Draft version July 19, 2021
Typeset using LaTeX twocolumn style in AASTeX631

Convective penetration probably parameterizes convective overshoot

Evan H. Anders,[1] Adam S. Jermyn,[2] Daniel Lecoanet,[1,3] and Benjamin P. Brown[4]

[1] CIERA, Northwestern University, Evanston IL 60201, USA
[2] Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA
[3] Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston IL 60208, USA
[4] Department Astrophysical and Planetary Sciences & LASP, University of Colorado, Boulder, CO 80309, USA

(Received; Revised; Accepted; Published)

Submitted to ApJ

ABSTRACT

Most stars host convection zones in which heat is transported directly by fluid motion. Parameterizations like mixing length theory adequately describe convective flows in the bulk of these regions, but the behavior of convective boundaries is not well understood. Here we examine how convective motions mix entropy and in turn extend the size of the convection zone, a process referred to as convective penetration. We derive a "penetration parameter" $\mathcal{P}$ which compares how far the radiative gradient deviates from the adiabatic gradient on either side of the Schwarzschild convective boundary. Following Roxburgh (1989) and Zahn (1991), we construct a model of penetration based on an energy balance, and find that the extent of penetration is controlled by $\mathcal{P}$. We perform 3D numerical simulations with Dedalus using a simplified Boussinesq model of stellar convection with a radiative flux and find good agreement with the derived theory. We evolve these simulations for thousands of overturn times; over these long timescales our simulations consistently develop large penetrative regions. In stellar contexts, we expect $\mathcal{P} \approx 1$ and in this regime our results suggest that convection zones may extend beyond the Schwarzschild boundary by up to ∼20-30% of a mixing length. We present a MESA solar model which employs our parameterization of convective penetration, and find that the base of the solar convection zone extends by (TODO X amount for some choice of parameters). We further discuss prospects for extending these results to more realistic stellar contexts.

Keywords: UAT keywords

1. INTRODUCTION

1.1. *Context*

Convection is a crucial mechanism for transporting heat in stars (Woosley et al. 2002; Hansen et al. 2004; Christensen-Dalsgaard 2021), and convective dynamics influence many poorly-understood stellar phenomena. For example, convection drives the magnetic dynamo of the Sun, leading to a whole host of emergent phenomena collectively known as solar activity (Brun & Browning 2017). Convection also mixes chemical elements in stars, which can modify observed surface abundances or inject additional fuel into their cores, thereby extending stellar lifetimes (Salaris & Cassisi 2017). Furthermore, convective motions excite waves, which can be observed and used to constrain the thermodynamic structure of stars (Aerts et al. 2010; Basu 2016). A complete and nuanced understanding of convection is therefore crucial for understanding stellar structure and evolution, and for connecting this understand to observations.

Corresponding author: Evan H. Anders
evan.anders@northwestern.edu

Despite decades of study, robust parameterizations for the mechanisms broadly referred to as "convective overshoot" remain elusive, and improved parameterizations could resolve many discrepancies between observations and structure models. In the stellar structure literature, "convective overshoot" refers to any convectively-driven mixing which occurs beyond the boundaries of the Ledoux-unstable zone. This mixing can influence, for example, observed surface lithium abundances in the Sun and solar-type stars, which align poorly with theoretical predictions (Pinsonneault 1997; Carlos et al. 2019; Dumont et al. 2021). Furthermore, modern spectroscopic observations suggest a lower solar metallicity than previously thought, and models computed with modern metallicity estimates and opacity tables have shallower convection zones than helioseismic observations suggest (Basu & Antia 2004; Bahcall et al. 2005; Bergemann & Serenelli 2014; Vinyoles et al. 2017; Asplund et al. 2021); modeling and observational discrepancies can be reduced with additional mixing below the convective boundary (Christensen-Dalsgaard et al. 2011).

Beyond the Sun, overshooting in massive stars with convective cores must be finely tuned as a function of stellar mass, again pointing to missing physics in our current parameterizations (Claret & Torres 2018; Jermyn et al. 2018; Viani & Basu 2020; Martinet et al. 2021; Pedersen et al. 2021). Since core convective overshoot increases the reservoir of fuel available for nuclear fusion at each stage in stellar evolution, improved models of core convective boundary mixing could have profound impacts on the post-main sequence evolution and remnant formation of massive stars (Farmer et al. 2019; Higgins & Vink 2020).

In order to ensure that models can be evolved on fast (human) timescales, 1D stellar evolution codes rely on simple parameterizations of convection (e.g., mixing length theory, Böhm-Vitense 1958) and convective overshoot (Shaviv & Salpeter 1973; Maeder 1975; Herwig 2000; Paxton et al. 2011, 2013, 2018, 2019). While some preliminary work has been done to couple 3D dynamical convective simulations with 1D stellar evolution codes (Jørgensen & Weiss 2019), these calculations are prohibitively expensive to perform at every timestep in a stellar evolution simulation. In order to resolve discrepancies between stellar evolution models and observations, a more complete and *parameterizeable* understanding of convective overshoot is required.

The broad category of "convective overshoot" in the stellar literature is an umbrella term for a few hydrodynamical processes (Zahn 1991; Brummell et al. 2002; Korre et al. 2019). The first process is confusingly also called "*convective overshoot*" in the fluid dynamics literature, and refers to motions which extend beyond the convective boundary but do not adjust the thermodynamic profiles. The second process is called "*convective penetration*," and refers to convective motions which mix the thermodynamic profiles beyond the convective boundary, thereby extending the adiabatically-mixed convective region. Many studies have also examined *entrainment*, the process by which convection zones expand by eroding composition gradients or modifying the radiative gradient (Meakin & Arnett 2007; Viallet et al. 2013; Cristini et al. 2017; Fuentes & Cumming 2020; Horst et al. 2021). The primary focus of this work is convective penetration.

Convective overshoot, penetration, and entrainment have been studied in the laboratory and through numerical simulations for decades, and the state of the field has been regularly reviewed (e.g., Marcus et al. 1983; Zahn 1991; Browning et al. 2004; Rogers et al. 2006; Viallet et al. 2015; Korre et al. 2019). Modern numerical experiments often examine the importance of the "stiffness" of a radiative-convective interface. The stiffness measures the relative stability of a radiative zone and an adjacent convection zone according to some measure like a dynamical frequency or characteristic entropy gradient. Experiments exhibiting extensive expansion of convection zones via entrainment have a long history (dating back to e.g., Musman 1968; Deardorff et al. 1969; Moore & Weiss 1973, and this process is often confusingly called "penetration"). Some recent studies in simplified Boussinesq setups exhibit stiffness-dependent convection zone expansion via entrainment (Couston et al. 2017; Toppaladoddi & Wettlaufer 2018); others find stiffness-dependent pure overshoot (Korre et al. 2019). While a link between stiffness and entrainment and overshoot processes has seemingly emerged in previous work, a mechanism for penetration remains elusive.

Studies which more closely model the conditions inside of stars have presented confusing results. Many studies have exhibited hints of penetrative convection – that is, mixing of the entropy gradient towards the adiabatic beyond the convective boundary (Hurlburt et al. 1986, 1994; Singh et al. 1995; Saikia et al. 2000; Browning et al. 2004; Rogers et al. 2006; Kitiashvili et al. 2016; Brun et al. 2017; Pratt et al. 2017; Higl et al. 2021). These simulations include simple plane-parallel studies of compressible convection as well as global 3D simulations which employ profiles from stellar structure codes as background states. Unfortunately, other fundamental studies in both Cartesian and spherical geometries have exhibited little to no dependence of penetration on the stiffness (Brummell et al. 2002; Rogers & Glatzmaier 2005). This host of experiments suggests that convective penetration is *probably* a process that happens in stars, but no clear model has emerged which explains whether convective penetration depends on stiffness or some other quantity.

There are hints in the literature that convective penetration may be dependent upon energy fluxes. Roxburgh

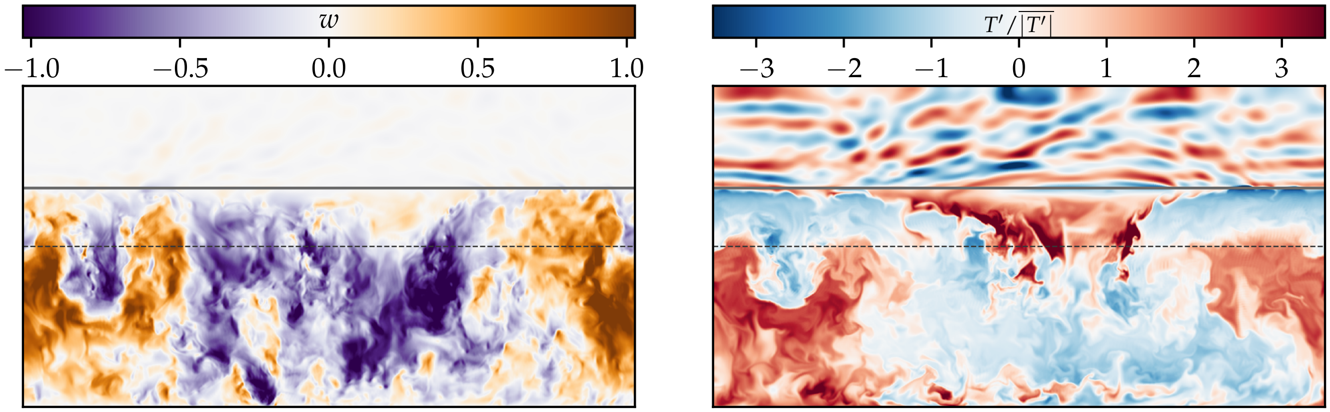

Figure 1. Vertical slice through a simulation with $\mathcal{R} = 6.4 \times 10^3$, $\mathcal{P}_D = 4$ and $\mathcal{S} = 10^3$ (see Sec. 4). The Schwarzschild boundary of the convection zone where $\nabla_{\rm ad} = \nabla_{\rm rad}$ is displayed as a dashed horizontal line. The top of the penetrative zone ($\delta_{0.1}$, see Sec. 4) is shown by a solid horizontal line. (Left) Vertical velocity is shown; orange convective upflows extend far past the Schwarzschild boundary of the convection zone but stop abruptly at the top of the penetration zone where ∇ departs from $\nabla_{\rm ad}$. (Right) Temperature fluctuations, normalized by their average magnitude at each height to clearly display all dynamical features.

