assumptions:

$$\begin{cases}
(b1)\ \ T_0,S_0\in H^4(\Omega)\\
(b2)\ \ (\boldsymbol{\kappa}\boldsymbol{\nabla}T_0(q))\cdot\hat{\boldsymbol{n}}(q)=\begin{cases}g_A^T(T^*(x,y)-T_0(q)) & \text{if } q\in\sigma_1\\ 0 & \text{else}\end{cases}\\
(b3)\ \ (\boldsymbol{\kappa}\boldsymbol{\nabla}S_0(q))\cdot\hat{\boldsymbol{n}}(q)=\begin{cases}g_A^S S^*(x,y) & \text{if } q\in\sigma_1\\ 0 & \text{else}\end{cases}
\end{cases}.\tag{C4}$$

We further define the notion of temporally global solutions as follows:

Definition 4 (T,S) *is a global weak/strong/classical solution of* $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ *if, for any* $\mathcal{T}>0$*, it is a weak/strong/classical solution on* $[0,\mathcal{T}]$*, respectively.*

Theorem 1 *Let* $\boldsymbol{\kappa}$*,* Γ*,* α*,* β*,* a_E*,* $\boldsymbol{u}_E$*,* $\boldsymbol{u}_I$*,* f_T*,* f_S*,* g_A^T*,* g_A^S*,* T^**,* S^**,* T_0*, and* S_0 *be given as in* (C1)*. Then:*

(i.) The nonlinear problem $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ *has a unique, global strong solution* (T,S)*.*

(ii.) This solution depends continuously on initial conditions, boundary conditions and sources in the sense described by equation (C26)*, below.*

(iii.) For every $\bar{\mathcal{T}}>0$*, the solution satisfies the following bounds:*

$$\begin{aligned}
&\sup_{t\in[0,\bar{\mathcal{T}}]}\|T\|^2(t)\le C_1^T;\quad\int_0^{\bar{\mathcal{T}}}\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}T\right\|^2(t)dt\le C_2^T+\bar{\mathcal{T}}C_3^T;\\
&\sup_{t\in[0,\bar{\mathcal{T}}]}\|S\|^2(t)\le C_1^S;\quad\int_0^{\bar{\mathcal{T}}}\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S\right\|^2(t)dt\le C_2^S+\bar{\mathcal{T}}C_3^S;
\end{aligned}\tag{C5}$$

where

$$
\begin{aligned}
C_1^T &\equiv \|T_0\|^2 + \frac{g_A^T}{2\lambda(1-a)}\tilde{\|}T^*\tilde{\|}^2 + \frac{1}{4\lambda^2 a(1-a)}\|f_T\|^2\,;\\
C_2^T &\equiv \frac{1}{2}\|T_0\|^2\,;\quad C_3^T \equiv \frac{\lambda}{2}\|T_0\|^2 + g_A^T\tilde{\|}T^*\tilde{\|}^2 + \frac{1}{\lambda}\|f_T\|^2\,;\\
C_1^S &\equiv \|S_0\|^2 + \frac{1}{2\pi\kappa_{min}^2}\left(\frac{1+\frac{2}{\pi}}{b}\left(g_A^S\right)^2\tilde{\|}S^*\tilde{\|}^2 + \frac{1}{\pi(\frac{1}{2}-b)}\|f_S\|^2\right);\\
C_2^S &\equiv \frac{\pi\kappa_{min}}{2}\|S_0\|^2\,;\quad C_3^S \equiv \frac{\pi\kappa_{min}}{2}\|S_0\|^2 + 6\frac{g_A^S}{\kappa_{min}}\left(g_A^S\tilde{\|}S^*\tilde{\|}^2 + \frac{2}{\pi\kappa_{min}}\|f_S\|^2\right),
\end{aligned}
\tag{C6}
$$

for $a = \left(1+\sqrt{1+2g_A^T\frac{\tilde{\|}T^*\tilde{\|}^2}{\|f_T\|^2}}\right)^{-1}$ *and* $b = \frac{1}{2}\left(1+\frac{1}{g_A^S\sqrt{2+\pi}}\frac{\|f_S\|}{\tilde{\|}S^*\tilde{\|}}\right)^{-1}$.

(iv.) If, in addition, the initial conditions satisfy (C4) *then* $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ *has a unique, global classical* L^2*-solution.*

2. Proof of Theorem 1

a. *Construction of an iterative sequence of approximate solutions*

The proof will consist of several major steps, which we will describe in the next subsections.

Let us start by constructing an iterative scheme, that will result in a sequence of solutions to an iteratively defined linear problem that approximates the nonlinear problem $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$. To this end, let $\mathcal{T} > 0$, and let $T_m, S_m \in C([0,\mathcal{T}];L^2(\Omega))$ be given. Then we present the following linear problem, defined for $t \in (0,\mathcal{T}]$:

$$
\begin{cases}
(P1^l)\ \ \partial_t T(t,\boldsymbol{r}) - \boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}T(t,\boldsymbol{r})) + (\boldsymbol{u}_m(\boldsymbol{r},\langle T_m\rangle,\langle S_m\rangle)\cdot\boldsymbol{\nabla})T(t,\boldsymbol{r}) = f_T(\boldsymbol{r}), & t>0\,,\ \ \boldsymbol{r}\in\Omega\\
(P2^l)\ \ \partial_t S(t,\boldsymbol{r}) - \boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}S(t,\boldsymbol{r})) + (\boldsymbol{u}_m(\boldsymbol{r};\langle T_m\rangle,\langle S_m\rangle)\cdot\boldsymbol{\nabla})S(t,\boldsymbol{r}) = f_S(\boldsymbol{r}), & t>0\,,\ \ \boldsymbol{r}\in\Omega\\
\\
(P3^l)\ \ a_I(\langle T_m\rangle,\langle S_m\rangle) = \Gamma\left(-\alpha(\langle T_m\rangle_2 - \langle T_m\rangle_1) + \beta(\langle S_m\rangle_2 - \langle S_m\rangle_1)\right), & t>0\\
(P4^l)\ \ \boldsymbol{u}_m(\boldsymbol{r};\langle T_m\rangle,\langle S_m\rangle) = a_E\boldsymbol{u}_E(\boldsymbol{r}) + a_I(\langle T_m\rangle,\langle S_m\rangle)\boldsymbol{u}_I(\boldsymbol{r}), & t>0\,,\ \ \boldsymbol{r}\in\Omega\\
\\
(P5^l)\ \ (\boldsymbol{\kappa}\boldsymbol{\nabla}T(t,q))\cdot\hat{\boldsymbol{n}}(q) = \begin{cases} g_A^T(T^*(x,y) - T(t,q)) & \text{if } q\in\sigma_1\\ 0 & \text{else}\end{cases}, & t>0\,,\ \ q\in\partial\Omega\\
(P6^l)\ \ (\boldsymbol{\kappa}\boldsymbol{\nabla}S(t,q))\cdot\hat{\boldsymbol{n}}(q) = \begin{cases} g_A^S S^*(x,y) & \text{if } q\in\sigma_1\\ 0 & \text{else}\end{cases}, & t>0\,,\ \ q\in\partial\Omega\\
\\
(P7^l)\ \ T(t=0,\boldsymbol{r}) = T_0(\boldsymbol{r}), & \boldsymbol{r}\in\Omega\\
(P8^l)\ \ S(t=0,\boldsymbol{r}) = S_0(\boldsymbol{r}). & \boldsymbol{r}\in\Omega
\end{cases}
\tag{C7}
$$

We call equations $(P1^l)$-$(P8^l)$ equipped with assumptions (C1) as $\left(P^{linear}_{T_0,S_0,T_m,S_m}\right)$. Note that in this notation we explicitly specify only the initial conditions T_0, S_0 and the functions $T_m, S_m \in C([0,\mathcal{T}];L^2(\Omega))$; the problem parameters, boundary functions, and source functions are the same as are given above in the theorem formulation. We remark that this problem is almost identical to the nonlinear problem, except that the velocity field is predetermined by the given functions T_m, S_m. Hence, this problem is indeed linear. In fact, one can decouple $\left(P^{linear}_{T_0,S_0,T_m,S_m}\right)$ into two separate linear problems: equations $(P1^l), (P3^l), (P5^l)$ for $T(t,\boldsymbol{r})$, and equations $(P2^l), (P4^l), (P6^l)$ for $S(t,\boldsymbol{r})$.

Following exactly as before, we define three notions of solutions to $\left(P^{linear}_{T_0,S_0,T_m,S_m}\right)$:

Definition 5 *Let* $\mathcal{T} > 0$ *and* $T_m, S_m \in C([0,\mathcal{T}];L^2(\Omega))$ *be given.* $(T,S) \in W_T \times W_S$ *is called a weak solution on* $[0,\mathcal{T}]$ *of the*

linear problem $\left(P^{linear}_{T_0,S_0,T_m,S_m}\right)$ *if, for all test functions* $\psi \in H^1([0,\mathcal{T}];H^1(\Omega))$ *with* $\psi(\mathcal{T})=0$, *the following holds:*

$$\begin{aligned}&-\int_0^{\mathcal{T}}\int_\Omega T(t)\partial_t\psi(t)dVdt+\int_0^{\mathcal{T}}\left(\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}T(t)-\boldsymbol{u}_m\,T(t))\cdot\boldsymbol{\nabla}\psi(t)dV+\int_{\sigma_1}g_A^T T(t)\psi(t)dxdy\right)dt\\&=\int_\Omega T_0\psi(0)dV+\int_0^{\mathcal{T}}\int_\Omega f_T\psi(t)dVdt+\int_0^{\mathcal{T}}\int_{\sigma_1}g_A^T T^*(x,y)\psi(t)dxdydt;\end{aligned}\tag{C8}$$

$$\begin{aligned}&-\int_0^{\mathcal{T}}\int_\Omega S(t)\partial_t\psi(t)dVdt+\int_0^{\mathcal{T}}\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}S(t)-\boldsymbol{u}_m\,S(t))\cdot\boldsymbol{\nabla}\psi(t)dVdt\\&=\int_\Omega S_0\psi(0)dV+\int_0^{\mathcal{T}}\int_\Omega f_S\psi(t)dVdt+\int_0^{\mathcal{T}}\int_{\sigma_1}g_A^S S^*(x,y)\psi(t)dxdydt.\end{aligned}\tag{C9}$$

Definition 6 *Let* $\mathcal{T}>0$, $T_m,S_m\in C([0,\mathcal{T}];L^2(\Omega))$ *be given.* (T,S) *is a strong solution on* $[0,\mathcal{T}]$ *of the linear problem* $\left(P^{linear}_{T_0,S_0,T_m,S_m}\right)$ *if it is a weak solution on* $[0,\mathcal{T}]$, *and*

$$T,S\in H^1([0,\mathcal{T}];L^2(\Omega)).$$

Definition 7 *Let* $\mathcal{T}>0$, $T_m,S_m\in C([0,\mathcal{T}];L^2(\Omega))$ *be given.* (T,S) *is a classical* L^2*-solution on* $[0,\mathcal{T}]$ *of the linear problem* $\left(P^{linear}_{T_0,S_0,T_m,S_m}\right)$ *if:*

1. $(T,S)\in C^1([0,\mathcal{T}];L^2(\Omega))\cap C([0,\mathcal{T}];H^1(\Omega))\times C^1([0,\mathcal{T}];\dot{L}^2(\Omega))\cap C([0,\mathcal{T}];\dot{H}^1(\Omega))$
2. $\{\boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}T),\boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}S)\}\in(C((0,\mathcal{T}];L^2(\Omega)))^2$
3. $\{(\boldsymbol{\kappa}\boldsymbol{\nabla}T)\cdot\hat{\boldsymbol{n}}|_{\partial\Omega},(\boldsymbol{\kappa}\boldsymbol{\nabla}S)\cdot\hat{\boldsymbol{n}}|_{\partial\Omega}\}\in(C([0,\mathcal{T}];L^2(\partial\Omega)))^2$
4. $(T(t=0),S(t=0))=(T_0,S_0)$
5. T,S *satisfy equations* $(P1^l,P2^l,P5^l,P6^l)$ *as functions in* $C([0,\mathcal{T}];L^2(\Omega))$, $C([0,\mathcal{T}];L^2(\partial\Omega))$. *In particular, thanks to (iii) above,* $(\partial_tT,\partial_tS)\in\left(C((0,\mathcal{T}];L^2(\Omega))\right)^2$

In particular, a classical L^2-solution is a strong solution.

Proposition 1 *Let* $\mathcal{T}>0$, $T_m,S_m\in C([0,\mathcal{T}];L^2(\Omega))$ *be given. Then:*

1. $\left(P^{linear}_{T_0,S_0,T_m,S_m}\right)$ *has a unique strong solution on* $[0,\mathcal{T}]$.
2. *If, in addition, the initial conditions* T_0,S_0 *satisfy assumptions (C4), then* $\left(P^{linear}_{T_0,S_0,T_m,S_m}\right)$ *has a unique classical* L^2*-solution on* $[0,\mathcal{T}]$.