(1978, 1989, 1992, 1998) derived an "integral constraint" from the energy equation and found that a spatial integral of the flux puts an upper limit on the size of a theoretical penetrative region. Zahn (1991) theorized that convective penetration should depend only on how steeply the radiative temperature gradient varies at the convective boundary. Following Zahn (1991)'s work, Rempel (2004) derived a semianalytic model and suggested that inconsistencies seen in simulations of penetrative dynamics can be explained by the magnitude of the fluxes or luminosities driving the simulations. Indeed, some simulations have tested this idea, and found that penetration lengths depend strongly on the input flux (Singh et al. 1998; Käpylä et al. 2007; Tian et al. 2009; Hotta 2017; Käpylä 2019). Furthermore, in the limit of low stiffness, the simulations of Hurlburt et al. (1994) and Rogers & Glatzmaier (2005) may agree with Zahn's theory (although at high stiffness they disagree). In light of these results, and the possible importance of energy fluxes, Roxburgh's integral constraint and Zahn's theory deserve to be revisited.

1.2. *Convective penetration & this study's findings*

Convective penetration is the process by which convective motions extend beyond the Schwarzschild-stable boundary and mix the entropy gradient to be nearly adiabatic.

> *In this paper, we present simulations which exhibit convective penetration.*

This process is phenomenologically described in Sec. 2.

In order to understand this phenomenon, we derive theoretical predictions for the size of the penetrative zone based on the ideas of Roxburgh (1989) and Zahn (1991).

> *We find that the extent of convective penetration depends strongly on the shape and magnitude of the radiative gradient near the convective boundary.*

Thus, the penetration length can be calculated using the radiative conductivity (or opacity) *profile* near the convective boundary.

We present these findings as follows. In Sec. 2, we present the central finding of this work: penetration zones in nonlinear convective simulations. In Sec. 3, we describe the equations used and derive a parameterized theory of convective penetration. In Sec. 4, we describe our simulation setup and parameters. In Sec. 5, we present the results of these simulations, with a particular focus on the height of the penetrative regions. In Sec. 6, we create and discuss a solar MESA model which uses this theory to determine the bottom of the solar convection zone. Finally, we discuss how future simulations can better constrain this theory in Sec. 7.

2. CENTRAL RESULT: CONVECTIVE PENETRATION

In Fig. 1, we display a snapshot of dynamics in an evolved simulation which exhibits convective penetration. The simulation domain is a 3D Cartesian box, and this figure shows a vertical slice through the center of the domain. In the left panel, we display the vertical velocity. We see that convective motions extend beyond the Schwarzschild boundary of the convection zone, which is denoted by a horizontal dashed grey line. These mo-

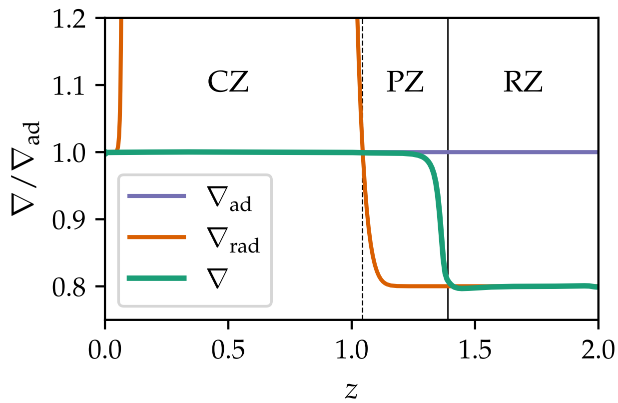

Figure 2. Horizontally- and temporally-averaged profiles of the thermodynamic gradients from the simulation in Fig. 1. We plot ∇ (green) compared to $\nabla_{\rm ad}$ (purple, a constant) and $\nabla_{\rm rad}$ (orange); note the extended penetration zone (PZ) where $\nabla \approx \nabla_{\rm ad} > \nabla_{\rm rad}$.

tions come to a halt at the top of a penetration zone, denoted by a solid horizontal line, where the temperature gradient departs from adiabatic towards the radiative gradient. In the right panel, we display temperature perturbations away from the time-evolving mean temperature profile. We see that warm upwellings in the Schwarzschild-unstable convection zone (below the dashed line) become cold upwellings in the penetration zone (above the dashed line), and these motions excite gravity waves in the stable radiative zone (above the solid line).

We further explore the simulation from Fig. 1 in Fig. 2 by displaying time- and horizontally-averaged 1D profiles of the thermodynamic gradient ∇ (defined in Sec. 3). The adiabatic gradient $\nabla_{\rm ad}$ is shown in purple and is a constant value in the simulation. The radiative gradient $\nabla_{\rm rad}$ is shown in orange. The domain exhibits a classical Schwarzshild-unstable convection zone (CZ) for $z \lesssim 1.04$ where $\nabla_{\rm rad} > \nabla_{\rm ad}$; the upper boundary of this region is denoted by a dashed vertical line. Above this point, $\nabla_{\rm rad} < \nabla_{\rm ad}$ and the domain would be considered stable by the Schwarzschild criterion. However, the evolved convective dynamics in Fig. 1 have raised $\nabla \rightarrow \nabla_{\rm ad}$ in an extended penetration zone (PZ) which extends from roughly $1.04 \leq z \leq 1.4$. Above this point, ∇ departs from $\nabla_{\rm ad}$, returning to $\nabla_{\rm rad}$ in a classical stable radiative zone (RZ).

Our goals in this paper are to understand how these PZs form, and to parameterize this effect so that it can be included in 1D stellar evolution calculations.

3. THEORY

In this section we derive a theoretical model of convective penetration by examining the energetics and energy fluxes in the convection and penetration zones. In Sec. 3.1, we describe our equations and problem setup and define the heat fluxes. In Sec. 3.2, we build a parameterized theory based on the kinetic energy (KE) equation. We find that excess KE in the convection zone can cause a penetration zone to form. By balancing the excess KE in the convection zone with a buoyancy-braking work term in the PZ, we are able to derive the size of the PZ. We find that a description of the size of a theoretical penetration zone does not depend on the often-considered stiffness, which measures the relative stability between the convection zone and an adjacent radiative zone.

3.1. *Equations & flux definitions*

Throughout this work, we will utilize a modified version of the incompressible Boussinesq equations,

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0 \tag{1}$$

$$\partial_t \boldsymbol{u} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} = -\frac{1}{\rho_0}\boldsymbol{\nabla} p + \frac{\rho_1}{\rho_0}\boldsymbol{g} + \nu\boldsymbol{\nabla}^2\boldsymbol{u} \tag{2}$$

$$\partial_t T + \boldsymbol{u}\cdot\boldsymbol{\nabla} T + w\nabla_{\rm ad} + \boldsymbol{\nabla}\cdot[-k\boldsymbol{\nabla}\overline{T}] = \chi\boldsymbol{\nabla}^2 T' + Q \tag{3}$$

$$\frac{\rho_1}{\rho_0} = -|\alpha| T. \tag{4}$$

Here, the density is decomposed into a constant background ρ_0 with fluctuations ρ_1 which appear only in the buoyancy force and depend on the temperature T and the coefficient of thermal expansion $\alpha = \partial \ln\rho/\partial T$. We define the velocity vector $\boldsymbol{u}$, the viscous diffusivity ν, the thermal diffusivity χ, the bulk internal heating Q, the adiabatic gradient $\nabla_{\rm ad}$, and a height-dependent thermal conductivity k. We will consider Cartesian coordinates (x, y, z) with a constant vertical gravity $\boldsymbol{g} = -g\hat{z}$. Throughout this work, we will represent horizontal averages with bars $(\overline{\cdot})$ and fluctuations away from those averages with primes $(')$. Thus, in Eqn. 3, $\overline{T}$ is the horizontally averaged temperature and T' are fluctuations away from that; both of these fields evolve in time according to Eqn. 3.