Proof:

1. Existence and uniqueness of weak solutions of the linear problem are guaranteed directly by Theorem 2.11 in [43]. Moreover, existence and uniqueness of strong solutions are given by Remark 2.15 in [43]. We note that in order to use Remark 2.15 to prove the existence of a strong solution, one may construct a function $G\in H^1([0,\mathcal{T}];H^2(\Omega))$ whose restriction to the boundary satisfies the boundary conditions. In our problem setting, due to the compatibility conditions $(a8)$, $(a9)$ in (C1) this task can be easily achieved by setting $G(x,y,z,t)=T^*(x,y,t)$ for T, and $G(x,y,z,t)=g_A^S S^*(x,y,t)z^2/2$ for S, where T^* and S^* are lifted from the boundary to the full domain and are constant with respect to the z variable. Due to the trace theorem, G has the required regularity, i.e. $G\in H^1([0,\mathcal{T}];H^2(\Omega))$.
2. Since the boundary condition functions T^*, S^*, and the source functions f_T, f_S, are time-independent, this follows directly from Proposition 2.7(b) in [43].

□

Equipped with the above, we can now construct a sequence of strong solutions to an iteratively defined linear problem.

1. We initialize the sequence with (T_0, S_0). Note that these are constant functions in time, and therefore by Proposition 1 the problem $\left(P^{linear}_{T_0,S_0,T_0,S_0}\right)$ has a unique strong solution, that we denote (T_1, S_1). Then $T_1, S_1 \in C([0,\mathcal{T}]; L^2(\Omega))$ from the definition of a strong solution to the linear problem, and thus, by Proposition 1 , $\left(P^{linear}_{T_0,S_0,T_1,S_1}\right)$ has a strong solution, that we denote (T_2, S_2).

2. Let $n > 1$, and assume (T_n, S_n) is a strong solution of $\left(P^{linear}_{T_0,S_0,T_{n-1},S_{n-1}}\right)$ for given $T_{n-1}, S_{n-1} \in C([0,\mathcal{T}]; L^2(\Omega))$. Then $T_n, S_n \in C([0,\mathcal{T}]; L^2(\Omega))$ from the definition of a strong solution to the linear problem, and thus, by Proposition 1, $\left(P^{linear}_{T_0,S_0,T_n,S_n}\right)$ has a strong solution, that we denote (T_{n+1}, S_{n+1}).

By repeating Step (ii) iteratively we construct a sequence $\{(T_n, S_n)\}_{n=1}^{\infty}$ of strong solutions to the corresponding sequence of linear problems $\left\{\left(P^{linear}_{T_0,S_0,T_{n-1},S_{n-1}}\right)\right\}_{n=1}^{\infty}$ for $n \geq 1$. Under the additional assumptions (C4), by the same induction steps, the sequence $\{(T_n, S_n)\}_{n=1}^{\infty}$ is a sequence of classical L^2-solutions to $\left\{\left(P^{linear}_{T_0,S_0,T_{n-1},S_{n-1}}\right)\right\}_{n=1}^{\infty}$.

b. Global uniform bounds on the sequence of approximate solutions

Now that we have established a sequence of approximate solutions $\{(T_n, S_n)\}_{n=1}^{\infty}$, we will use energy estimates to establish some uniform in n estimates for $\{(T_n, S_n)\}_{n=1}^{\infty}$. To this end, let $\mathcal{T} > 0$, and $n \geq 1$.

Lemma 1 *Let $\phi \in H^1([0,\mathcal{T}]; L^2(\Omega)) \cap L^2(0,\mathcal{T}; H^1(\Omega))$, $\phi(\mathcal{T}) = 0$. A weak solution (T_n, S_n) of $\left(P^{linear}_{T_0,S_0,T_{n-1},S_{n-1}}\right)$ satisfies equations (C8), (C9) with ϕ in place of the test function $\psi \in H^1([0,\mathcal{T}]; H^1(\Omega))$.*

Proof: Since $H^1(\Omega)$ is dense in $L^2(\Omega)$, there exists a sequence $\{\phi^m\}_{m=1}^{\infty}$, $\phi^m \in H^1([0,\mathcal{T}]; H^1(\Omega))$ with $\phi^m(\mathcal{T}) = 0$, converging to ϕ in $H^1([0,\mathcal{T}]; L^2(\Omega)) \cap L^2(0,\mathcal{T}; H^1(\Omega))$. Thus, it is easy to see that equation (C8) with $\psi = \phi^m$ converges, term by term, to the same equation with ϕ in place of ψ as $m \to \infty$. □

Lemma 2 *Let $\psi \in H^1([0,\mathcal{T}]; L^2(\Omega)) \cap L^2(0,\mathcal{T}; H^1(\Omega))$ with $\psi(\mathcal{T}) \in \mathbb{R}$. A weak solution (T_n, S_n) of $\left(P^{linear}_{T_0,S_0,T_{n-1},S_{n-1}}\right)$ satisfies:*

$$\begin{aligned}&\int_\Omega T_n(\mathcal{T})\psi(\mathcal{T})dV - \int_0^{\mathcal{T}}\int_\Omega T_n(t)\partial_t\psi(t)dVdt + \int_0^{\mathcal{T}}\left(\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}T_n(t) - \boldsymbol{u}_{n-1}T_n(t))\cdot\boldsymbol{\nabla}\psi(t)dV + \right.\\ &\left.\int_{\sigma_1} g_A^T T_n(t)\psi(t)dxdy\right)dt = \int_\Omega T_0\psi(0)dV + \int_0^{\mathcal{T}}\left(\int_\Omega f_T\psi(t)dV + \int_{\sigma_1} g_A^T T^*(x,y)\psi(t)dxdy\right)dt.\end{aligned} \tag{C10}$$

$$\begin{aligned}&\int_\Omega S_n(\mathcal{T})\psi(\mathcal{T})dV - \int_0^{\mathcal{T}}\int_\Omega S_n(t)\partial_t\psi(t)dVdt + \int_0^{\mathcal{T}}\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}S_n(t) - \boldsymbol{u}_{n-1}S_n(t))\cdot\boldsymbol{\nabla}\psi(t)dVdt\\ &= \int_\Omega S_0\psi(0)dV + \int_0^{\mathcal{T}}\left(\int_\Omega f_S\psi(t)dV + \int_{\sigma_1} g_A^S S^*(x,y)\psi(t)dxdy\right)dt.\end{aligned} \tag{C11}$$

Proof: Set $\tilde{\mathcal{T}} = \mathcal{T}+1$. Let $(\tilde{T}_n, \tilde{S}_n)$ be a strong solution of $\left(P^{linear}_{T_0,S_0,T_{n-1},S_{n-1}}\right)$ on $[0,\tilde{\mathcal{T}}]$. Define $\phi(t) = \begin{cases}\psi(t) & t \in [0,\mathcal{T}]\\ \psi(\mathcal{T})\varphi(t) & t \in (\mathcal{T},\tilde{\mathcal{T}}]\end{cases}$,

where $\varphi(\mathcal{T}) = 1$, and φ is a smooth function decreasing to 0 over the interval $[\mathcal{T},\tilde{\mathcal{T}}]$. Then ϕ satisfies the conditions for a test function in the weak solution according to Lemma 1, i.e. $(\tilde{T}_n, \tilde{S}_n)$ satisfy equations (C8), (C9) with ϕ instead of ψ. Also, $(\tilde{T}_n, \tilde{S}_n)$ agrees with (T_n, S_n) on $[0,\mathcal{T}]$ from uniqueness of the strong solution.

Set $\mathcal{T}$ as the new initial time, and let $\{\tilde{\tilde{T}}_n, \tilde{\tilde{S}}_n\}$ be the solution of $\left(P^{linear}_{\tilde{T}_n(\mathcal{T}),\tilde{S}_n(\mathcal{T}),T_{n-1},S_{n-1}}\right)$. If the sources and boundary functions are not autonomous, shift their time parameter corresppondingly. Then $(\tilde{\tilde{T}}_n, \tilde{\tilde{S}}_n)$ agrees with $(\tilde{T}_n, \tilde{S}_n)$ on $[\mathcal{T},\tilde{\mathcal{T}}]$ from uniqueness of the strong solution, and therefore the following is satisfied:

$$\begin{aligned}&-\int_{\mathcal{T}}^{\tilde{\mathcal{T}}}\int_\Omega \tilde{T}_n(t)\partial_t\phi(t)dVdt + \int_{\mathcal{T}}^{\tilde{\mathcal{T}}}\left(\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}\tilde{T}_n(t) - \boldsymbol{u}_{n-1}\tilde{T}_n(t))\cdot\boldsymbol{\nabla}\phi(t)dV + \int_{\sigma_1} g_A^T\tilde{T}_n(t)\phi(t)dxdy\right)dt\\ &= \int_\Omega \tilde{T}(\mathcal{T})\phi(\mathcal{T})dV + \int_{\mathcal{T}}^{\tilde{\mathcal{T}}}\left(\int_\Omega f_T\phi(t)dV + \int_{\sigma_1} g_A^T T^*(t,x,y)\phi(t)dxdy\right)dt.\end{aligned} \tag{C12}$$

On the other hand, $\tilde{T}_n$ satisfies

$$\begin{aligned}
&-\int_0^{\tilde{\mathcal{T}}}\int_\Omega \tilde{T}_n(t)\partial_t\phi(t)dVdt+\int_0^{\tilde{\mathcal{T}}}\left(\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}\tilde{T}_n(t)-\boldsymbol{u}_{n-1}\tilde{T}_n(t))\cdot\boldsymbol{\nabla}\phi(t)dV+\int_{\sigma_1}g_A^T\tilde{T}_n(t)\phi(t)dxdy\right)dt\\
&=\int_\Omega T_0\phi(0)dV+\int_0^{\tilde{\mathcal{T}}}\left(\int_\Omega f_T\phi(t)dV+\int_{\sigma_1}g_A^T T^*(t,x,y)\phi(t)dxdy\right)dt.
\end{aligned}\tag{C13}$$

Subtracting (C12) from (C13), and separating the time integral to two parts, $\int_0^{\tilde{\mathcal{T}}}=\int_0^{\mathcal{T}}+\int_{\mathcal{T}}^{\tilde{\mathcal{T}}}$, we are left with:

$$\begin{aligned}
&\int_\Omega\tilde{T}_n(\mathcal{T})\phi(\mathcal{T})dV-\int_0^{\mathcal{T}}\int_\Omega\tilde{T}_n(t)\partial_t\phi(t)dVdt+\int_0^{\mathcal{T}}\left(\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}\tilde{T}_n(t)-\boldsymbol{u}_{n-1}\tilde{T}_n(t))\cdot\boldsymbol{\nabla}\phi(t)dV\right.\\
&\left.+\int_{\sigma_1}g_A^T\tilde{T}_n(t)\phi(t)dxdy\right)dt=\int_\Omega T_0\phi(0)dV+\int_0^{\mathcal{T}}\left(\int_\Omega f_T\phi(t)dV+\int_{\sigma_1}g_A^T T^*(t,x,y)\phi(t)dxdy\right)dt.
\end{aligned}$$

Now, since $\tilde{T}_n(t)=T_n(t)$ on $[0,\mathcal{T}]$ and $\phi(t)=\psi(t)$ on $[0,\mathcal{T}]$, this equals

$$\begin{aligned}
&\int_\Omega T_n(\mathcal{T})\psi(\mathcal{T})dV-\int_0^{\mathcal{T}}\int_\Omega T_n(t)\partial_t\psi(t)dVdt+\int_0^{\mathcal{T}}\left(\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}T_n(t)-\boldsymbol{u}_{n-1}T_n(t))\cdot\boldsymbol{\nabla}\psi(t)dV\right.\\
&\left.+\int_{\sigma_1}g_A^T T_n(t)\psi(t)dxdy\right)dt=\int_\Omega T_0\psi(0)dV+\int_0^{\mathcal{T}}\left(\int_\Omega f_T\psi(t)dV+\int_{\sigma_1}g_A^T T^*(t,x,y)\psi(t)dxdy\right)dt.
\end{aligned}$$

The same steps may be followed for S_n. □

Remark 6 *Following similar arguments as in Lemma 1 and Lemma 2 above, one can show that a weak solution (T,S) of the nonlinear problem $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ satisfies equations analogous to equations (C10) and (C11), with T in place of T_n, S in place of S_n, and, in place of $\boldsymbol{u}_{n-1}$, $\boldsymbol{u}=\boldsymbol{u}(\boldsymbol{r};\langle T\rangle,\langle S\rangle)$ as in equation $(P4)$ in (C2).*

Next we present a few additional lemmas that will be useful for proving our global bounds.

Lemma 3 *(Poincaré inequality) For a function $T\in H^1(\Omega)$, the following inequality holds:*

$$\frac{g_A^T}{2}\tilde{\|}T\tilde{\|}^2+\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}T\right\|^2\geq\lambda\|T\|^2.$$

Lemma 4 *Let $S\in H^1(\Omega)$. For any $\epsilon>0$,*

$$\tilde{\|}S\tilde{\|}^2\leq\left(1+\frac{1}{\epsilon}\right)\|S\|^2+\frac{\epsilon}{\kappa_{min}}\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S\right\|^2.$$

Remark 7 *Lemma 4 is a version of the trace theorem customized for this problem.*

Lemma 5 *Let $S\in H^1(\Omega)$. Then*

$$\|S-\langle S\rangle_\Omega\|^2\leq\frac{1}{\pi\kappa_{min}}\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S\right\|^2\quad\text{where }\langle S\rangle_\Omega\equiv\frac{1}{|\Omega|}\int SdV.$$

Proofs of these useful Lemmas 3-5 are left to the last section C 4. Now, we can prove several global bounds on the strong solutions (T_n,S_n) of $\left(P^{linear}_{T_0,S_0,T_{n-1},S_{n-1}}\right)$, that we have established in section C 2 a.