Assuming convection reaches a time-stationary state, the heat fluxes are found by horizontally-averaging then vertically integrating Eqn. 3 to find

$$\overline{F_{\rm tot}} = \overline{F_{\rm rad}} + \overline{F_{\rm conv}} = \int Q dz + F_{\rm bot}, \tag{5}$$

where $F_{\rm bot}$ is the flux carried at the bottom of the domain, and $\overline{F_{\rm tot}}$ is the total flux, which can vary in height due to the heating Q. The mean temperature profile $\overline{T}$ carries the radiative flux $\overline{F_{\rm rad}} = -k\boldsymbol{\nabla}\overline{T}$. We note that k

and $-\partial_z \overline{T}$ fully specify $\overline{F_{\rm rad}}$ and in turn the convective flux, $\overline{F_{\rm conv}} = \overline{F_{\rm tot}} - \overline{F_{\rm rad}}$. We define the temperature gradient and radiative temperature gradient

$$\nabla \equiv -\partial_z \overline{T} \qquad \nabla_{\rm rad} \equiv \frac{\overline{F_{\rm tot}}}{k}. \tag{6}$$

We have defined the ∇'s as positive quantities to align with stellar structure conventions and intuition. Marginal stability is achieved when $\nabla = \nabla_{\rm ad}$, which we take to be a constant. We note that the classical Schwarzschild boundary of the convection zone is the height $z = L_s$ at which $\nabla_{\rm rad} = \nabla_{\rm ad}$ and $\overline{F_{\rm conv}} = 0$.

The addition of a nonzero $\nabla_{\rm ad}$ to Eqn. 3 was derived by Spiegel & Veronis (1960) and utilized by e.g., Korre et al. (2019). In this work, we have decomposed the radiative diffusivity into a background portion ($\boldsymbol{\nabla}\cdot\overline{F_{\rm rad}}$) and a fluctuating portion ($\chi\boldsymbol{\nabla}^2 T'$); by doing so, we have introduced a height-dependent $\nabla_{\rm rad}$ to the equation set while preserving the diffusive behavior on fluctuations felt by classical Rayleigh-Bénard convection. Here, we will assume a model in which an unstable convection zone ($\nabla_{\rm rad} > \nabla_{\rm ad}$) sits below a stable radiative zone ($\nabla_{\rm rad} < \nabla_{\rm ad}$), but in this incompressible model where there is no density stratification to break the symmetry of upflows and downflows precisely the same arguments can be applied to the inverted problem.

3.2. *Kinetic energy & the dissipation-flux link*

Taking a dot product of the velocity and Eqn. 2 reveals the kinetic energy equation,

$$\frac{\partial \mathcal{K}}{\partial t} + \boldsymbol{\nabla}\cdot\mathcal{F} = \mathcal{B} - \Phi, \tag{7}$$

where we define the kinetic energy $\mathcal{K} \equiv |\boldsymbol{u}|^2/2$, the fluxes of kinetic energy $\mathcal{F} \equiv [\boldsymbol{u}(\mathcal{K} + p/\rho_0) - \nu\boldsymbol{u}\times\boldsymbol{\omega}]$, the buoyant energy generation rate $\mathcal{B} \equiv |\alpha| g w T'$, and the viscous dissipation rate $\Phi \equiv \nu|\boldsymbol{\omega}|^2$ where $\boldsymbol{\omega} = \boldsymbol{\nabla}\times\boldsymbol{u}$ is the vorticity and $|\boldsymbol{u}|^2 = \boldsymbol{u}\cdot\boldsymbol{u}$ & $|\boldsymbol{\omega}|^2 = \boldsymbol{\omega}\cdot\boldsymbol{\omega}$. We next take a horizontal- and time-average of Eqn. 7 (we absorb the time-average into the horizontal-average $\overline{\cdot}$ notation for simplicity). Assuming that $\overline{\mathcal{K}}$ reaches a statistically stationary state, convective motions satisfy

$$\frac{d\overline{\mathcal{F}}}{dz} = \overline{\mathcal{B}} - \overline{\Phi}. \tag{8}$$

Each profile in Eqn. 8 is shown in Fig. 3 for the simulation whose dynamics are displayed in Fig. ??. As in Fig. 2, the Schwarzschild CZ boundary is plotted as a dashed line, and the top of the PZ is plotted as a solid vertical line. In the top panel, we display $\overline{\mathcal{F}}$, neglecting the viscous terms which are only nonzero near the bottom boundary. We see that $\overline{\mathcal{F}}$ is zero at the bottom boundary (left edge of plot) and at the top of the PZ. In the bottom panel, we plot $\overline{\mathcal{B}}$ and $\overline{\Phi}$; we see that $\overline{\mathcal{B}}$ changes sign at the Schwarzschild CZ boundary, and that $\overline{\Phi}$ is positive-definite.

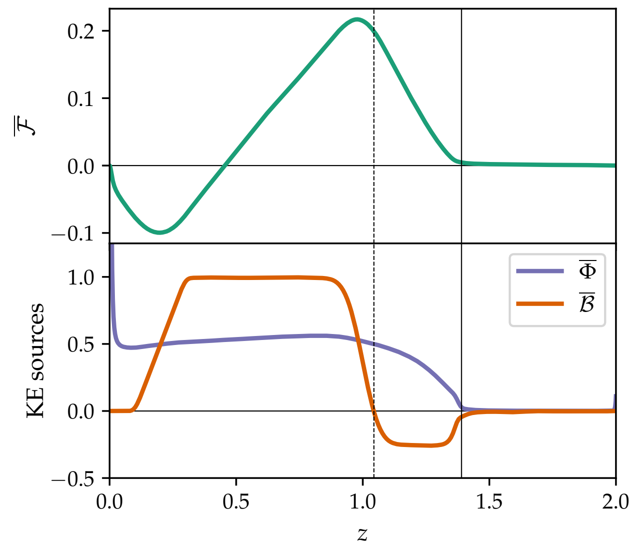

Figure 3. Temporally- and horizontally-averaged profiles in an evolved simulation from the kinetic energy equation (Eqn. 8). The vertical dashed line corresponds to the Schwarzschild CZ boundary, the vertical solid line corresponds to the top of the PZ. (upper) Kinetic energy fluxes $\overline{\mathcal{F}}$, which go to zero at the bottom boundary and the top of the PZ. (bottom) Source terms from Eqn. 8 normalized by the maximum of $\overline{\mathcal{B}}$ ($\overline{\mathcal{F}}$ in the upper panel is similarly normalized). The buoyancy source $\overline{\mathcal{B}}$ changes sign at the Schwarzschild boundary, and $\overline{\Phi}$ is positive-definite.

At the boundaries of the convecting region, $\overline{\mathcal{F}}$ is zero (Fig. 3, upper panel). We integrate Eqn. 8 vertically between these zeros to find

$$\int \overline{\mathcal{B}}\, dz = \int \overline{\Phi}\, dz. \tag{9}$$

Integral constraints of this form are the basis for a broad range of analyses in Boussinesq convection (see e.g., Ahlers et al. 2009; Goluskin 2016) and were considered in the context of penetrative stellar convection by Roxburgh (1989). Eqn. 9 is the straightforward statement that work by buoyancy on large scales must be balanced by viscous dissipation on small scales.

We break up the convecting region into a Schwarzschild-unstable "convection zone" (CZ) and a extended "penetration zone" (PZ); we assume that convective motions efficiently mix $\nabla \to \nabla_{\rm ad}$ in both the CZ and PZ. The buoyant energy generation is proportional to the convective flux, $\overline{\mathcal{B}} = |\alpha| g \overline{w T'} = |\alpha| g \overline{F_{\rm conv}}$. Thus, in the PZ where $\nabla_{\rm ad} > \nabla_{\rm rad}$ (see Fig. 3, middle panel), we know that $\overline{F_{\rm conv}} < 0$ and so $\overline{\mathcal{B}} < 0$ too (see the discussion around Eqn. 6). Breaking up Eqn. 9, we see that

$$\int_{\rm CZ} \overline{\mathcal{B}}\, dz \;=\; \int_{\rm CZ} \overline{\Phi}\, dz + \int_{\rm PZ} \overline{\Phi}\, dz + \int_{\rm PZ} (-\overline{\mathcal{B}})\, dz. \quad (10)$$

Eqn. 10 is arranged so that the (positive) buoyant engine of convection is on the left-hand side, and the (positive) sinks of work are on the RHS. If viscous dissipation in the CZ does not balance the buoyant generation of energy in the CZ, the kinetic energy of the convective flows grows, resulting in a penetrative region. This region grows with time until Eqn. 10 is satisfied. We see that the viscous dissipation and buoyancy breaking felt by flows in the PZ determine its size. We now parameterize the fraction of the buoyant engine consumed by CZ dissipation,

$$f \equiv \frac{\int_{\rm CZ} \overline{\Phi}\, dz}{\int_{\rm CZ} \overline{\mathcal{B}}\, dz}. \quad (11)$$

Under this parameterization, Eqn. 10 can be written

$$\frac{\int_{\rm PZ} (-\overline{\mathcal{B}})\, dz}{\int_{\rm CZ} \overline{\mathcal{B}}\, dz} + \frac{\int_{\rm PZ} \overline{\Phi}\, dz}{\int_{\rm CZ} \overline{\mathcal{B}}\, dz} = (1-f). \quad (12)$$

We will measure and report the values of f achieved in our simulations in this work. Eqn. 12 provides two limits on a hypothetical PZ:

1. In the limit that $f \to 0$, viscous dissipation is inefficient. Reasonably if we also assume that $\int_{\rm PZ} \overline{\Phi}\, dz \to 0$, Eqn. 12 states that the PZ must be large enough for its negative buoyant work to be equal in magnitude to the excess positive buoyant work of the CZ. This is the integral constraint on the maximum size of the PZ that Roxburgh (1989) derived.

2. In the limit that $f \to 1$, viscous dissipation efficiently counteracts the buoyancy work in the CZ. Per Eqn. 12, the positive-definite PZ terms must approach zero and no PZ develops in this limit. This is mathematically equivalent to standard boundary-driven convection experiments.

In general, we anticipate from the results of e.g., Currie & Browning (2017) that f is closer to 1 than 0, but its precise value must be measured from simulations. Indeed, we find that $f \gg 0$ but $f < 1$ in our simulations (see e.g., Fig. 3, bottom panel[1]). Our simulations produce typical values of $f \sim 0.7$.

Assuming that a PZ of height $\delta_{\rm p}$ develops above a CZ of depth $L_{\rm CZ}$, we model the PZ dissipation as

$$\int_{\rm PZ} \overline{\Phi}\, dz = \xi \frac{\delta_{\rm p}}{L_{\rm CZ}} \int_{\rm CZ} \overline{\Phi}\, dz = \xi \delta_{\rm p} \Phi_{\rm CZ}. \quad (13)$$

[1] the bulk dynamics suggest by eye $f \sim 0.5$, but due to e.g., the height dependence of $\overline{\mathcal{B}}$ in our simulations we measure $f \approx 0.74$.

Here $\Phi_{\rm CZ}$ is the volume-averaged dissipation rate in the CZ and ξ is a measurable parameter in $[0,1]$ that describes the shape of the dissipation profile as a function of height in the PZ. In words, we assume that $\overline{\Phi}(z = L_s) \approx \Phi_{\rm CZ}$ at the CZ-PZ boundary and that $\overline{\Phi}$ decreases with height in the PZ. The shape of $\overline{\Phi}$ determines ξ; a linear falloff gives $\xi = 1/2$, a quadratic falloff gives $\xi = 2/3$, and $\xi = 1$ assumes no falloff. With this parameterization, and $\overline{\mathcal{B}} \propto \overline{F_{\rm conv}}$, we rewrite Eqn. 12

$$-\frac{\int_{\rm PZ} \overline{F_{\rm conv}}\, dz}{\int_{\rm CZ} \overline{F_{\rm conv}}\, dz} + f\xi \frac{\delta_{\rm p}}{L_{\rm CZ}} = (1-f). \quad (14)$$

The fundamental result of this theory is Eqn. 14, which is a parameterized and generalized form of Roxburgh (1989)'s integral constraint. This equation is also reminiscent of Zahn (1991)'s theory, and says that the size of a PZ is set by the profile of $\nabla_{\rm rad}$ near the convective boundary. A parameterization like Eqn. 14 can be implemented in stellar structure codes and used to find the extent of penetration zones under the specification of f and ξ. The parameters f and ξ are measurables which can be constrained by direct numerical simulations, and we will measure their values in this work. In general, we expect that f and ξ should not change too drastically with other simulation parameters.

In order to derive a specific prediction for the PZ height, one must specify the vertical shape of $\overline{F_{\rm conv}}$. We will study two cases in this work, laid out below. In both of these cases, we define a nondimensional "Penetration Parameter" whose magnitude is set by the ratio of the convective flux slightly above and below the Schwarzschild convective boundary L_s (assuming $\nabla = \nabla_{\rm ad}$ in the CZ and PZ),

$$\mathcal{P} \equiv -\frac{\overline{F_{\rm conv}}_{\rm CZ}}{\overline{F_{\rm conv}}_{\rm PZ}}. \quad (15)$$

Since $F_{\rm conv} < 0$ in the PZ, the sign of $\mathcal{P}$ is positive. Intuitively, $\mathcal{P}$ describes which terms are important in Eqn. 12. When $\mathcal{P} \ll 1$, the buoyancy term dominates in the PZ and dissipation can be neglected there. When $\mathcal{P} \gg 1$, buoyancy is negligible and dissipation constrains the size of the PZ. When $\mathcal{P} \sim 1$, both terms matter.

3.2.1. *Case I: Discontinuous flux*

We first consider a model which satisfies

$$\overline{F_{\rm conv}}(z) = F_{\rm cz} \begin{cases} 1 & z \leq L_s, \\ -\mathcal{P}_D^{-1} & z > L_s \end{cases}. \quad (16)$$

Here, $F_{\rm cz}$ is a constant value of flux carried in the convection zone and $\mathcal{P}_D$ is the penetration parameter (subscript D for discontinuous case). Plugging this functional form of the flux into Eqn. 14, and integrating the

CZ over a depth $L_{\rm CZ}$ below L_s and the PZ over a height $\delta_{\rm p}$ above L_s, we predict

$$\frac{\delta_{\rm p}}{L_{\rm CZ}} = \mathcal{P}_D \frac{1-f}{1+\xi f \mathcal{P}_D}. \tag{17}$$

Assuming that f and ξ are weak functions of $\mathcal{P}_D$, we see that, for small $\mathcal{P}_D$, the size of the penetration region is linearly proportional to $\mathcal{P}_D$, but saturates as $\mathcal{P}_D \to \infty$ due to dissipation. Intuitively, this result makes sense: as $\mathcal{P}_D$ grows, the magnitude of $\overline{F_{\rm conv}}$ and the breaking force of buoyancy in the PZ shrink, resulting in larger penetrative regions (but this growth cannot extend indefinitely).

3.2.2. *Case II: Piecewise linear flux*

We next assume that $\overline{F_{\rm conv}}(z)$ is not discontinuous at the CZ-PZ boundary, but that its derivative may be,

$$\overline{F_{\rm conv}}(z) = \left.\frac{\partial F_{\rm rad}}{\partial z}\right|_{\rm CZ} \begin{cases} (L_s - z) & z \le L_s \\ -\mathcal{P}_L^{-1}(z - L_s) & z > L_s \end{cases}, \tag{18}$$

where $(\partial F_{\rm rad}/\partial z)|_{\rm CZ}$ is a constant and $\mathcal{P}_L$ is the penetration parameter (subscript L for linear case). When $\mathcal{P}_L = 1$, $\overline{F_{\rm conv}}$ is just a linear profile that crosses through zero at $z = L_s$. Again, solving Eqn. 14 with this functional form of the flux and integrating over $L_{\rm CZ}$ in the CZ and $\delta_{\rm p}$ in the PZ, we retrieve a quadratic equation. This equation has two solution branches, only one of which corresponds to a positive value of $\delta_{\rm p}$. On that branch, we find

$$\frac{\delta_{\rm p}}{L_{\rm CZ}} = \sqrt{\mathcal{P}_L(1-f)}\,\left(\sqrt{\zeta^2+1} - \zeta\right), \tag{19}$$

where $\zeta \equiv (\xi f/2)\sqrt{\mathcal{P}_L/(1-f)}$. We expect the penetration height to be roughly proportional to $\sqrt{\mathcal{P}_L}$ for small values of $\mathcal{P}_L$, and to again saturate at large values of $\mathcal{P}_L$.

In this work, we will test the predictions of Eqns. 17 and 19. Our goals are to see if the predicted scalings with the penetration parameter $\mathcal{P}$ are realized in simulations, and to measure the values of f and ξ.

4. SIMULATION DETAILS

We will now describe a set of simulations that test the predictions in Sec. 3. While many simulations of convection interacting with radiative zones have been performed by previous authors, ours differ in two crucial ways. First, we construct our experiments so that $\mathcal{P}$ and $\mathcal{S}$ can be varied separately. $\mathcal{P}$ is the "Penetration Parameter," defined in Eqn. 15, which compares the magnitude of the convective flux in the CZ and PZ; $\mathcal{S}$ is the "stiffness," defined in Eqn. 25, and compares the buoyancy frequency in the stable radiative zone to the convective frequency. We suspect that some past experiments have implicitly set $\mathcal{P} \approx \mathcal{S}^{-1}$, which would result in negligible penetration for high stiffness. Second, as we will show in Sec. 5, the development of penetrative zones is a slow process and many prior studies did not evolve simulations for long enough to see these regions grow and saturate.

We nondimensionalize Eqns. 1-4 on the length scale of the Schwarzschild-unstable convection zone L_s, the timescale of freefall across that convection zone $\tau_{\rm ff}$, and the temperature scale of the internal heating over that freefall time ΔT,

$$\begin{aligned} &T^* = (\Delta T)T = Q_0 \tau_{\rm ff} T, \qquad Q^* = Q_0 Q, \\ &\partial_{t^*} = \tau_{\rm ff}^{-1}\partial_t = \left(\frac{|\alpha| g Q_0}{L_s}\right)^{1/3} \partial_t, \quad \boldsymbol{\nabla}^* = L_s^{-1}\boldsymbol{\nabla}, \\ &\boldsymbol{u}^* = u_{\rm ff}\boldsymbol{u} = \left(|\alpha| g Q_0 L_s^2\right)^{1/3}\boldsymbol{u}, \qquad p^* = \rho_0 u_{\rm ff}^2 \varpi, \\ &k^* = (L_s^2 \tau_{\rm ff}^{-1})k, \qquad \mathcal{R} = \frac{u_{\rm ff} L_s}{\nu}, \qquad \mathrm{Pr} = \frac{\nu}{\chi}. \end{aligned} \tag{20}$$

For convenience, here we define quantities with $*$ (e.g., T^*) as being the "dimensionful" quantities of Eqns. 1-4. Henceforth, quantities without $*$ (e.g., T) are dimensionless. The dimensionless equations of motion are