Proposition 2 *Suppose (T_n,S_n) is a strong solution on $[0,\mathcal{T}]$ of $\left(P^{linear}_{T_0,S_0,T_{n-1},S_{n-1}}\right)$, as established in section C 2 a. Then T_n, S_n satisfy the following bounds, that are independent of n:*

1. $\sup_{t\in[0,\mathcal{T}]}\|T_n\|^2(t)\leq C_1^T$;
2. $\int_0^{\mathcal{T}}\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}T_n\right\|^2(t)dt\leq C_2^T+\mathcal{T}C_3^T$;
3. $\sup_{t\in[0,\mathcal{T}]}\|S_n\|^2(t)\leq C_1^S$;

4. $\int_0^{\mathcal{T}} \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla} S_n\right\|^2 (t)dt \leq C_2^S + \mathcal{T} C_3^S$;

where C_i^j *for* $i \in \{1,2,3\}$, $j \in \{T,S\}$ *are given explicitly in equation* (C6).

Proof:
(i) Uniform in n, $L^\infty([0,\mathcal{T}];L^2(\Omega))$ bounds on T_n:
According to Lemma 2, T_n can be used as a test function in equation (C10) in place of ψ, and thus satisfies:

$$\int_\Omega (T_n(\mathcal{T}))^2 dV - \int_0^{\mathcal{T}} \int_\Omega T_n(t)\partial_t T_n(t) dV dt + \int_0^{\mathcal{T}} \left(\int_\Omega (\boldsymbol{\kappa}\boldsymbol{\nabla} T_n(t) + \boldsymbol{u}_{n-1} T_n(t)) \cdot \boldsymbol{\nabla} T_n(t) dV + \int_{\sigma_1} g_A^T (T_n(t))^2 dxdy \right) dt = \int_\Omega T_0^2 dV + \int_0^{\mathcal{T}} \left(\int_\Omega f_T T_n(t) dV + \int_{\sigma_1} g_A^T T^*(t,x,y) T_n(t) dxdy \right) dt.$$

Therefore, since $T_n \in L^2(0,\mathcal{T};H^1(\Omega)) \cap C([0,\mathcal{T}];L^2(\Omega))$, one can employ assumptions $(a4)$, $(a5)$ along with a generalized version of the divergence theorem [12] to obtain:

$$\int_0^{\mathcal{T}} \left(\frac{1}{2}\frac{d}{dt} \|T_n\|^2 + g_A^T |\!|\!| T_n |\!|\!|^2 + \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla} T_n\right\|^2 \right) dt = \int_0^{\mathcal{T}} \left(g_A^T \int_{\sigma_1} T_n T^*(x,y) d\sigma + \int_\Omega T_n f_T \right) dt.$$

Thus, the integrands of time are equal almost everywhere. The next steps will be performed inside the integrands, and eventually will be integrated over time again. We use the Cauchy-Schwarz and Young inequalities to obtain:

$$\begin{aligned} \frac{1}{2}\frac{d}{dt} \|T_n\|^2 + g_A^T |\!|\!| T_n |\!|\!|^2 + \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla} T_n\right\|^2 &\leq g_A^T |\!|\!| T_n |\!|\!| |\!|\!| T^* |\!|\!| + \|2\epsilon T_n\| \left\| \frac{1}{2\epsilon} f_T \right\| \\ &\leq \frac{g_A^T}{2} |\!|\!| T_n |\!|\!|^2 + \frac{g_A^T}{2} |\!|\!| T^* |\!|\!|^2 + \epsilon \|T_n\|^2 + \frac{1}{4\epsilon} \|f_T\|^2 \end{aligned}$$

and after reordering:

$$\frac{1}{2}\frac{d}{dt} \|T_n\|^2 + \frac{g_A^T}{2} |\!|\!| T_n |\!|\!|^2 + \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla} T_n\right\|^2 \leq \epsilon \|T_n\|^2 + \frac{g_A^T}{2} |\!|\!| T^* |\!|\!|^2 + \frac{1}{4\epsilon} \|f_T\|^2 \tag{C14}$$

for any $\epsilon > 0$. Recall that all of the parameters are rescaled to render the entire problem, and each of its constituents, dimensionless, therefore ϵ is also dimensionless.

Apply Lemma 3 to the left-hand side, to deduce:

$$\frac{d}{dt} \|T_n\|^2 + 2(\lambda - \epsilon) \|T_n\|^2 \leq g_A^T |\!|\!| T^* |\!|\!|^2 + \frac{1}{2\epsilon} \|f_T\|^2 \equiv \zeta. \tag{C15}$$

By choosing $0 < \epsilon < \lambda$, and defining $\eta \equiv 2(\lambda - \epsilon) > 0$, we obtain a bound on the growth of T_n, $\frac{d}{dt} \|T_n\|^2 + \eta \|T_n\|^2 \leq \zeta$, and using Gronwall's inequality,

$$\|T_n\|^2 \leq e^{-\eta t} \|T_0\|^2 + \frac{\zeta}{\eta}(1 - e^{-\eta t}).$$

This means that $\|T_n\|$ is bounded, and the bound is independent of n, and for all $t \in [0,\mathcal{T}]$:

$$\|T_n\|^2 (t) \leq C_1^T \equiv \|T_0\|^2 + \frac{g_A^T}{2\lambda(1-a)} |\!|\!| T^* |\!|\!|^2 + \frac{1}{4\lambda^2 a(1-a)} \|f_T\|^2 \tag{C16}$$

where $a \in (0,1)$, Observe that C_1^T obtains its minimal value for $a = \frac{1}{1+\sqrt{1+2g_A^T \frac{|\!|\!| T^* |\!|\!|^2}{\|f_T\|^2}}}$.

(ii) Uniform in n, $L^2(0,\mathcal{T};H^1(\Omega))$ bounds on T_n:
Combining equations (C14) and (C16) yields:

$$\frac{1}{2}\frac{d}{dt} \|T_n\|^2 + \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla} T_n\right\|^2 \leq a_2 \lambda C_1^T + \frac{g_A^T}{2} |\!|\!| T^* |\!|\!|^2 + \frac{1}{4a_2\lambda} \|f_T\|^2$$

for any $a_2 \in (0,1)$. Integrate over $[0,\mathcal{T}]$, and set $a_2 = a$:

$$\frac{1}{2}\|T\|^2(\mathcal{T}) + \int_0^{\mathcal{T}} dt \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}T\right\|^2 \leq \frac{1}{2}\|T_0\|^2 + \mathcal{T}\left(a\lambda\|T_0\|^2 + \frac{g_A^T}{2(1-a)}\tilde{\|}T^*\tilde{\|}^2 + \frac{1}{4(1-a)a\lambda}\|f_T\|^2\right).$$

This is true for any value of $a \in (0,1)$, and the value of a can be tweaked to obtain an optimal bound. However, we do not care about the exact value of this bound, so to simplify notation, in this subsection we set $a = 1/2$ to obtain:

$$\frac{1}{2}\|T\|^2(\mathcal{T}) + \int_0^{\mathcal{T}} dt \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}T\right\|^2 \leq C_2^T + \mathcal{T}C_3^T$$

where

$$C_2^T \equiv \frac{1}{2}\|T_0\|^2; \quad C_3^T \equiv \frac{\lambda}{2}\|T_0\|^2 + g_A^T\tilde{\|}T^*\tilde{\|}^2 + \frac{1}{\lambda}\|f_T\|^2.$$

(iii) Uniform in n, $L^\infty([0,\mathcal{T}];L^2(\Omega))$ bounds on S_n:

Taking the same approach for S_n here as we did for T_n, we note that the only difference between the two functions is in the boundary conditions, therefore we can immediately write:

$$\frac{1}{2}\frac{d}{dt}\|S_n\|^2 - g_A^S\int_{\sigma_1} S_n(x,y,z=1,t)S^*(x,y)d\sigma + \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S_n\right\|^2 = \int_\Omega S_n f_S dV$$

for almost all times $[0,\mathcal{T}]$. Rearranging the terms and using Cauchy-Schwarz and Young inequalities on the boundary and source terms, we obtain:

$$\frac{1}{2}\frac{d}{dt}\|S_n\|^2 + \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S_n\right\|^2 \leq \frac{g_A^S}{2\epsilon_1}\tilde{\|}S^*\tilde{\|}^2 + \frac{1}{2\epsilon_2}\|f_S\|^2 + \frac{g_A^S}{2}\epsilon_1\tilde{\|}S_n\tilde{\|}^2 + \frac{\epsilon_2}{2}\|S_n\|^2,$$

for any $\epsilon_1, \epsilon_2 > 0$. Using Lemma 4, we are left with:

$$\frac{1}{2}\frac{d}{dt}\|S_n\|^2 + \left(1 - \frac{g_A^S}{2\kappa_{min}}\epsilon_1\right)\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S_n\right\|^2 \leq \frac{g_A^S}{2\epsilon_1}\tilde{\|}S^*\tilde{\|}^2 + \frac{1}{2\epsilon_2}\|f_S\|^2 + \left(\frac{\epsilon_2}{2} + g_A^S\epsilon_1\right)\|S_n\|^2, \tag{C17}$$

where $\epsilon_1 < \frac{2\kappa_{min}}{g_A^S}$. Since we set $\langle S_n\rangle_\Omega = 0$, Lemma 5 gives us

$$\frac{d}{dt}\|S_n\|^2 + \nu\|S_n\|^2 \leq \frac{g_A^S}{\epsilon_1}\tilde{\|}S^*\tilde{\|}^2 + \frac{1}{\epsilon_2}\|f_S\|^2,$$

where $\nu \equiv 2\pi\kappa_{min} - (\pi+2)g_A^S\epsilon_1 - \epsilon_2$. Choosing ϵ_2 small enough such that $\nu > 0$, we can use the Gronwall inequality to establish a bound on $\|S_n\|(t)$ for all $t \in [0,\mathcal{T}]$:

$$\|S_n\|^2(t) \leq \|S_0\|^2 + \frac{g_A^S}{\epsilon_1\nu}\tilde{\|}S^*\tilde{\|}^2 + \frac{1}{\epsilon_2\nu}\|f_S\|^2. \tag{C18}$$

Taking $\epsilon_1 = 2\frac{\kappa_{min}}{g_A^S}\frac{b}{1+2/\pi}$, $\epsilon_2 = 2\kappa_{min}\pi d$ for $b, d > 0$ satisfying $b + d < 1$, the requirements on ϵ_1, ϵ_2 are guaranteed. One can check that the optimal values for b and d are given by $b = \left(2 + \frac{2}{g_A^S\sqrt{2+\pi}}\frac{\|f_S\|}{\tilde{\|}S^*\tilde{\|}}\right)^{-1}$, $d = \frac{1}{2} - b$, and the bound simplifies to:

$$\|S_n\|^2(t) \leq C_1^S \equiv \|S_0\|^2 + \frac{1}{2\pi\kappa_{min}^2}\left(\frac{1+\frac{2}{\pi}}{b}(g_A^S)^2\tilde{\|}S^*\tilde{\|}^2 + \frac{1}{\pi(\frac{1}{2}-b)}\|f_S\|^2\right).$$

(iv) Uniform in n, $L^2(0,\mathcal{T};\dot{H}^1(\Omega))$ bounds on S_n:

Using the bound on $\|S_n\|^2$ obtained above with equation (C17), we deduce:

$$\frac{1}{2}\frac{d}{dt}\|S_n\|^2 + (1 - \frac{g_A^S}{2\kappa_{min}}\epsilon_1)\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S_n\right\|^2 \leq \frac{g_A^S}{2\epsilon_1}\tilde{\|}S^*\tilde{\|}^2 + \frac{1}{2\epsilon_2}\|f_S\|^2 + (\frac{\epsilon_2}{2} + g_A^S\epsilon_1)\|S_n\|^2$$

for any ϵ_1, ϵ_2. Integrate over the time interval $[0, \mathcal{T}]$, and use equation (C18) to deduce

$$\int_0^{\mathcal{T}} dt \left\| \boldsymbol{\kappa}^{1/2} \boldsymbol{\nabla} S \right\|^2 \leq \frac{\kappa_{min}}{2\kappa_{min} - g_A^S \epsilon_1} \|S_0\|^2 +$$
$$+ \mathcal{T} \frac{(2g_A^S \epsilon_1 + \epsilon_2)\kappa_{min}}{2\kappa_{min} - g_A^S \epsilon_1} \|S_0\|^2 + \frac{\kappa_{min}}{\epsilon_1 (2\kappa_{min} - g_A^S \epsilon_1 (1 + 2/\pi) - \epsilon_2/\pi)} (\frac{1}{\epsilon_2} \|f_S\|^2 + g_A^S \tilde{\|} S^* \tilde{\|}^2).$$

One can enlarge the bounds and choose values for ϵ_1 and ϵ_2 small enough to obtain

$$\int_0^{\mathcal{T}} \left\| \boldsymbol{\kappa}^{1/2} \boldsymbol{\nabla} S \right\|^2 dt \leq C_2^S + \mathcal{T} C_3^S,$$

where

$$C_2^S \equiv \frac{\pi \kappa_{min}}{2} \|S_0\|^2 ; \quad C_3^S \equiv \frac{\pi \kappa_{min}}{2} \|S_0\|^2 + 6 \frac{g_A^S}{\kappa_{min}} \left(g_A^S \tilde{\|} S^* \tilde{\|}^2 + \frac{2}{\pi \kappa_{min}} \|f_S\|^2 \right).$$

□

c. *Convergence of the iterative approximate sequences*

In this section we prove that the sequence $\{(T_n, S_n)\}_{n=1}^{\infty}$ of strong solutions to the corresponding iteratively defined problems $\left\{ \left(P^{linear}_{T_0, S_0, T_{n-1}, S_{n-1}} \right) \right\}_{n=1}^{\infty}$ established in section C 2 a converges to limit functions, that we denote (T_∞, S_∞). Eventually, we will prove that these limit functions are solutions to the nonlinear problem $\left(P_{T_0, S_0, T^*, S^*, f_T, f_S} \right)$.