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0 \tag{21}$$

$$\partial_t \boldsymbol{u} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} = -\boldsymbol{\nabla}\varpi + T\hat{z} + \mathcal{R}^{-1}\boldsymbol{\nabla}^2\boldsymbol{u} \tag{22}$$

$$\begin{aligned} \partial_t T + \boldsymbol{u}\cdot\boldsymbol{\nabla}T + w\boldsymbol{\nabla}_{\rm ad} + \boldsymbol{\nabla}\cdot[-k\boldsymbol{\nabla}\overline{T}] \\ = (\mathrm{Pr}\mathcal{R})^{-1}\boldsymbol{\nabla}^2 T' + Q. \end{aligned} \tag{23}$$

We construct a domain in the range $z \in [0, L_z]$ and choose $L_z \ge 2$ so that the domain is at least twice as deep as the Schwarzschild-unstable convection zone. We decompose the temperature field into a time-stationary initial background profile and fluctuations, $T(x,y,z,t) = T_0(z) + T_1(x,y,z,t)$. T_0 is constructed with $\nabla = \nabla_{\rm ad}$ for $z \le L_s$, and $\nabla = \nabla_{\rm rad}$ above $z > L_s$. We impose a fixed-flux boundary at the bottom of the box ($\partial_z T_1 = 0$ at $z = 0$) and a fixed temperature boundary at the top of the domain ($T_1 = 0$ at $z = L_z$). We generally impose impenetrable, no-slip boundary conditions at the top and bottom of the box so that $\boldsymbol{u} = 0$ at $z = [0, L_z]$. For a select few simulations, we impose stress-free instead of no-slip boundary conditions ($w = 0$ and $\partial_z u = \partial_z v = 0$ at $z = [0, L_z]$).

We impose a constant internal heating which spans only part of the convection zone,

$$Q = \begin{cases} 0 & z < 0.1 \text{ or } \mathrm{z} \ge 0.1 + \Delta_{\rm H}, \\ Q_{\rm mag} & 0.1 \le z \le 0.1 + \Delta_H \end{cases}. \tag{24}$$

The integrated flux through the system from heating is $F_H = \int_0^{L_z} Q_{\rm mag} dz = Q_{\rm mag}\Delta_H$. Throughout this work we choose $Q_{\rm mag} = 1$ and $\Delta_H = 0.2$ so $F_H = 0.2$. We offset this heating from the bottom boundary to $z = 0.1$ to avoid heating within the bottom impenetrable boundary layer where velocities go to zero and k is small; this prevents strong temperature gradients from establishing there. Furthermore, since the conductivity is not zero at the bottom boundary, the adiabatic temperature gradient there carries some flux, $F_{\rm bot} = \mu F_H$ and we choose $\mu = 10^{-3}$ so that most of the flux in the convection zone is carried by the convection.

The average convective velocity depends on the magnitude of the heating, $\langle|\boldsymbol{u}|\rangle \approx Q_{\rm mag}^{1/3} \approx 1$, so the characteristic convective frequency is $f_{\rm conv} \approx \langle|\boldsymbol{u}|\rangle/L_s \approx 1$. The stiffness is defined,

$$\mathcal{S} \equiv \frac{N^2}{f_{\rm conv}^2} \approx N^2, \tag{25}$$

where N^2 is the Brunt-Väisälä frequency in the radiative zone. In our nondimensionalization, $N^2 = \nabla_{\rm ad} - \nabla_{\rm rad}$. We treat $\mathcal{S}$ as a control parameter; by choosing a value of the stiffness we set the magnitude of $\nabla_{\rm rad}$, which in turn sets the value of k in the radiative zone.

Aside from $\mathcal{S}$ and $\mathcal{P}$, the two remaining control parameters $\mathcal{R}$ and Pr determine the properties of the turbulence. The value of $\mathcal{R}$ roughly corresponds to the value of the Reynolds number $\mathrm{Re} = \mathcal{R}|\boldsymbol{u}|$, and we set the ratio of the diffusivities $\mathrm{Pr} = 0.5$ throughout this work. Astrophysical convection is in the limit of $\mathrm{Pr} \ll 1$ (Garaud 2021); we choose a modest value of Pr which slightly separates the scales between thermal and viscous structures while still allowing us to achieve convection with large Reynolds and Péclet numbers.

4.1. *Case I: Discontinuous flux*

Most of the simulations in this paper have a discontinuous convective flux at the Schwarzschild convective boundary. We achieve this by constructing a discontinuous radiative conductivity,

$$k = \begin{cases} k_{\rm CZ} & z < 1 \\ k_{\rm RZ} & z \geq 1 \end{cases}, \tag{26}$$

where CZ refers to the convection zone and RZ refers to the radiative zone (some of which will be occupied by the penetrative zone PZ). Leaving $\mathcal{S}$ and $\mathcal{P}_D$ as free parameters and requiring that $\nabla_{\rm ad}$ carries $F_{\rm bot}$ at $z = 0$ and that $\nabla_{\rm rad}$ carries $F_H + F_{\rm bot}$ for $z \geq 1$ specifies this system fully,

$$\begin{aligned} k_{\rm RZ} &= \frac{\Delta_H}{\mathcal{S}\mathcal{P}_D}, \\ k_{\rm CZ} &= k_{\rm RZ}\frac{1}{1+\mu+\mathcal{P}_D^{-1}}, \\ \nabla_{\rm ad} &= Q_{\rm mag}\mathcal{S}\mathcal{P}_D(1+\mu+\mathcal{P}_D^{-1}), \\ \nabla_{\rm rad} &= \nabla_{\rm ad} - Q_{\rm mag}\mathcal{S}. \end{aligned} \tag{27}$$

We study three sweeps through the ($\mathcal{P}_D$, $\mathcal{S}$, $\mathcal{R}$) parameter space in this paper in which we hold two of these parameters constant and vary the other. We study an additional sweep through $\mathcal{R}$ parameter space using stress-free boundaries to compare to our no-slip cases. According to Eqn. 17, we expect $\delta_{\rm p} \propto \mathcal{P}_D$.

4.2. *Case II: Piecewise linear flux*

We also study simulations where the flux's gradient may be discontinuous at the Schwarzschild convective boundary. We achieve this by constructing a radiative conductivity with a piecewise discontinuous gradient,

$$\partial_z k = \partial_z k_0 \begin{cases} 1 & z < 1 \\ \mathcal{P}_L^{-1} & z \geq 1 \end{cases} \tag{28}$$

Since k varies with height, the value of $\mathcal{S}$ and $\mathcal{P}$ also vary with height; we specify their values at $z = 2$. By this choice, we require

$$\partial_z k_0 = \frac{\Delta_H}{L_s \mathcal{S}\psi}, \qquad k_b = \frac{\Delta_H \mu}{\mathcal{S}\psi}, \qquad \nabla_{\rm ad} = Q\mathcal{S}\psi, \tag{29}$$

where $\psi \equiv 1 + \mathcal{P}_L(1+\mu)$. We will study one sweep through $\mathcal{P}_L$ space at fixed $\mathcal{R}$ and $\mathcal{S}$. According to Eqn. 19, we expect $\delta_p \propto \mathcal{P}_L^{1/2}$.

4.3. *Numerics*

We time-evolve equations 21-23 using the Dedalus pseudospectral solver (Burns et al. 2020)[2] using timestepper SBDF2 (Wang & Ruuth 2008) and safety factor 0.35. All fields are represented as spectral expansions of n_z Chebyshev coefficients in the vertical (z) direction and as (n_x,n_y) Fourier coefficients in the horizontal (x,y) directions; our domains are therefore horizontally periodic. We use a domain aspect ratio of two so that $x \in [0, L_x]$ and $y \in [0, L_y]$ with $L_x = L_y = 2L_z$. To avoid aliasing errors, we use the 3/2-dealiasing rule in all directions. To start our simulations, we add random noise temperature perturbations with a magnitude

[2] we use commit efb13bd; the closest stable release to this commit is v2.2006.

of 10^{-3} to a background temperature profile $\overline{T}$; we discuss the choice of $\overline{T}$ in appendix A. In some simulations we start with $\overline{T} = T_0$, described above, and in others we impose an established penetrative zone in the initial state $\overline{T}$ according to Eqn. A1.

Spectral methods with finite coefficient expansions cannot capture true discontinuities. In order to approximate discontinuous functions such as Eqns. 24, 26, and 28, we must use smooth transitions. To create these smooth transitions, we define an approximate Heaviside step function using the error function,

$$H(z; z_0, d_w) = \frac{1}{2}\left(1 + \mathrm{erf}\left[\frac{z - z_0}{d_w}\right]\right). \tag{30}$$

In the limit that $d_w \to 0$, this function behaves identically to the classical Heaviside function centered at z_0. For Eqn. 24 and Eqn. 28, we use $d_w = 0.02$; while for Eqn. 26 we use $d_w = 0.075$. In all other cases, we use $d_w = 0.05$.

A table describing all of the simulations presented in this work can be found in Appendix C. We produce the figures in this paper using matplotlib (Hunter 2007; Caswell et al. 2021). All of the Python scripts used to run the simulations in this paper and to create the figures in this paper are publicly available in a git repository, found at [TODO CITE].

4.4. *Penetration height measurements*

In our evolved simulations, we find that the penetrative region has a nearly adiabatic stratification $\nabla \approx \nabla_{\mathrm{ad}}$. In order to characterize the vertical extent of the penetrative region, we measure how drastically ∇ has departed from ∇_{ad}. We define the difference between the adiabatic and radiative gradient,

$$\Delta \equiv \nabla_{\mathrm{ad}}(z) - \nabla_{\mathrm{rad}}(z). \tag{31}$$

We measure penetration and overshoot heights in terms of "departure points," or heights at which the realized temperature gradient ∇ has evolved away from the adiabatic ∇_{ad} by some fraction $h < 1$. Specifically,

$$L_s + \delta_h = \max(z) \;\mid\; \nabla > (\nabla_{\mathrm{ad}} - h\,\Delta). \tag{32}$$

In this work, we measure the 10% ($\delta_{0.1}$, $h = 0.1$), 50% ($\delta_{0.5}$, $h = 0.5$), and 90% ($\delta_{0.9}$, $h = 0.9$) departure points. Using Zahn (1991)'s terminology, $\delta_{0.5}$ is the mean value of the top of the PZ while $\delta_{0.9} - \delta_{0.1}$ represents the width of the "thermal adjustment layer." We will simply refer to these as three different measures of the top of the PZ. We find that these measurements based on the (slowly-evolving) thermodynamic profile are more robust and straightforward than many previous dynamically-based prescriptions (see e.g., Pratt et al. 2017, for a nice discussion).

5. RESULTS

We will now present on the results of the 3D dynamical simulations described in the previous section. Dynamics in these simulations resemble those previously shown in Fig. 1. To quantitatively compare our simulations to the theory of Sec. 3, we will present time- and volume-averaged scalar quantities, as well as some select horizontally-averaged vertical profiles.

5.1. *Qualitative description of simulation evolution*

In Fig. 4, we show the time evolution of a Case I (discontinuous conductivity) simulation with $\mathcal{R} = 400$, $\mathcal{S} = 10^3$, and $\mathcal{P}_D = 4$ whose temperature profile has $\nabla = \nabla_{\mathrm{ad}}$ in the convection zone ($z \lesssim 1$) and $\nabla = \nabla_{\mathrm{rad}}$ in the radiative zone ($z \gtrsim 1$). In the top left panel, we display the height of the penetrative region $\delta_{0.5}$ vs. time. This region initially grows quickly over hundreds of freefall times, but this evolution slows down; reaching the final equilibrium takes tens of thousands of convective overturn (freefall) times. The evolution of the other parameters in our theory (f, ξ) are shown in the middle and bottom left panels of Fig. 4. We plot their rolling mean, averaged over 200 freefall time units. We see that the values of f and ξ reach their final values ($f \approx 0.67$, $\xi \approx 0.58$) faster than the penetration zone evolves to its full height. We quantify this fast evolution by plotting a vertical line in each of the left three panels corresponding to the first time at which the rolling average converges to within 1% of its final value (averaged over the final 1000 freefall times of the simulation and plotted as a grey horizontal line). The evolved value of f indicates that roughly 2/3 of the buoyancy driving is dissipated in the bulk CZ, so that 1/3 is available for PZ dissipation and negative buoyancy work. The evolved value of ξ indicates that the shape of dissipation in the PZ is slightly steeper than linear.

In the right panel of Fig. 4, we plot the profile of $\nabla/\nabla_{\mathrm{ad}}$ in our simulation at regular time intervals, where the color of the profile corresponds to time, as in the left panels. ∇_{ad} is plotted as a dashed horizontal line while ∇_{rad} is plotted as a grey solid line which decreases with height around $z \approx 1$ and satures to a constant above $z \gtrsim 1.1$. The location of the Schwarzschild boundary, L_s, is overplotted as a black vertical dashed line and does not evolve over the course of the simulation. We note that the Schwarzschild boundary does not move over the course of our simulation, so the extention of the convection zone past this point is true penetration and not the result of entrainment-induced changes in the Schwarzschild (or Ledoux) convective boundaries. The traces of $\delta_{0.1}$ and $\delta_{0.9}$ are plotted as red lines while that of $\delta_{0.5}$ is plotted as a black line. We see that the fast

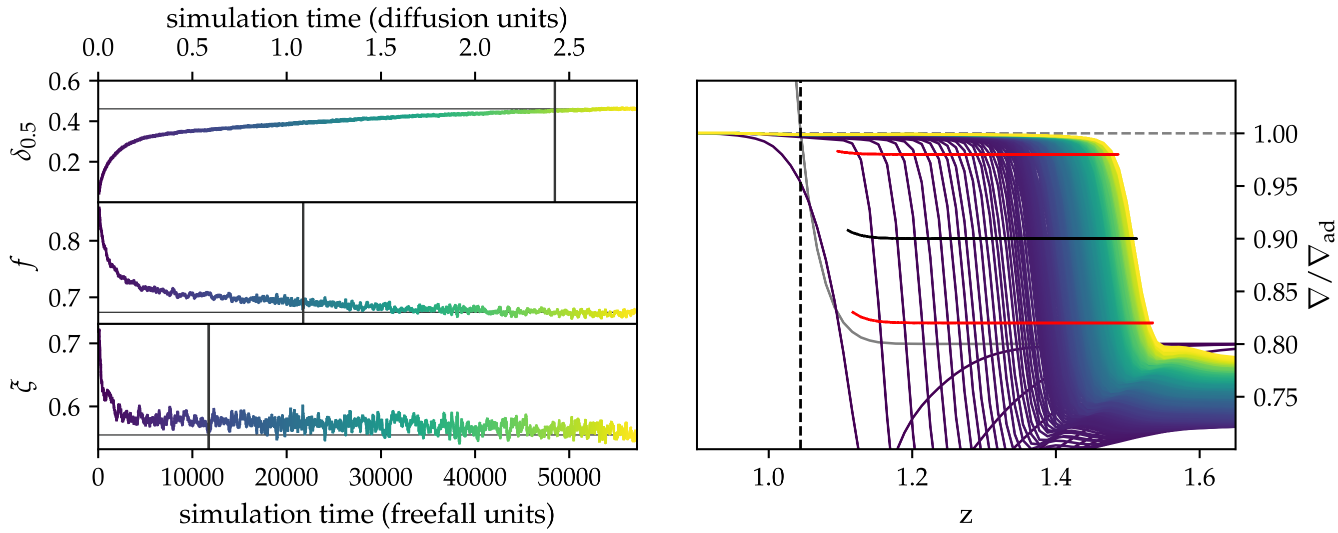

Figure 4. (left top panel) The PZ height $\delta_{0.5}$ vs. time. We also show the time evolution of f (left middle panel, defined in Eqn. 11) and ξ (left bottom panel, defined in Eqn. 13). The vertical lines denote the times at which the traces first converge to within 1% of their final values, which are denoted by solid horizontal lines. (right panel) The vertical profile of $\nabla/\nabla_{\rm ad}$ is plotted against height at regular time intervals. The line color denotes the time, following the time traces in the left panels. The constant value of $\nabla_{\rm ad}$ is denoted by the horizontal dashed grey line, and the value of $\nabla_{\rm rad}$ is denoted by the solid grey line. The location of the Schwarzschild convective boundary is displayed as a vertical dashed black line and denotes the place where $\nabla_{\rm ad} = \nabla_{\rm rad}$. The top-of-PZ departure points (Eqn. 32) are plotted over the profile evolution ($\delta_{0.1}$ and $\delta_{0.9}$ as red lines, $\delta_{0.5}$ as black lines).

initial evolution establishes a sizeable PZ (denoted by purple ∇ profiles), but its final equilibration takes much longer (indicated by the separation between the purple, green, and yellow profiles decreasing over time).

This long evolution is computationally expensive; for this modest simulation (256x64^2 coefficients), the evolution represents roughly 20 days of evolution on 1024 cores for a total of $\sim$500,000 cpu-hours. It is not feasible to perform simulations of this length for a full parameter space study, and so we accelerate the evolution of most of the simulations in this work. To do so, we take advantage of the nearly monotonic nature of the evolution of $\delta_{\rm p}$ vs. time displayed in Fig. 4. We measure the instantaneous values of $(\delta_{0.1}, \delta_{0.5}, \delta_{0.9})$, as well as their instantaneous time derivatives. Using these values, we take a large "time step" forward, advancing the values of $\delta_{\rm p}$ according to their current trend. While doing so, we preserve the width of the transition from the PZ to the RZ, and we also adjust the solution so that $\nabla = \nabla_{\rm rad}$ in the RZ, effectively equilibrating the RZ instantaneously. In other words, we reinitialize the simulation's temperature profile with a better guess at its evolved state based on its current dynamical evolution. For details on how this procedure is carried out, see Appendix A.

5.2. *Dependence on $\mathcal{P}$*

We find that the height of the penetration zone is strongly dependent on $\mathcal{P}$. In the upper two panels of Fig. 5, we plot the penetration height ($\delta_{0.1}$, $\delta_{0.5}$, $\delta_{0.9}$ from Eq. 32) from Case I simulations (discontinuous k, upper left) and Case II simulations (discontinuous $\partial_z k$, upper right). The fixed values of $\mathcal{R}$ and $\mathcal{S}$ are shown above these panels. We find that the leading-order $\mathcal{P}$ scaling predictions of Eqns. 17 & 19 describe the data extremely well (orange lines). At small values of $\mathcal{P}$ we see somewhat weaker scalings than these predictions, because the profiles of k and $\partial_z k$ are not truly discontinuous but jump from one value in the CZ to another in the RZ over a finite width (see e.g., the $\nabla_{\rm rad}$ profile in Fig. 4 & Sec. 4.3). At large values of $\mathcal{P}$, penetration height falls off of these predicted scaling laws. In this regime, dissipation dominates over buoyancy in the PZ and PZ heights saturate.