Proposition 3 *Let $\mathcal{T} > 0$ be given, and let $\{(T_n, S_n)\}_{n=1}^{\infty}$ be strong solutions on $[0, \mathcal{T}]$ of the iteratively defined sequence of problems $\left\{ \left(P^{linear}_{T_0, S_0, T_{n-1}, S_{n-1}} \right) \right\}_{n=1}^{\infty}$, as established in C 2 a. There exists $\tau \in (0, \mathcal{T}]$ such that $\{T_n\}_{n=1}^{\infty}$, $\{S_n\}_{n=1}^{\infty}$ converge to limit functions, respectively T_∞ and S_∞, strongly in $L^\infty([0, \tau]; L^2(\Omega))$.*

Proof: It is enough to show that there exists $\tau \in (0, \mathcal{T}]$ for which $\{(T_n, S_n)\}_{n=1}^{\infty}$ is a Cauchy sequence in the Banach space $(L^\infty([0, \tau]; L^2(\Omega)))^2$, hence it has a limit $(T_\infty, S_\infty) \in (L^\infty([0, \tau]; L^2(\Omega)))^2$.

Define $\delta T_n \equiv T_n - T_{n-1}$, $\delta S_n \equiv S_n - S_{n-1}$, $\eta_n(t) \equiv \|\delta T_n\|^2(t) + \|\delta S_n\|^2(t)$, $n \in \mathbb{N}$. Let $m \in \mathbb{N}$. Then δT_{m+1} can be used as a test function for weak solutions in equation (C10) according to Lemma 2. Thus, we subtract equation (C10) for T_{m+1} in place of T_n and δT_{m+1} in place of ψ, from equation (C10) for T_m in place of T_n and δT_{m+1} in place of ψ, deducing:

$$\int_0^{\mathcal{T}} dt \left(\frac{1}{2} \frac{d}{dt} \|\delta T_{m+1}\|^2 + \left\| \boldsymbol{\kappa}^{1/2} \boldsymbol{\nabla} \delta T_{m+1} \right\| + g_A^T \tilde{\|} \delta T_{m+1} \tilde{\|} + \int_\Omega \delta T_{m+1} \delta \boldsymbol{u}_m \cdot (\boldsymbol{\nabla} T_m) \right) = 0,$$

where we define $\delta \boldsymbol{u}_m \equiv \boldsymbol{u}_m - \boldsymbol{u}_{m-1}$, add and subtract $T_m \boldsymbol{u}_m \cdot (\boldsymbol{\nabla} T_m)$, and use the generalized divergence theorem [12] and assumptions $(a4)$, $(a5)$ in (C1). Therefore,

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\delta T_{m+1}\|^2 (s) &\leq \frac{1}{2} \frac{d}{dt} \|\delta T_{m+1}\|^2 (s) + \left\| \boldsymbol{\kappa}^{1/2} \boldsymbol{\nabla} \delta T_{m+1} \right\|^2 (s) + g_A^T \tilde{\|} \delta T_{m+1} \tilde{\|}^2 (s) \\ &\leq \int_\Omega |\delta T_{m+1}| |\delta \boldsymbol{u}_m| |\boldsymbol{\nabla} T_m| (t) dV, \end{aligned} \tag{C19}$$

where as before, we find bounds for a general time $s \in [0, \mathcal{T}]$, and will eventually integrate over time. We note that $\delta \boldsymbol{u}_m = (a_I^m - a_I^{m-1}) \boldsymbol{u}_I$. Since $\boldsymbol{u}_I$ is a known, bounded function by construction, we can bound it with its L^∞ norm $u_I^{\max} = \max_{\boldsymbol{r} \in \Omega} |\boldsymbol{u}_I(\boldsymbol{r})| < \infty$. Thus,

$$\begin{aligned} \|\delta \boldsymbol{u}_m\|_{L^\infty(\Omega)} &\leq u_I^{\max} \Gamma \left(\frac{1}{|D_1|} \int_{D_1} (\alpha |\delta T_m| + \beta |\delta S_m|) dV + \frac{1}{|D_2|} \int_{D_2} (\alpha |\delta T_m| + \beta |\delta S_m|) dV \right) \\ &\leq c \int_\Omega (|\delta T_m| + |\delta S_m|), \end{aligned} \tag{C20}$$

where $c = u_I^{\max}\Gamma\frac{\max\{\alpha,\beta\}}{\min\{|D_1|,|D_2|\}}$. By (C19), (C20), and Cauchy-Schwarz, we obtain

$$\frac{d}{dt}\left\|\delta T_{m+1}\right\|^2 = 2\left\|\delta T_{m+1}\right\|\frac{d}{dt}\left\|\delta T_{m+1}\right\| \leq 2c\left\|\delta T_{m+1}\right\|\left\|\boldsymbol{\nabla} T_m\right\|\left(\int_\Omega (|\delta T_m| + |\delta S_m|)dV\right),$$

therefore

$$\frac{d}{dt}\left\|\delta T_{m+1}\right\| \leq c\left\|\boldsymbol{\nabla} T_m\right\|\left(\int_\Omega (|\delta T_m| + |\delta S_m|)dV\right).$$

Integrate over the time interval $[0,t]$ for $t \in [0,\mathcal{T}]$ to obtain:

$$\left\|\delta T_{m+1}\right\|(t) \leq c\sup_{0\leq s\leq t}\left\{\int_\Omega (|\delta T_m|(s) + |\delta S_m|(s))\,dV\right\}\int_0^t \left\|\boldsymbol{\nabla} T_m\right\|dt,$$

while observing that $\left\|\delta T_{m+1}\right\|(0) = 0$. By the Cauchy-Schwarz inequality,

$$\left\|\delta T_{m+1}\right\|(t) \leq c\sup_{0\leq s\leq t}\left\{\int_\Omega (|\delta T_m| + |\delta S_m|)\,dV\right\}\left(\int_0^t \left\|\boldsymbol{\nabla} T_m\right\|^2 dt\right)^{1/2} t^{1/2}.$$

By virtue of bound (ii) in Proposition 2,

$$\left\|\delta T_{m+1}\right\|(t) \leq b_T\sup_{0\leq s\leq t}\left\{\int_\Omega (|\delta T_m| + |\delta S_m|)\,dV\right\}(1+t)^{1/2}t^{1/2}$$

where $b_T = \frac{c\max\{C_2^T, C_3^T\}^{1/2}}{\min\{\kappa_x,\kappa_y,\kappa_z\}}$. Observe that

$$\left(\int_\Omega (|\delta T_m| + |\delta S_m|)dV\right)^2 \leq 2\left(\int_\Omega |\delta T_m|dV\right)^2 + 2\left(\int_\Omega |\delta S_m|dV\right)^2 \leq 2|\Omega|\eta_m,$$

therefore

$$\left\|\delta T_{m+1}\right\|^2(t) \leq 2|\Omega|b_T^2\sup_{0\leq s\leq t}\{\eta_m(s)\}(1+t)t. \tag{C21}$$

Similar calculations may be performed for δS_{m+1}, yielding

$$\left\|\delta S_{m+1}\right\|^2(t) \leq 2|\Omega|b_S^2\sup_{0\leq s\leq t}\{\eta_m(s)\}(1+t)t, \tag{C22}$$

where $b_S = \frac{c\max\{C_2^S, C_3^S\}^{1/2}}{2\min\{\kappa_x,\kappa_y,\kappa_z\}}$. Combining (C21) and (C22) yields, for any $t \in [0,\mathcal{T}]$ and $\tau \in [t,\mathcal{T}]$,

$$\begin{aligned}\eta_{m+1}(t) &\leq 2|\Omega|\max\{b_T^2, b_S^2\}(1+t)t\sup_{0\leq s\leq t}\{\eta_m(s)\}\\ &\leq 2|\Omega|\max\{b_T^2, b_S^2\}(1+\tau)\tau\sup_{0\leq s\leq \tau}\{\eta_m(s)\}.\end{aligned} \tag{C23}$$

Hence, (C23) implies

$$\sup_{0\leq s\leq \tau}\{\eta_{m+1}(s)\} \leq \theta\sup_{0\leq s\leq \tau}\{\eta_m(s)\}$$

where

$$\theta = b(1+\tau)\tau,\ with\ b = 2|\Omega|\max\{b_T^2, b_S^2\}.$$

Let τ be chosen small enough such that $\theta = 1/2$. This yields

$$\sup_{0\leq s\leq \tau}\{\eta_{m+1}\} \leq 2^{-m}\sup_{0\leq s\leq \tau}\{\eta_1\},$$

which implies that the sequence $\{(T_n, S_n)\}_{n=1}^\infty$ is a Cauchy sequence in $(L^\infty([0,\tau]; L^2(\Omega)))^2$. □

d. (T_∞, S_∞) is a global solution to the nonlinear model $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$

Proposition 4 *Let (T_∞, S_∞) and $\tau > 0$ be as in Proposition 3. Then (T_∞, S_∞) is a strong solution to $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ on $[0,\tau]$. Furthermore, T_∞ and S_∞ satisfy bounds (i), (ii) and (iii), (iv), respectively, of Proposition 2, where n is replaced by ∞ and $\mathcal{T}$ is replaced by τ.*

Proof: In this proof, we focus on the time interval $[0,\tau]$. By Proposition 1, $\{T_n\}_{n=1}^\infty, \{S_n\}_{n=1}^\infty \subset C([0,\tau];L^2(\Omega))$. Moreover, by the proof of Proposition 3 the sequences are Cauchy in $L^\infty([0,\tau];L^2(\Omega))$. Therefore, since the convergence is uniform in time we conclude that $T_\infty, S_\infty \in C([0,\tau];L^2(\Omega))$.

By Proposition 2, the sequence $\{T_n\}_{n=1}^\infty$ is bounded in the Hilbert space $L^2(0,\tau;H^1(\Omega))$; hence it has a subsequence $\{T_{n_k}\}_{k=1}^\infty$ that converges weakly to $\bar{T}$ in $L^2(0,\tau;H^1(\Omega))$ and in $L^2(0,\tau;L^2(\Omega))$ [62]. On the other hand, from Proposition 3, T_∞ is a strong limit of $\{T_{n_k}\}_{k=1}^\infty$ in $L^\infty([0,\tau];L^2(\Omega))$. Since the L^∞ norm is stronger than the L^2 norm on a bounded interval, and from uniqueness of the weak limit, it follows that $\bar{T} = T_\infty$, therefore $T_\infty \in L^2(0,\tau;H^1(\Omega)) \cap C([0,\tau];L^2(\Omega))$. The same argument holds for S_∞. Therefore the limit functions T_∞, S_∞ satisfy the regularity conditions of weak solutions.

Given a test function $\psi \in H^1([0,\tau];H^1(\Omega))$, we want to show that T_∞, S_∞ satisfy equation (C3). For all $n \in \mathbb{N}$, T_n, S_n satisfy equations (C8), (C9) respectively, with T_n replacing T, S_n replacing S, and $\boldsymbol{u}_{n-1}$ in place of $\boldsymbol{u}_m$. Due to the strong convergence of $\{T_n\}_{n=1}^\infty, \{S_n\}_{n=1}^\infty$ in $C([0,\tau];L^2(\Omega))$, the first term in each of the equations (C8), (C9) converges to, respectively, $\int_0^\tau \int_\Omega T_\infty(t)\partial_t\psi(t)dtdV$, $\int_0^\tau \int_\Omega S_\infty(t)\partial_t\psi(t)dtdV$.