The middle and bottom panels of Fig. 5 demonstrate that that f and ξ are to leading order constant with $\mathcal{P}$. However, we find that f has slightly smaller values in the Case I simulations (left) than in the Case II simulations (right). We measure characteristic values of $f \in [0.6, 0.9]$, signifying that 60-90% of the buoyant work is balanced by dissipation in the convection zone, depending on the simulation. We note a weak trend where f decreases as $\mathcal{P}$ increases. As $\mathcal{P}$ incraeses, we find that CZ velocities decrease, leading to a decrease in the dissipation rate. When $\mathcal{P}$ is small, the PZ-RZ boundary (which acts like a wall, left panel of Fig. 1) ef-

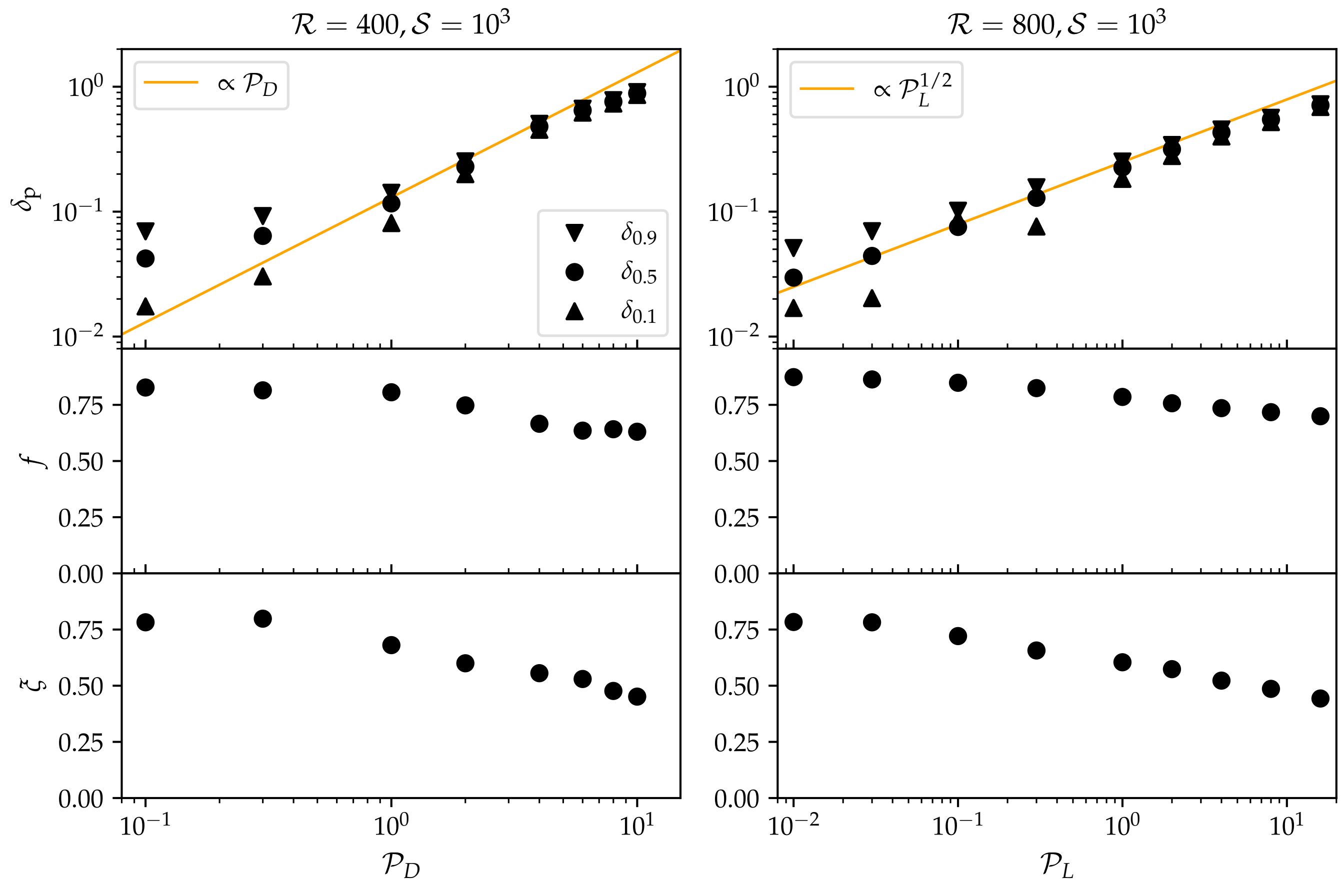

Figure 5. Results from simulations with a Schwarzschild boundary characterized by a discontinuous conductivity (left, Case I, k defined by Eqn. 26) and with a discontinuous conductivity gradient (right, Case II, k defined by Eqn. 28). The top panels show the penetration height according to Eqn. 32. The Case I penetration heights (upper left) vary linearly with $\mathcal{P}$, in line with the prediction of Eqn. 17. The Case II penetration heights (upper right) vary like $\sqrt{\mathcal{P}}$, in line with the prediction of Eqn. 19. In the middle panels, we measure f according to Eqn. 11. We find values of $f \in [0.5, 0.9]$, and changes in f are secondary to changes in $\mathcal{P}$ for determining penetration heights. In the bottom panels, we measure ξ according to Eqn. 13. We find characteristic values of $\xi \in [0.5, 0.75]$, suggesting that the falloff of the $\overline{\Phi}$ in the PZ is well described by a linear function (at high $\mathcal{P}$ when $\xi = 1/2$), or by a cubic function (at low $\mathcal{P}$ when $\xi \approx 3/4$).

ficiently deflects convective velocities sideways resulting in increased bulk-CZ velocities. As $\mathcal{P}$ grows, the velocities have access to an extended PZ in which to buoyantly break before deflection, resulting in slightly lower bulk velocities. A similar trend of ξ decreasing as $\mathcal{P}$ increases can be seen. Recall that smaller values of ξ indicate the dissipative dynamics are rather different in the PZ and CZ. As the size of the PZ grows, the dynamical structures of the PZ shift from what is found in the CZ, and so ξ shrinks.

5.3. *Dependence on $\mathcal{S}$*

We find that the height of the penetration zone is weakly dependent on $\mathcal{S}$. In the left panel of Fig. 6, we plot the penetration height of a few discontinuous-k simulations with $\mathcal{P}_D = 4$ and $\mathcal{R} = 400$ but with different values of $\mathcal{S}$. We find that the mean penetration height $\delta_{0.5}$ varies only weakly with changing $\mathcal{S}$, but that the values of $\delta_{0.1}$ and $\delta_{0.9}$ vary more strongly. In other words, the transition region in which ∇ varies from $\nabla_{\rm ad}$ in the PZ to $\nabla_{\rm rad}$ in the RZ becomes narrower as $\mathcal{S}$ increases. To quantify this effect, we plot $\delta_{0.9} - \delta_{0.1}$ in the righthand panel of Fig. 6. We find that the width of this region varies according to a $\mathcal{S}^{-1/2}$ scaling law, reminiscent of the pure-overshoot law described by Korre et al. (2019).

Note that if the enstrophy, ω^2 in the convection zone exceeds the value of the square buoyancy frequency N^2 in the radiative zone, the gravity waves in the RZ became nonlinear. As a result, we have restricted the simulations in this study to relatively large (yet modest compared to astrophysical values) of $10^2 \leq \mathcal{S} < 10^4$ in order to ensure $N^2 > \omega^2$ even in our highest enstrophy simulations.

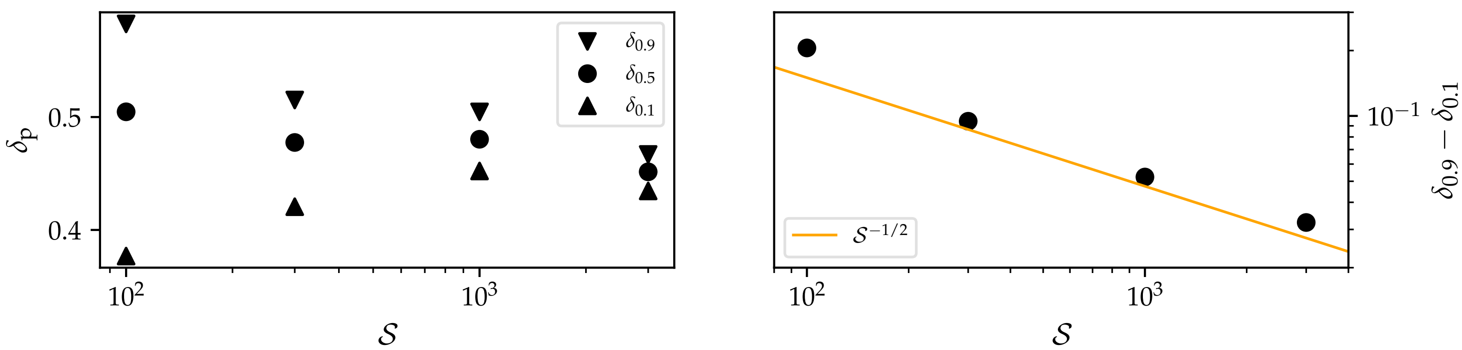

Figure 6. (Left panel) Penetration heights are shown as a function of stiffness for simulations with a discontinuous conductivity. While the width of the transition layer from the adiabatic PZ to the stable RZ shrinks as a function of $\mathcal{S}$, the mean penetration height ($\delta_{0.5}$) is roughly constant. (Right panel) The width of the thermal transition layer ($\delta_{0.9} - \delta_{0.1}$) is shown as a function of stiffness. We roughly observe a $\mathcal{S}^{-1/2}$ scaling, similar to that expected by e.g., Korre et al. (2019).