Let us move into the subsequence $\{T_{n_k}\}_{k=1}^\infty$. Then

$$\int_0^\tau \int_\Omega \left(\boldsymbol{\kappa}(\boldsymbol{\nabla} T_{n_k} - \boldsymbol{\nabla} T_\infty)\cdot(\boldsymbol{\nabla}\psi)\right) dVdt \underset{k\to\infty}{\to} 0$$

due to the weak convergence in $L^2(0,\tau;H^1(\Omega))$. The equivalent argument shows that the same thing is true for S_∞ as well. Regarding the velocity term, observe that

$$|\boldsymbol{u}_{n_k-1}| \le \mathcal{U} \equiv |a_E| u_E^{\max} + u_I^{\max} \frac{\Gamma|\Omega|^{1/2}}{\min\{|D_1|,|D_2|\}} (\alpha (C_1^T)^{1/2} + \beta (C_1^S)^{1/2}),$$

where $u_E^{\max} = \sup_{\boldsymbol{r}\in\Omega} |\boldsymbol{u}_E|$. Therefore, by the strong convergence of T_n to T_∞ in $C([0,\tau];L^2(\Omega))$ and the Cauchy-Schwarz inequality,

$$\left|\int_0^\tau \int_\Omega \boldsymbol{u}_{n_k-1}(T_{n_k} - T_\infty)\cdot\boldsymbol{\nabla}\psi dVdt\right| \le \mathcal{U}\int_0^\tau\int_\Omega |T_{n_k} - T_\infty||\boldsymbol{\nabla}\psi| dVdt \underset{k\to\infty}{\to} 0.$$

From equation (C20),

$$\begin{aligned}
&\left|\int_0^\tau \int_\Omega T_\infty(\boldsymbol{u}_{n_k-1} - \boldsymbol{u}_\infty)\cdot(\boldsymbol{\nabla}\psi) dVdt\right| \\
&\le c\int_0^\tau \left(\|T_\infty\|\,\|\boldsymbol{\nabla}\psi\| \int_\Omega \left(|T_{n_k-1} - T_\infty| + |S_{n_k-1} - S_\infty|\right) dV\right) dt \\
&\le 2cd|\Omega|^{1/2}\left(\int_0^\tau \|\boldsymbol{\nabla}\psi\| \left(\int_\Omega \left(|T_{n_k-1} - T_\infty|^2 + |S_{n_k-1} - S_\infty|^2\right) dV\right)^{1/2} dt\right) \\
&\le 2cdg|\Omega|^{1/2}\left(\int_0^\tau\int_\Omega \left(|T_{n_k-1} - T_\infty|^2 + |S_{n_k-1} - S_\infty|^2\right) dVdt\right)^{1/2} \underset{k\to\infty}{\to} 0,
\end{aligned}$$

where c, d, g are some positive constants, and we used Cauchy-Schwarz and the regularity of T_∞ and ψ. The same calculation holds for S_∞.

The last term we need to take care of is the boundary term, $\int_0^\tau \int_{\sigma_1} g_A^T T_n(t)\psi(t)dxdydt$ in equation (C8), relevant only to the temperature equation. We define an auxilliary function $\beta_T(q) = \begin{cases} g_A^T & \text{if } q \in \sigma_1 \\ 0 & \text{if } q \in \partial\Omega \setminus \sigma_1 \end{cases}$. Thus,

$$\int_0^\tau \int_{\partial\Omega} \beta_T(T_{n_k} - T_\infty)\psi d\sigma dt \underset{k\to\infty}{\to} 0,$$

since T_{n_k} converges weakly in $L^2([0,\tau];H^1(\Omega))$ to T_∞. Therefore, (T_∞, S_∞) is a weak solution of $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ on $[0,\tau]$.

Next, we prove that the weak solution (T_∞, S_∞) established above is indeed a strong solution of $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ on $[0,\tau]$. Let us define $\boldsymbol{u}_\infty \equiv \boldsymbol{u}_\infty(\boldsymbol{r};\langle T_\infty\rangle,\langle S_\infty\rangle)$ following the notation of $(P4^l)$ in (C7). Then $\boldsymbol{u}_\infty \in L^\infty((0,\tau);L^\infty(\Omega))^3$ is given, and can be used to define the linear problem $\left(P^{linear}_{T_0,S_0,T_\infty,S_\infty}\right)$ on $[0,\tau]$. Due to Remark 2.15 in [43], this problem has a strong solution (T,S). Observe that (T_∞,S_∞) is also a weak solution of the linear problem $\left(P^{linear}_{T_0,S_0,T_\infty,S_\infty}\right)$ on $[0,\tau]$, since it solves the nonlinear problem. Since a strong solution to the linear problem is also a weak solution, and weak solutions of the linear problem are unique, the strong solution (T,S) must equal (T_∞,S_∞) on the time interval $[0,\tau]$.

Regarding the bounds of Proposition 2, since T_∞, S_∞ are strong limits of T_{n_k}, S_{n_k} in $C([0,\tau];L^2(\Omega))$ and weak limits in $L^2(0,\tau;H^1(\Omega))$, respectively, then T_∞, S_∞ enjoy the same bounds as the sequence itself, as established in Proposition 2. □

Next, we prove that the solution (T_∞,S_∞) as established in Proposition 4 is unique, and that it has a continuous dependence on the data of the system, namely the initial and boundary conditions.

Proposition 5 *There exists $\bar\tau > 0$ such that $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ has a unique strong solution on $[0,\bar\tau]$, with a continuous dependence on the problem's data - the initial conditions, the boundary conditions, and the sources.*

Proof: Let $(T_1^{\rm nl},S_1^{\rm nl})$ and $(T_2^{\rm nl},S_2^{\rm nl})$ be strong solutions of the nonlinear problems $\left(P_{T_{0,i},S_{0,i},T_i^*,S_i^*,f_{T,i},f_{S,i}}\right)$ on $[0,\tau_i]$, $\tau_i>0$, for $i=1$ and $i=2$, respectively, as established in Proposition 4. Define $\bar\tau \equiv \min\{\tau_1,\tau_2\}$, $\delta T \equiv T_2^{\rm nl}-T_1^{\rm nl}$, $\delta S \equiv S_2^{\rm nl}-S_1^{\rm nl}$. Then δT, δS can be used as test functions for weak solutions on $[0,\bar\tau]$ according to Lemma 2.

Let us start with considering equation (C3) for the temperature, with $T_i^{\rm nl}$, $S_i^{\rm nl}$ in place of T, S for $i=1,2$ and δT as a test function in place of ψ. The difference between the equations satisfies

$$\int_0^{\bar\tau}\left(\frac{1}{2}\frac{d}{dt}\|\delta T\|^2+\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}\delta T\right\|^2-\int_\Omega(\boldsymbol{u}_2T_2^{\rm nl}-\boldsymbol{u}_1T_1^{\rm nl})\cdot(\boldsymbol{\nabla}\delta T)dV+g_A^T\tilde{\|}\delta T\tilde{\|}^2\right)dt=$$
$$\int_0^{\bar\tau}\left(\int_\Omega\delta f_T\,\delta T\,dV+\int_{\sigma_1}g_A^T\delta T^*\delta T\,dxdy\right)dt$$

where $\boldsymbol{u}_i=\boldsymbol{u}\left(\boldsymbol{r};\left\langle T_i^{\rm nl}\right\rangle,\left\langle S_i^{\rm nl}\right\rangle\right)$ for $i=1,2$, $\delta f_T = f_{T,2}-f_{T,1}$ and $\delta T^* = T_2^*-T_1^*$. Inside the time integrand we add and subtract $\int_\Omega(\boldsymbol{u}_2T_1^{\rm nl}\cdot(\boldsymbol{\nabla}\delta T))$. We also use (a4), (a5) in (C1) to simplify the velocity terms and deduce

$$\frac{1}{2}\frac{d}{dt}\|\delta T\|^2+\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}\delta T\right\|^2+g_A^T\tilde{\|}\delta T\tilde{\|}^2\le\int_\Omega\left(|\delta T||\delta\boldsymbol{u}||\boldsymbol{\nabla}T_1^{\rm nl}|+|\delta f_T||\delta T|\right)dV+g_A^T\int_{\sigma_1}|\delta T^*||\delta T|d\sigma$$

where $\delta\boldsymbol{u}\equiv\boldsymbol{u}_2-\boldsymbol{u}_1$. Employing Cauchy-Schwarz and using Lemma 3, we deduce:

$$\frac{1}{2}\frac{d}{dt}\|\delta T\|^2\le\int_\Omega|\delta T||\delta\boldsymbol{u}||\boldsymbol{\nabla}T_1^{\rm nl}|dV+\frac{1}{2\lambda}\|\delta f_T\|^2+\frac{g_A^T}{2}\tilde{\|}\delta T^*\tilde{\|}^2.$$

Use the bound on $\delta\boldsymbol{u}$ from (C20), and employ Cauchy-Schwarz again, to obtain:

$$\frac{d}{dt}\|\delta T\|^2\le h(\|\delta T\|+\|\delta S\|)\|\delta T\|\left\|\boldsymbol{\nabla}T_1^{\rm nl}\right\|+\frac{1}{\lambda}\|\delta f_T\|^2+g_A^T\tilde{\|}\delta T^*\tilde{\|}^2,\tag{C24}$$

where $h=2u_I^{\max}\Gamma\frac{\max\{\alpha,\beta\}}{\min\{|D_1|,|D_2|\}}|\Omega|^{1/2}$. Applying similar arguments on S and using Lemmas 4 and 5 , there is some constant $\epsilon>0$ small enough for which we have

$$\frac{d}{dt}\|\delta S\|^2\le h(\|\delta T\|+\|\delta S\|)\|\delta S\|\left\|\boldsymbol{\nabla}S_1^{\rm nl}\right\|+\frac{1}{2\epsilon}\|\delta f_S\|^2+\frac{g_A^S}{2\epsilon}\tilde{\|}\delta S^*\tilde{\|}^2.\tag{C25}$$

Defining $\eta\equiv\|\delta T\|^2+\|\delta S\|^2$, from (C24) and (C25) we deduce

$$\frac{d}{dt}\eta-\zeta(t)\eta\le\mu$$

where $\zeta(t)=h\left(\left\|\boldsymbol{\nabla}T_1^{\rm nl}\right\|+\left\|\boldsymbol{\nabla}S_1^{\rm nl}\right\|\right)$, and $\mu=\frac{1}{\lambda}\|\delta f_T\|^2+g_A^T\tilde{\|}\delta T^*\tilde{\|}^2+\frac{1}{2\epsilon}\|\delta f_S\|^2+\frac{g_A^S}{2\epsilon}\tilde{\|}\delta S^*\tilde{\|}^2$. Then, by Gronwall's inequality, for $t\in[0,\bar\tau]$,

$$\eta(t)\le(\eta(0)+\mu t)\exp\left(\int_0^t\zeta(t')dt'\right)\le(\eta(0)+\mu t)\exp\left(\left(\tilde at+\tilde bt^2\right)^{1/2}\right),$$

where we employ the bounds (ii) and (iv) from Proposition 2, and $\tilde{a}$ and $\tilde{b}$ are positive constants that depend on the readily established estimates on solutions, equations (C6). Hence,

$$\sup_{t\in[0,\bar{\tau}]}\eta \leq (\eta(0)+\mu\bar{\tau})\exp\left(\left(\tilde{a}\bar{\tau}+\tilde{b}\bar{\tau}^2\right)^{1/2}\right), \tag{C26}$$

implying uniqueness and continuous dependence on data on the time interval $[0,\bar{\tau}]$.

□

Proposition 6 *Let (T,S) be a strong solution of $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ on $[0,\mathcal{T}]$ for $\mathcal{T}>0$. Then, bounds (C5) are satisfied on the time interval $[0,\mathcal{T}]$.*

Proof: By Remark 6, a strong solution can be used as a test function in the analogous equations to (C10) and (C11), as described in the remark. Then, following similar arguments as in Proposition 2, one can conclude the proof of our proposition. □

Next, we establish global existence and uniqueness.

Proposition 7 *Let $\bar{\mathcal{T}}\in(0,\infty)$ be given and let (C1) hold. Then:*

(a) Problem $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ has a unique strong solution (T,S) on $[0,\bar{\mathcal{T}}]$, which depends continuously on the initial conditions, boundary functions and sources in the sense described in equation (C26), and satisfies the bounds from equations (C5) and (C6).

(b) If the initial conditions additionally satisfy assumptions (C4) then (T,S) is also a classical L^2-solution of $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ on $[0,\bar{\mathcal{T}}]$.

Proof:

(a) Let $\tau>0$ such that (T,S) is a unique strong solution of $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ on $[0,\tau]$, according to Propositions 4 and 5. Let $\mathcal{T}\geq\tau$ be the maximal interval of existence of the strong solution, $\mathcal{T}=\sup\{s\in[\tau,\infty]:\ \forall t<s,\ (T,S)$ is a unique strong solution on $[0,t$

If $\mathcal{T}=\infty$, the solution is global in time and we are done.

Assume by contradiction that $\mathcal{T}<\infty$. Then necessarily $\limsup_{t\to\mathcal{T}^-}\|T\|(t)=\infty$ or $\limsup_{t\to\mathcal{T}^-}\|S\|(t)=\infty$. Else, using Proposition 3, the solution can be extended beyond the maximal interval of existence $\mathcal{T}$, which is a contradiction. However, by Proposition 6 $\limsup_{t\to\mathcal{T}^-}\|T\|(t)<\infty$ and $\limsup_{t\to\mathcal{T}^-}\|S\|(t)<\infty$, which contradicts the fact that $\mathcal{T}<\infty$. Hence, $\mathcal{T}$ must be infinite.

Uniqueness and continuous dependence on initial conditions are established in the same manner as in Proposition 5.

(b) Since (T,S) is a strong solution of $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ on $[0,\bar{\mathcal{T}}]$, we can define $\boldsymbol{u}\equiv\boldsymbol{u}(\boldsymbol{r};\langle T\rangle,\langle S\rangle)$ from $(P4^l)$. Then $\boldsymbol{u}\in L^\infty(0,\bar{\mathcal{T}};L^\infty(\Omega))^3$, and can be considered as a given function and used to define the linear problem $\left(P^{linear}_{T_0,S_0,T,S}\right)$. Due to Proposition 2.7 in [43], under the additional assumptions (C4) this problem has a classical L^2-solution $(T^{\mathbf{C}},S^{\mathbf{C}})$ on $[0,\bar{\mathcal{T}}]$. Since a classical solution is also a strong solution, and strong solutions of the linear problem are unique, then $(T^{\mathbf{C}},S^{\mathbf{C}})=(T,S)$ on $[0,\bar{\mathcal{T}}]$.

□

To summarize, we have proved the main theorem, that the nonlinear problem $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ is well-posed in the sense of Hadamard: the model has

(i) a unique, global, strong solution,

(ii) with a smooth dependence on the problem's data: initial conditions, boundary conditions and sources.

(iii) This solution satisfies the bounds described in equation (C5).

(iv) With the additional regularity conditions on the initial conditions of (C4), the solution is also a classical L^2-solution.

□

Corollary 1 *Let $\boldsymbol{\kappa}$, Γ, α, β, a_E, $\boldsymbol{u}_E$, $\boldsymbol{u}_I$, f_T, f_S, g_A^T, g_A^S, T^*, S^*, T_0, and S_0 be as in (C1), and let (T,S) be the global strong solution of $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$, as established in Theorem 1. The dynamic weight of the velocity function, $a_I(\langle T\rangle,\langle S\rangle)$, as defined in equation (P3) in (C2), is bounded at all times:*

$$|a_I(\langle T\rangle,\langle S\rangle)|\leq\frac{\Gamma|\Omega|^{1/2}}{\min\{|D_1|,|D_2|\}}\left(\alpha(C_1^T)^{1/2}+\beta(C_1^S)^{1/2}\right),$$

where the values of C_1^T and C_1^S are given in equation (C6).

Proof: Using the Cauchy-Schwarz and the Young inequalities, and the bounds (C5) from Theorem 1, we obtain a bound on a_I:

$$|a_I(\langle T\rangle, \langle S\rangle)| \le \Gamma\left(\frac{1}{|D_1|}\int_{D_1} |-\alpha T+\beta S|dV + \frac{1}{|D_2|}\int_{D_2} |-\alpha T+\beta S|dV\right)$$
$$\le \frac{\Gamma|\Omega|^{1/2}}{\min\{|D_1|,|D_2|\}}\left(\alpha\|T\|+\beta\|S\|\right) \le \frac{\Gamma|\Omega|^{1/2}}{\min\{|D_1|,|D_2|\}}\left(\alpha(C_1^T)^{1/2}+\beta(C_1^S)^{1/2}\right).$$

□

3. The steady-state problem

This section investigates steady-state solutions to the nonlinear problem defined above. Consider the nonlinear time-independent problem

$$\left(P^{\text{SS}}\right)\begin{cases}
(P1^{\text{SS}}) & -\boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}T(\boldsymbol{r})) + (\boldsymbol{u}(\boldsymbol{r};\langle T\rangle,\langle S\rangle)\cdot\boldsymbol{\nabla})T(\boldsymbol{r}) = f_T(\boldsymbol{r}), & \boldsymbol{r}\in\Omega\\
(P2^{\text{SS}}) & -\boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}S(\boldsymbol{r})) + (\boldsymbol{u}(\boldsymbol{r};\langle T\rangle,\langle S\rangle)\cdot\boldsymbol{\nabla})S(\boldsymbol{r}) = f_S(\boldsymbol{r}), & \boldsymbol{r}\in\Omega\\
(P3^{\text{SS}}) & a_I(\langle T\rangle,\langle S\rangle) = \Gamma\left(-\alpha(\langle T\rangle_2-\langle T\rangle_1)+\beta(\langle S\rangle_2-\langle S\rangle_1)\right), & \\
(P4^{\text{SS}}) & \boldsymbol{u}(\boldsymbol{r};\langle T\rangle,\langle S\rangle) = a_E\boldsymbol{u}_E(\boldsymbol{r}) + a_I(\langle T\rangle,\langle S\rangle)\boldsymbol{u}_I(\boldsymbol{r}), & \boldsymbol{r}\in\Omega\\
(P5^{\text{SS}}) & (\boldsymbol{\kappa}\boldsymbol{\nabla}T(q))\cdot\hat{\boldsymbol{n}}(q) = \begin{cases} g_A^T(T^*(x,y)-T(q)) & \text{if } q\in\sigma_1\\ 0 & \text{else}\end{cases}, & q\in\partial\Omega\\
(P6^{\text{SS}}) & (\boldsymbol{\kappa}\boldsymbol{\nabla}S(q))\cdot\hat{\boldsymbol{n}}(q) = \begin{cases} g_A^S S^*(x,y) & \text{if } q\in\sigma_1\\ 0 & \text{else}\end{cases}, & q\in\partial\Omega.
\end{cases} \tag{C27}$$

where the parameters satisfy (C1). Since if S is a solution then so is $S+a$ for any constant a, then we restrict ourselves to the class where the average of S over Ω is zero as we did for the time-dependent problem. We thus define a weak solution to P^{SS} in the same spirit as of weak solutions of the nonlinear time-dependent problem:

Definition 8 *$(T,S)\in H^1(\Omega)\times\dot{H}^1(\Omega)$ is called a weak solution to $\left(P^{SS}\right)$ if, for all test functions $\varphi\in H^1(\Omega)$, $\psi\in\dot{H}^1(\Omega)$, the following holds:*

$$\int_\Omega((\boldsymbol{\kappa}\boldsymbol{\nabla}T)\cdot(\boldsymbol{\nabla}\varphi)+((\boldsymbol{u}\cdot\nabla)T)\varphi)\,dV + \int_{\sigma_1} g_A^T T\varphi dxdy = \int_\Omega f_T\varphi dV + \int_{\sigma_1} g_A^T T^*(x,y)\varphi dxdy;$$
$$\int_\Omega((\boldsymbol{\kappa}\boldsymbol{\nabla}S)\cdot(\boldsymbol{\nabla}\psi)+((\boldsymbol{u}\cdot\nabla)S)\psi)\,dV = \int_\Omega f_S\psi dV + \int_{\sigma_1} g_A^S S^*(x,y)\psi dxdy, \tag{C28}$$

where $\boldsymbol{u}=\boldsymbol{u}(\boldsymbol{r};\langle T\rangle,\langle S\rangle)$ is defined according to equation $(P4)$ in (C2).

Remark 8 *A weak solution (T,S) to the steady state problem $\left(P^{SS}\right)$ is a strong solution of the nonlinear time-dependent problem $\left(P_{T,S,T^*,S^*,f_T,f_S}\right)$ that is independent of time.*

To conclude this appendix, in Theorem 2 we show that for any set of parameters, boundary functions and source functions satisfying (C1), the nonlinear steady-state problem $\left(P^{\text{SS}}\right)$ has a weak solution with bounded norms. Note that we do not show uniqueness for the general problem; indeed for a given set of parameters, uniqueness of a weak solution to the nonlinear steady state problem $\left(P^{\text{SS}}\right)$ is not guaranteed nor expected in general. However, we show that given some restriction on the size of the parameters under which the system is not vigorously forced with respect to its dissipation, the steady-state solution *is* unique; furthermore, all solutions to the time-dependent problem converge to this unique steady state solution as $t\to\infty$, as could be expected from a dissipative dynamical system induced by advection-diffusion-type problems. In fact, one should be able to show that the infinite-dimensional dynamical system induced by the evolution of the nonlinear problem $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$ possess a finite-dimensional global attractor, a subject which is outside the scope of this article [53].

Theorem 2 *Let $\boldsymbol{\kappa}$, Γ, α, β, a_E, $\boldsymbol{u}_E$, $\boldsymbol{u}_I$, f_T, f_S, g_A^T, g_A^S, T^*, and S^* be as in (C1). Then:*

1. *The nonlinear steady-state problem* $\left(P^{SS}\right)$ *has a weak solution.*

2. *A solution* (T,S) *to* $\left(P^{SS}\right)$ *satisfies the following bounds:*

$$\|T\|^2 \le C_4^T, \quad \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}T\right\|^2 \le C_5^T, \quad \|S\|^2 \le C_4^S, \quad \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S\right\|^2 \le C_5^S; \tag{C29}$$

where

$$\begin{aligned} C_4^T &= \frac{1}{\lambda^2}\|f_T\|^2 + \frac{g_A^T}{\lambda}\tilde{\|}T^*\tilde{\|}^2, \quad C_5^T = \frac{1}{\lambda}\|f_T\|^2 + \frac{2g_A^T}{3}\tilde{\|}T^*\tilde{\|}^2, \\ C_4^S &= \frac{1}{\epsilon_1}\tilde{\|}S^*\tilde{\|}^2 + \frac{1}{\epsilon_2}\|f_S\|^2, \quad C_5^S = \frac{1}{\epsilon_1}\tilde{\|}S^*\tilde{\|}^2 + \frac{1}{\epsilon_3}\|f_S\|^2. \end{aligned} \tag{C30}$$

The constants are given by $\epsilon_1 = 2\kappa_{min}/\left(g_A^S\right)^2$, $\epsilon_2 = 4\kappa_{min}\left(3\pi/8-1\right)$, $\epsilon_3 = 2\left(\kappa_{min}(\pi-2)+1\right)$.

3. *Let* T_0, S_0 *be as* (C1)*, and let* $(\tilde{T},\tilde{S})$ *be a global strong solution to* $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$*, as established in Theorem 1. If the following condition is satisfied by the problem parameters:*

$$u_I^{max}\Gamma < \frac{\min\{|D_1|,|D_2|\}}{\max\{\alpha,\beta\}|\Omega|^{1/2}} \frac{\min\{\kappa_x,\kappa_y,\kappa_z\}\min\{2\pi\kappa_x,2\pi\kappa_y,\frac{g_A^T}{2},\frac{\kappa_z}{2}\}}{4\max\{C_5^T,C_5^S\}}; \quad u_I^{max} = \max_{\boldsymbol{r}}|\boldsymbol{u}_I(\boldsymbol{r})|,$$

where C_5^T, C_5^S *are given in equation* (C30)*, then* (T,S) *converges to a unique steady-state solution as* $t\to\infty$.

Proof:

(ii) We start by proving the second point, establishing the *a-priori* bounds (C29). To this end, assume (T,S) is a weak solution to $\left(P^{\mathrm{SS}}\right)$. Since $(T,S)\in H^1(\Omega)\times\dot{H}_1(\Omega)$, then each can be used as a test function in equation (C28). Thus, by using the divergence theorem, Cauchy-Schwarz, Young's inequality and Lemma 3 one can immediately conclude the equations for T in (C29) and (C30). Similarly, using Lemmas 4 and 5 instead of Lemma 3, one obtains the equations for S in (C29) and (C30).

(i) Next, we prove that the nonlinear steady-state problem $\left(P^{\mathrm{SS}}\right)$ has a weak solution.

Step 1. Induced linear problem.

Let $\bar{T},\bar{S}\in L^2(\Omega)$ be given. We introduce the following linear stationary problem induced by $\left(P^{\mathrm{SS}}\right)$:

$$\left(P^{\mathrm{ss,lin}}_{\bar{T},\bar{S}}\right)\begin{cases} (P1^{\mathrm{ss,lin}}) & -\boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}T(\boldsymbol{r})) + (\bar{\boldsymbol{u}}(\boldsymbol{r};\langle\bar{T}\rangle,\langle\bar{S}\rangle)\cdot\boldsymbol{\nabla})T(\boldsymbol{r}) = f_T(\boldsymbol{r}), & \boldsymbol{r}\in\Omega \\ (P2^{\mathrm{ss,lin}}) & -\boldsymbol{\nabla}\cdot(\boldsymbol{\kappa}\boldsymbol{\nabla}S(\boldsymbol{r})) + (\bar{\boldsymbol{u}}(\boldsymbol{r};\langle\bar{T}\rangle,\langle\bar{S}\rangle)\cdot\boldsymbol{\nabla})S(\boldsymbol{r}) = f_S(\boldsymbol{r}), & \boldsymbol{r}\in\Omega \\ (P3^{\mathrm{ss,lin}}) & a_I(\langle\bar{T}\rangle,\langle\bar{S}\rangle) = \Gamma\left(-\alpha(\langle\bar{T}\rangle_2 - \langle\bar{T}\rangle_1) + \beta(\langle\bar{S}\rangle_2 - \langle\bar{S}\rangle_1)\right), & \\ (P4^{\mathrm{ss,lin}}) & \bar{\boldsymbol{u}}(\boldsymbol{r};\langle\bar{T}\rangle,\langle\bar{S}\rangle) = a_E\boldsymbol{u}_E(\boldsymbol{r}) + a_I(\langle\bar{T}\rangle,\langle\bar{S}\rangle)\boldsymbol{u}_I(\boldsymbol{r}), & \boldsymbol{r}\in\Omega \\ (P5^{\mathrm{ss,lin}}) & (\boldsymbol{\kappa}\boldsymbol{\nabla}T(q))\cdot\hat{\boldsymbol{n}}(q) = \begin{cases} g_A^T(T^*(x,y)-T(q)) & \text{if } q\in\sigma_1 \\ 0 & \text{else} \end{cases}, & q\in\partial\Omega \\ (P6^{\mathrm{ss,lin}}) & (\boldsymbol{\kappa}\boldsymbol{\nabla}S(q))\cdot\hat{\boldsymbol{n}}(q) = \begin{cases} g_A^S S^*(x,y) & \text{if } q\in\sigma_1 \\ 0 & \text{else} \end{cases}, & q\in\partial\Omega. \end{cases} \tag{C31}$$

We define the bilinear forms:

$$\begin{aligned} B_T &: H^1(\Omega)\times H^1(\Omega)\to\mathbb{R}, \quad B_T(T,\varphi) = \int_\Omega\left((\boldsymbol{\kappa}\boldsymbol{\nabla}T)\cdot(\boldsymbol{\nabla}\varphi) + ((\bar{\boldsymbol{u}}\cdot\boldsymbol{\nabla})T)\varphi\right)dV + \int_{\sigma_1} g_A^T T\varphi dxdy, \\ B_S &: \dot{H}^1(\Omega)\times\dot{H}^1(\Omega)\to\mathbb{R}, \quad B_S(T,\varphi) = \int_\Omega\left((\boldsymbol{\kappa}\boldsymbol{\nabla}S)\cdot(\boldsymbol{\nabla}\varphi) + ((\bar{\boldsymbol{u}}\cdot\boldsymbol{\nabla})S)\varphi\right)dV, \end{aligned}$$

and the linear functionals:

$$\begin{aligned} l_T &: H^1(\Omega)\to\mathbb{R}, \quad l_T(\varphi) = \int_\Omega f_T\varphi dV + \int_{\sigma_1} g_A^T T^*\varphi dxdy, \\ l_S &: \dot{H}^1(\Omega)\to\mathbb{R}, \quad l_S(\varphi) = \int_\Omega f_S\varphi dV + \int_{\sigma_1} g_A^S S^*\varphi dxdy. \end{aligned}$$

We define a weak solution to $\left(P^{\text{ss,lin}}_{\bar{T},\bar{S}}\right)$ to be $(T,S)\in H^1(\Omega)\times\dot{H}^1(\Omega)$ such that

$$B_T(T,\varphi)=l_T(\varphi),\quad B_S(S,\psi)=l_S(\psi) \tag{C32}$$

for every test function $\varphi\in H^1(\Omega),\psi\in\dot{H}^1(\Omega)$.

Next, we show that the linear problem has a unique weak solution, employing the Lax-Milgram theorem, see, e.g., Evans [16], chapter 6.2. To this end, we need to check that the conditions of the Lax-Milgram theorem are valid, namely the boundedness of the bilinear forms B_T and B_S and their coercivity, as well as the boundedness of the linear functionals l_T and l_S.

Using the Cauchy-Schwarz inequality and the trace theorem, we have

$$\begin{aligned}|B_T(\rho,\varphi)|&=\left|\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}\rho-\bar{\boldsymbol{u}}\rho)\cdot(\boldsymbol{\nabla}\varphi)dV+\int_{\sigma_1}g_A^T\rho\varphi dxdy\right|\\&\le\left(\left(\int_\Omega|\boldsymbol{\kappa}\boldsymbol{\nabla}\rho|^2\right)^{1/2}+\int_\Omega\left(|\bar{\boldsymbol{u}}\rho|^2\right)^{1/2}\right)\left(\int_\Omega|\boldsymbol{\nabla}\varphi|^2\right)^{1/2}+g_A^T\left(\int_{\sigma_1}\rho^2\right)^{1/2}\left(\int_{\sigma_1}\varphi^2\right)^{1/2}\\&\le M_T\left\|\rho\right\|_{H^1(\Omega)}\left\|\varphi\right\|_{L^2(\Omega)}+g_A^T\left\|\rho\right\|_{L^2(\partial\Omega)}\left\|\varphi\right\|_{L^2(\partial\Omega)}\le\alpha_T\left\|\rho\right\|_{H^1(\Omega)}\left\|\varphi\right\|_{H^1(\Omega)},\end{aligned}$$

where $M_T>0$ exists because $\boldsymbol{\kappa}$ is constant and $\bar{\boldsymbol{u}}$ is bounded by construction (equation (C1)), and $\alpha_T>0$. Using similar arguments,

$$\begin{aligned}|B_S(\rho,\varphi)|&=\left|\int_\Omega(\boldsymbol{\kappa}\boldsymbol{\nabla}\rho-\bar{\boldsymbol{u}}\rho)\cdot(\boldsymbol{\nabla}\varphi)dV\right|\le\left(\left(\int_\Omega|\boldsymbol{\kappa}\boldsymbol{\nabla}\rho|^2\right)^{1/2}+\int_\Omega\left(|\bar{\boldsymbol{u}}\rho|^2\right)^{1/2}\right)\left(\int_\Omega|\nabla\varphi|^2\right)^{1/2}\\&\le M_S\left\|\rho\right\|_{H^1(\Omega)}\left\|\varphi\right\|_{L^2(\Omega)}\le\alpha_S\left\|\rho\right\|_{H^1(\Omega)}\left\|\varphi\right\|_{H^1(\Omega)}.\end{aligned}$$

Next, to show coercivity, we use the divergence theorem and Lemma 3 to deduce

$$B_T(\varphi,\varphi)=\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}\varphi\right\|^2_{L_2(\Omega)}+g_A^T\tilde{\|}\varphi\tilde{\|}^2\ge\beta_T\left\|\varphi\right\|^2_{H^1(\Omega)}$$

where $\beta_T=\frac{\min\{\lambda,\kappa_{min}\}}{2}$. Moreover, using Lemma 5, we conclude

$$B_S(\varphi,\varphi)=\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}\varphi\right\|^2_{L_2(\Omega)}\ge\beta_S\left\|\varphi\right\|^2_{H^1(\Omega)} \tag{C33}$$

where $\beta_S=\frac{\min\{1,\kappa_{min}\}}{1+1/(\pi\kappa_{min})}$. Finally, to show boundedness of l_T and l_S, one uses Cauchy-Schwarz and Lemmas 3 and 5 as above.

Step 2. Existence of a steady state solution as a fixed point of a continuous operator.

We denote the function $Q:L^2\times\dot{L}^2\to L^2\times\dot{L}^2$ as follows: $Q((\bar{T},\bar{S}))=(T,S)$ where (T,S) is the solution of $\left(P^{\text{ss,lin}}_{\bar{T},\bar{S}}\right)$. Notice that a fixed point of the mapping Q is exactly a steady-state solution of the nonlinear stationary problem $\left(P^{\text{ss}}\right)$. Next, we show that the mapping Q indeed has a fixed point, employing the Schauder-Tychonof fixed point theorem (see [63], Corollary 2.13). Thus, in order to conclude the proof we have left only to check that the conditions of the Schauder-Tychonof fixed point theorem are valid, namely that Q is a continuous operator from a compact, convex subset of L^2 to itself.

Indeed, let $(\bar{T},\bar{S})$, (T,S) satisfy $Q((\bar{T},\bar{S}))=(T,S)$. Then (T,S) satisfy the bounds from equation (C29). Therefore, (T,S) is inside the closed ball in $H^1(\Omega)\times\dot{H}^1(\Omega)$ defined by these bounds, and this closed ball is compact in $L^2(\Omega)\times\dot{L}^2(\Omega)$ by the Rellich-Kondrachov Lemma [15].

To prove continuity of Q in the sense of L^2, let $\bar{T}_i,\bar{S}_i,T_i,S_i$ satisfy $Q((\bar{T}_i,\bar{S}_i))=(T_i,S_i)$ for $i=1,2$. Then T_i,S_i satisfy $B_T(T_i,\varphi)=l_T(\varphi)$, $B_S(S_i,\psi)=l_S(\psi)$, respectively. Using assumptions (a4), (a5) in (C1), and subtracting equation (C32) for T_1 from equation (C32) for T_2, both with the same test function $\varphi=T_2-T_1\equiv\delta T$, we deduce

$$\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}\delta T\right\|_{L^2(\Omega)}+\int_\Omega((\delta\bar{\boldsymbol{u}}\cdot\nabla)T_1)\delta TdV+g_A^T\tilde{\|}\delta T\tilde{\|}^2=0$$

where $\bar{\boldsymbol{u}}_i=\bar{\boldsymbol{u}}(\boldsymbol{r};\langle\bar{T}_i\rangle,\langle\bar{S}_i\rangle)$ is given by equation $(P4^{\text{ss,lin}})$ in for $i=1,2$ in (C31), $\delta\bar{\boldsymbol{u}}=\bar{\boldsymbol{u}}_2-\bar{\boldsymbol{u}}_2$. Therefore, using Lemma 3 and equation (C20), and defining $\delta\bar{T}\equiv\bar{T}_2-\bar{T}_1$, $\delta\bar{S}\equiv\bar{S}_2-\bar{S}_1$, we obtain

$$\lambda\left\|\delta T\right\|^2_{L^2(\Omega)}\le b\int_\Omega(|\delta\bar{T}|+|\delta\bar{S}|)dV\int_\Omega|\boldsymbol{\nabla}T_1||\delta T|dV$$

where $b > 0$ is some constant. Using Cauchy-Schwarz, equation (C29) and Young's inequality, we deduce

$$\|\delta T\|^2_{L^2(\Omega)} \leq \tilde{b}\left(\left\|\delta\bar{T}\right\|_{L^2(\Omega)} + \left\|\delta\bar{S}\right\|_{L^2(\Omega)}\right)^2 \tag{C34}$$

where $\tilde{b} > 0$ is some constant. We repeat similar steps for S, using Lemma 5 and defining $\delta S \equiv S_2 - S_1$, to obtain

$$\|\delta S\|^2_{L^2(\Omega)} \leq \tilde{c}\left(\left\|\delta\bar{T}\right\|_{L^2(\Omega)} + \left\|\delta\bar{S}\right\|_{L^2(\Omega)}\right)^2 \tag{C35}$$

where $\tilde{c} > 0$ is some constant. Equations (C34) and (C35) prove continuity of Q in the sense of $L^2(\Omega)$ since small $\left\|\delta\bar{T}\right\|_{L^2(\Omega)}$, $\left\|\delta\bar{S}\right\|_{L^2(\Omega)}$ imply small $\|\delta T\|_{L^2(\Omega)}$, $\|\delta S\|_{L^2(\Omega)}$.

(iii) To conclude the proof, we derive a condition on the parameters under which solutions to the time-dependent problem converge to a unique steady state solution. Let T_0, S_0 be as (C1), and let $(\tilde{T}, \tilde{S})$ be a global strong solution to $\left(P_{T_0,S_0,T^*,S^*,f_T,f_S}\right)$, as established in Theorem 1. Further, let (T, S) be a steady state solution to $\left(P_{T,S,T^*,S^*,f_T,f_S}\right)$ as established above. By Remark 6, using the analogous equation to (C10) for T and $\tilde{T}$ both with $\delta T \equiv \tilde{T} - T$ as a test function, we deduce

$$\frac{1}{2}\frac{d}{dt}\|\delta T\|^2 + g_A^T\|\delta T\|^2 + \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}\delta T\right\|^2 \leq \int_\Omega |\delta T||\delta u||\boldsymbol{\nabla} T|.$$

By Lemma 3, equation (C20) and the Cauchy-Schwarz inequality, we obtain

$$\frac{1}{2}\frac{d}{dt}\|\delta T\|^2 + \lambda\|\delta T\|^2 \leq b\frac{C_5^T}{\kappa_{min}}(\|\delta T\| + \|\delta S\|)\|\delta T\|$$

where $\delta S \equiv \tilde{S} - S$ and $b \equiv u_I^{\max}\Gamma\frac{\max\{\alpha,\beta\}}{\min\{|D_1|,|D_2|\}}|\Omega|^{1/2}$. Following similar arguments for S,

$$\frac{1}{2}\frac{d}{dt}\|\delta S\|^2 + \pi\kappa_{min}\|\delta S\|^2 \leq b\frac{C_5^S}{\kappa_{min}}(\|\delta T\| + \|\delta S\|)\|\delta S\|.$$

Hence, by defining $\tilde{\lambda} = \min\{2\pi\kappa_{min}, \frac{g_A^T}{2}, \frac{\kappa_z}{2}\}$, $c = 4\frac{b}{\kappa_{min}}\max\{C_5^T, C_5^S\}$, and $\eta(t) = \|\delta T\|^2(t) + \|\delta S\|^2(t)$, we obtain

$$\frac{d}{dt}\eta + \tilde{\lambda}\eta \leq \frac{c}{2}(\|\delta T\| + \|\delta S\|)^2 \leq c\eta.$$

This allows employing Gronwall's inequality to obtain the following bound:

$$\eta(t) \leq \eta(0)e^{-(\tilde{\lambda}-c)t}.$$

Thus, if $\tilde{\lambda} > c$, i.e.

$$u_I^{\max}\Gamma < \frac{\min\{|D_1|,|D_2|\}}{\max\{\alpha,\beta\}|\Omega|^{1/2}}\frac{\kappa_{min}\min\{4\pi\kappa_{min}, g_A^T, \kappa_z\}}{4\max\{C_5^T, C_5^S\}},$$

then $(\tilde{T}, \tilde{S})$ converges to the steady-state solution (T, S) at a rate of at least $1/(\tilde{\lambda} - c)$, independent of the initial conditions T_0, S_0. Thus, in this case, the steady-state solution is unique. □

4. Proof of useful Lemmas

a. *Proof of Lemma 3 (Poincaré inequality)*

Using the Cauchy-Schwarz inequality and Young's inequality,

$$\begin{aligned}
T^2(x,y,z) &= T^2(x,y,1) + \int_1^z \partial_z T^2(x,y,z')dz' = T^2(x,y,1) + 2\int_1^z T\partial_z T(x,y,z')dz' \\
&\leq T^2(x,y,1) + 2\int_1^z |T|\left|\frac{\partial T}{\partial z}\right|dz' \leq T^2(x,y,1) + \left(\int_1^z |T|^2 dz'\right)^{1/2}\left(\int_1^z 4\left|\frac{\partial T}{\partial z}\right|^2 dz'\right)^{1/2} \\
&\leq T^2(x,y,1) + \frac{1}{2}\int_1^0 |T|^2 dz + 2\int_1^0\left|\frac{\partial T}{\partial z}\right|^2 dz'.
\end{aligned}$$

By integrating over the domain Ω, we obtain $\|T\|^2 \leq 2\tilde{\|}T\tilde{\|}^2 + 4\|\partial_z T\|^2$, therefore

$$\lambda\|T\|^2 \leq 2\lambda\tilde{\|}T\tilde{\|}^2 + 4\lambda\|\partial_z T\|^2 \leq \frac{g_A^T}{2}\tilde{\|}T\tilde{\|}^2 + \kappa_z\|\partial_z T\|^2 \leq \frac{g_A^T}{2}\tilde{\|}T\tilde{\|}^2 + \left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}T\right\|.$$

□

b. Proof of Lemma 4

Using the Cauchy-Schwarz inequality and Young's inequality,

$$\begin{aligned} S^2(x,y,1) &= S^2(x,y,z) - 2\int_1^z S\partial_z S(x,y,z')dz' \leq S^2(x,y,z) + 2\int_1^z |S|\,|\partial_z S|\,dz' \\ &\leq S^2(x,y,z) + \frac{1}{\epsilon}\int_1^0 |S|^2 dz' + \epsilon\int_1^0 |\partial_z S|^2 dz'. \end{aligned}$$

Integrating over the domain Ω, we deduce

$$\tilde{\|}S\tilde{\|}^2 \leq \left(1+\frac{1}{\epsilon}\right)\|S\|^2 + \frac{\epsilon}{\kappa_{min}}\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S\right\|^2.$$

□

c. Proof of Lemma 5

Note that the inequality $\|S - \langle S\rangle_\Omega\|^2 \leq c\|\boldsymbol{\nabla}S\|^2$ is the Poincaré-Wirtinger inequality for $p = 2$, where c is a constant determined only by the domain and p. In fact, since Ω is a smooth bounded cube with side lengths 1, we can calculate the constant: it equals $1/\lambda_1$, where λ_1 is the smallest eigenvalue of minus the Laplacian, solving $-\boldsymbol{\nabla}^2 S = \lambda_1 S$, and equals $\pi/\max\{L_x, L_y, L_z\}$. In this simple case, $\lambda_1 = \pi$, therefore $c = 1/\pi$, and

$$\|S - \langle S\rangle_\Omega\|^2 \leq \frac{1}{\pi\kappa_{min}}\left\|\boldsymbol{\kappa}^{1/2}\boldsymbol{\nabla}S\right\|^2. \tag{C36}$$

□

[1] Aharon, R., Rom-Kedar, V. & Gildor, H. 2012 When complexity leads to simplicity: Ocean surface mixing simplified by vertical convection. *Phys. Fluids* **24** (5), 056603.

[2] Aref, H. 1984 Stirring by chaotic advection. *J. Fluid Mech.* **143**, 1–21.

[3] Aref, H. et al. 2017 Frontiers of chaotic advection. *Rev. Mod. Phys.* **89** (2), 025007.

[4] Ashkenazy, Y. & Tziperman, E. 2007 A wind-induced thermohaline circulation hysteresis and millennial variability regimes. *J. Phys. Oceanogr.* **37** (10), 2446–2457.

[5] Barkan, R., Winters, K. B. & Llewellyn Smith, S. G. 2013 Rotating horizontal convection. *J. Fluid Mech.* **723**, 556–586.

[6] Brett, G. J., Pratt, L., Rypina, I. & Wang, P. 2019 Competition between chaotic advection and diffusion: Stirring and mixing in a 3-d eddy model. *Nonlinear Proc. Geog.* **26** (2).

[7] Buckley, M. W. & Marshall, J. 2016 Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review. *Rev. Geophys.* **54** (1), 5–63.

[8] Butler, E. D., Oliver, K. I. C., Hirschi, J. M. & Mecking, J. V. 2016 Reconstructing global overturning from meridional density gradients. *Clim. Dyn.* **46** (7-8), 2593–2610.

[9] Calmanti, S., Artale, V. & Sutera, A. 2006 North Atlantic MOC variability and the Mediterranean Outflow: A box-model study. *Tellus A* **58** (3), 416–423.

[10] Carton, J. A. & Giese, B. S. 2008 A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Mon. Weather Rev.* **136** (8), 2999–3017.

[11] Cessi, P. 1994 A simple box model of stochastically forced thermohaline flow. *J. Phys. Oceanogr.* **24** (9), 1911–1920.

[12] Constantin, P. & Foias, C. 1988 *Navier-Stokes equations*. University of Chicago Press.

[13] Courant, R., Friedrichs, K. & Lewy, H. 1967 On the partial difference equations of mathematical physics. *IBM J. Res. Dev.* **11** (2), 215–234.

[14] Drévillon, M. et al. 2008 The GODAE/Mercator-Ocean global ocean forecasting system: Results, applications and prospects. *J. Operational Oceanogr.* **1** (1), 51–57.

[15] Edmunds, D. E. & Evans, W. D. 1987 *Spectral theory and differential operators*. Oxford University Press.

[16] Evans, L. C. 2010 *Partial differential equations*, , vol. 19. American Mathematical Society.

[17] Ferreira, D. et al. 2018 Atlantic-Pacific asymmetry in deep water formation. *Annu. Rev. Earth Planet Sci.* .

[18] Frankignoul, C., de Coëtlogon, G., Joyce, T. M. & Dong, S. 2001 Gulf Stream variability and ocean–atmosphere interactions. *J. Phys. Oceanogr.* **31** (12), 3516–3529.

[19] Gargett, A. E. 1984 Vertical eddy diffusivity in the ocean interior. *J. Mar. Res.* **42** (2), 359–393.

[20] Gent, P. R. & McWilliams, J. C. 1990 Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.* **20** (1), 150–155.

[21] Ghil, M. 2017 The wind-driven ocean circulation: Applying dynamical systems theory to a climate problem. *Discrete Contin. Dyn. Syst.* **37** (1), 189–228.

[22] Gildor, H. & Tziperman, E. 2001 A sea ice climate switch mechanism for the 100-kyr glacial cycles. *J. Geophys. Res. Oceans* **106** (C5), 9117–9133.

[23] Gill, A. E. 2016 *Atmosphere – Ocean Dynamics*, chap. A3. Elsevier.

[24] Gough, W. A. 1997 Isopycnal mixing and convective adjustment in an ocean general circulation model. *Atmosphere-Ocean* **35** (4), 495–511.

[25] Griffies, S. M. & Tziperman, E. 1995 A linear thermohaline oscillator driven by stochastic atmospheric forcing. *J. Climate* **8** (10), 2440–2453.

[26] Hadamard, M. I. 1902 On problems in partial derivatives, and their physical significance. *Princeton University Bulletin* **13** (49-52), 28.

[27] Haney, R. L. 1971 Surface thermal boundary condition for ocean circulation models. *J. Phys. Oceanogr.* **1** (4), 241–248.

[28] Huang, R. X., Luyten, J. R. & Stommel, H. M. 1992 Multiple equilibrium states in combined thermal and saline circulation. *J. Phys. Oceanogr.* **22** (3), 231–246.

[29] Hughes, G. O. & Griffiths, R. W. 2008 Horizontal convection. *Annu. Rev. Fluid Mech.* **40**, 185–208.

[30] Ivanovic, R. F., Valdes, P. J., Gregoire, L., Flecker, R. & Gutjahr, M. 2014 Sensitivity of modern climate to the presence, strength and salinity of Mediterranean-Atlantic exchange in a global general circulation model. *Climate Dyn.* **42** (3-4), 859–877.

[31] Johnson, H. L., Cessi, P., Marshall, D. P., Schloesser, F. & Spall, M. A. 2019 Recent contributions of theory to our understanding of the Atlantic Meridional Overturning Circulation. *J. Geophys. Res.* **124** (8), 5376–5399.

[32] Koshel, K. V. & Prants, S. V. 2006 Chaotic advection in the ocean. *Phys-Usp* **49** (11), 1151.

[33] Ladyzhenskaya, O. A. 1985 *The Boundary Value Problems of Mathematical Physics*, , vol. 49. Springer Science & Business Media.

[34] Li, J., Chen, L. & Surulescu, C. 2019 Global existence, asymptotic behavior, and pattern formation driven by the parametrization of a nonlocal Fisher – KPP problem. *arXiv preprint arXiv:1909.07934* .

[35] Li, J. & Titi, E. S. 2018 *Recent Advances Concerning Certain Class of Geophysical Flows*, pp. 933–971. Cham: Springer International Publishing.

[36] Lozier, M. S. & Stewart, N. M. 2008 On the temporally varying northward penetration of Mediterranean Overflow Water and eastward penetration of Labrador Sea water. *J. Phys. Oceanogr.* **38** (9), 2097–2103.

[37] Majda, A. J. & Kramer, P. R. 1999 Simplified models for turbulent diffusion: Theory, numerical modelling, and physical phenomena. *Phys. Rep.* **314** (4-5), 237–574.

[38] Marshall, J. & Schott, F. 1999 Open-ocean convection: Observations, theory, and models. *Rev. Geophys.* **37** (1), 1–64.

[39] McCartney, M. S. & Mauritzen, C. 2001 On the origin of the warm inflow to the Nordic Seas. *Prog. Oceanogr.* **51** (1), 125–214.

[40] Mogilner, A. & Edelstein-Keshet, L. 1999 A non-local model for a swarm. *J. Math. Biol.* **38** (6), 534–570.

[41] Mullarney, J. C., Griffiths, R. W. & Hughes, G. O. 2004 Convection driven by differential heating at a horizontal boundary. *J. Fluid Mech.* **516**, 181.

[42] Mullarney, J. C., Griffiths, R. W. & Hughes, G. O. 2007 The role of freshwater fluxes in the thermohaline circulation: Insights from a laboratory analogue. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **54** (1), 1–21.

[43] Nittka, R. 2014 Inhomogeneous parabolic Neumann problems. *Czech. Math. J.* **64** (3), 703–742.

[44] Pasquero, C. & Tziperman, E. 2004 Effects of a wind-driven gyre on thermohaline circulation variability. *J. Phys. Oceanogr.* **34** (4), 805–816.

[45] Rahmstorf, S. 2005 Thermohaline circulation hysteresis: A model intercomparison. *Geophys. Res. Lett.* **32** (23).

[46] Reid, J. L. 1979 On the contribution of the Mediterranean Sea outflow to the Norwegian-Greenland Sea. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **26** (11), 1199–1223.

[47] Sheard, G. J. & King, M. P. 2011 Horizontal convection: Effect of aspect ratio on Rayleigh number scaling and stability. *Appl. Math. Model.* **35** (4), 1647–1655.

[48] Shishkina, O. 2017 Mean flow structure in horizontal convection. *J. Fluid Mech.* **812**, 525.

[49] Sijp, W. P., Gregory, J. M., Tailleux, R. & Spence, P. 2012 The key role of the western boundary in linking the AMOC strength to the north–south pressure gradient. *J. Phys. Oceanogr.* **42** (4), 628–643.

[50] Stommel, H. M. 1961 Thermohaline convection with two stable regimes of flow. *Tellus* **13** (2), 224–230.

[51] Sweby, P. K. 1984 High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM J. Numer. Anal.* **21** (5), 995–1011.

[52] Tailleux, R. & Rouleau, L. 2010 The effect of mechanical stirring on horizontal convection. *Tellus A* **62** (2), 138–153.

[53] Temam, R. 2012 *Infinite-dimensional dynamical systems in mechanics and physics*, , vol. 68. Springer Science & Business Media.

[54] Temam, R. & Ziane, M. 2005 Some mathematical problems in geophysical fluid dynamics. In *Handbook of mathematical fluid dynamics*, , vol. 3, pp. 535–658. Elsevier.

[55] Tziperman, E. 1997 Inherently unstable climate behaviour due to weak thermohaline ocean circulation. *Nature* **386** (6625), 592–595.

[56] Tziperman, E., Toggweiler, R. J., Bryan, K. & Feliks, Y. 1994 Instability of the thermohaline circulation with respect to mixed boundary conditions: Is it really a problem for realistic models? *J. Phys. Oceanogr.* **24** (2), 217–232.

[57] Vallis, G. K. 2017 *Atmospheric and Oceanic Fluid Dynamics*. Cambridge University Press.
[58] Weijer, W., Cheng, W., Drijfhout, S. S., Fedorov, A. V., Hu, A., Jackson, L. C., Liu, W., McDonagh, E. L., Mecking, J. V. & Zhang, J. 2019 Stability of the Atlantic Meridional Overturning Circulation: A review and synthesis. *J. Geophys. Res. Oceans* **124** (8), 5336–5375.
[59] Yang, H. 1996 The subtropical/subpolar gyre exchange in the presence of annually migrating wind and a meandering jet: Water mass exchange. *J. Phys. Oceanogr.* **26** (1), 115–130.
[60] Yang, H. & Liu, Z. 1994 Chaotic transport in a double gyre ocean. *Geophys. Res. Lett.* **21** (7), 545–548.
[61] Yang, H. & Liu, Z. 1997 The three-dimensional chaotic transport and the great ocean barrier. *J. Phys. Oceanogr.* **27** (7), 1258–1273.
[62] Yosida, K. 1971 *Functional Analysis*, , vol. 123. Springer-Verlag.
[63] Zeidler, E. 1986 *Nonlinear functional analysis and its applications: I: Fixed-point theorems*. Springer-Verlag.