5.4. *Dependence on $\mathcal{R}$*

We find that the height of the penetration zone is weakly dependent on $\mathcal{R}$. In the upper left panel of Fig. 7, we find a logarithmic decrease in the penetration height with the Reynolds number. In order to understand how this could happen at fixed $\mathcal{P}$, we also plot the output values of f (upper middle) and ξ (upper right). We find that f increases with increasing $\mathcal{R}$, but is perhaps leveling off as $\mathcal{R}$ becomes large. We find that ξ does not increase strongly with $\mathcal{R}$ except for in the case of laminar simulations with $\mathcal{R} < 200$. Eqn. 17 predicts that δ_p should change at fixed $\mathcal{P}$ and ξ if f is changing. In the bottom left panel, we show that the change in δ_p is due to this change in f. We find that this is true both for simulations with stress-free dynamical boundary conditions (open symbols, SF) and for no-slip conditions (closed symbols, NS).

We now examine why f increases as $\mathcal{R}$ increasesIn the SF simulations, within the CZ, we can reasonably approximate $\overline{\Phi}$ as a constant Φ_CZ in the bulk and zero within the viscous boundary layer,

$$\overline{\Phi}(z) = \begin{cases} \Phi_\mathrm{CZ} & z > \ell_\nu \\ 0 & z \leq \ell_\nu \end{cases}, \tag{33}$$

where ℓ_ν is the viscous boundary layer depth. We have visualized a NS dissipation profile in the bottom panel of Fig. 3; SF simulations look similar in the bulk, but drop towards zero at the bottom boundary rather than reaching a maximum. Then, we have

$$\int_\mathrm{CZ} \overline{\Phi}\, dz \approx \Phi_\mathrm{CZ}\,(L_s - \ell_\nu)\,, \tag{34}$$

and so per Eqn. 11,

$$f = f_\infty \left(1 - \frac{\ell_\nu}{L_s}\right), \tag{35}$$

where f_∞ is the expected value of f at $\mathcal{R} = \infty$ when $\ell_\nu = 0$. So we see that the CZ dissipation and therefore f vary linearly with ℓ_ν.

In the bottom middle panel of Fig. 7, we find that Eqn. 35 with $f_\infty = 0.755$ captures the high-$\mathcal{R}$ behavior. To measure ℓ_ν, we first measure the height of the extremum of the viscous portion of the kinetic energy flux $\overline{\mathcal{F}}$ near the boundary, and take ℓ_ν to be the twice that height. We find that Eqn. 11 is a slightly better description for the SF simulations than the NS simulations; NS simulations have maximized dissipation in the bounary layer, and therefore Eqn. 33 is a poor model for $z \leq \ell_\nu$. In the bottom right panel of Fig. 7, we demonstrate that the depth of the viscous boundary layer follows classical scaling laws from Rayleigh-Bénard convection[3] (Ahlers et al. 2009; Goluskin 2016). Combining these trends, we expect

$$f = f_\infty(1 - C\mathcal{R}^{-2/3}) \tag{36}$$

for a constant C. Thus as $\mathcal{R} \to \infty$, $f \to f_\infty$.

We use the fitted function of f from the bottom middle panel, along with Eqn. 17, to estimate $\delta_{0.5}$ in the bottom left panel. We need to multiply this equation by a factor of 0.9, which accounts for some differences between the simulations and the idealized "discontinuous flux" theoretical model. First, due to internal heating and the finite width of the conductivity transition around the Schwarzschild boundary, the convective flux is not truly constant through the full depth of the CZ. Thus, we expect L_CZ in Eqn. 17 to be smaller than 1. Furthermore,

[3] If you assume the Nusselt Number dependence on the Rayleigh number is throttled by the boundaries, $\mathrm{Nu} \propto \mathrm{Ra}^{1/3}$ (as is frequently measured), and the Reynolds number is $\mathrm{Re} \propto \mathrm{Ra}^{1/2}$, you retrieve $\mathrm{Nu} \propto \mathrm{Re}^{2/3}$. The Nusselt number generally varies like the inverse of the boundary layer depth, $\mathrm{Nu} \propto \delta^{-1}$, and so we expect $\delta_\nu \propto \mathcal{R}^{-2/3}$.

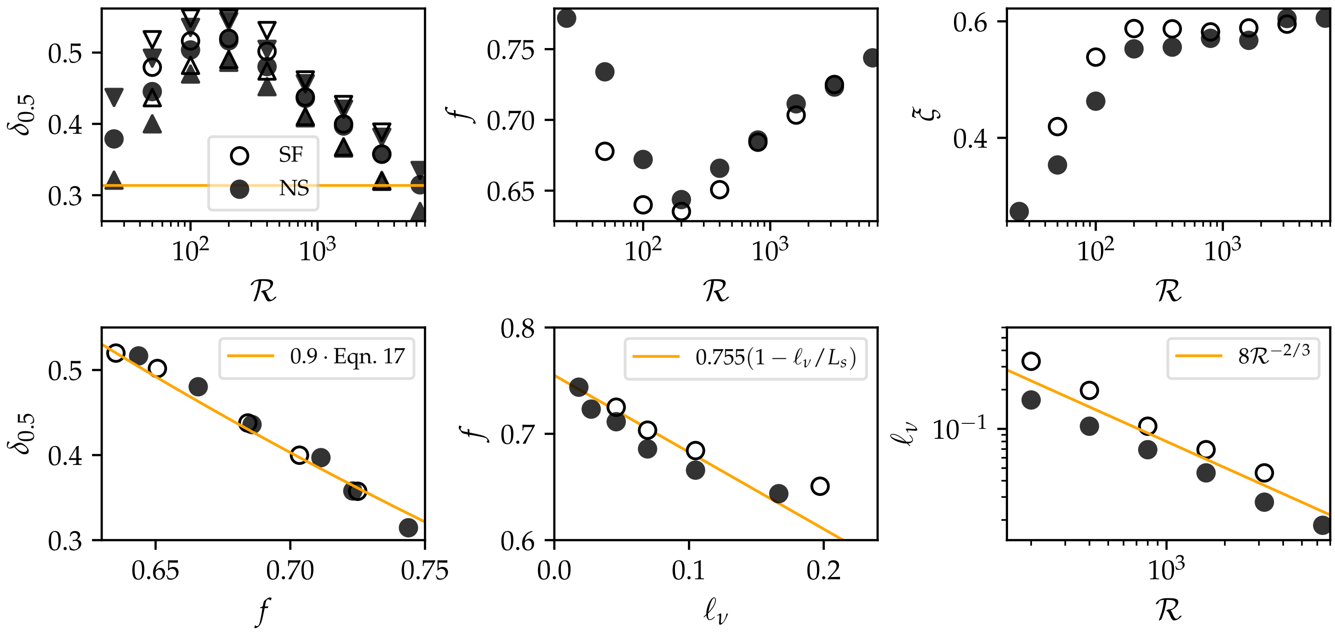

Figure 7. (Upper left panel) Penetration heights are shown as a function of increasing Reynolds number ($\mathcal{R}$) for simulations with a discontinuous conductivity. We plot results both for cases with stress-free boundary conditions (SF) and no-slip boundary conditions (NS). In both cases, we see a roughly logarithmic decrease of δ_{p} with $\mathcal{R}$. (Upper middle panel) f also increases with $\mathcal{R}$. (Upper right panel) ξ does not change appreciably with $\mathcal{R}$ for turbulent simulations with $\mathcal{R} \geq 200$. (Lower left panel) There is a strong correlation between $\delta_{0.5}$ and f, agreeing with our theoretical model of Eqn. 17. (Lower middle panel) Changes in f are roughly linearly proportional to the depth of the viscous boundary layer, ℓ_ν, at the bottom of the domain, at least for the open-circle SF cases. (Lower right panel) The viscous boundary layer follows a classical scaling law seen in Rayleigh-Bénard convection, so we anticipate that $\delta_{0.5}$ and f should saturate as $\mathcal{R} \to \infty$ and $\delta_\nu \to 0$.

the theory is derived in the limit of an instantaneous transition from ∇_{ad} to ∇_{rad} where $\delta_{0.1} = \delta_{0.5} = \delta_{0.9}$; our simulations have a finite transition width. Despite these subtle differences, we find good agreement.

Using $f_\infty = 0.755$ we estimate that $\delta_{0.5} \approx 0.31$ for $\mathcal{R} \to \infty$ and plot this as a horizontal orange line on the upper left panel of Fig. 7. This value is coincidentally very near the value of $\delta_{0.5}$ achieved in our highest-$\mathcal{R}$ simulations. Unfortunately, we cannot probe more turbulent simulations. We can only run the $\mathcal{R} = 6.4 \times 10^3$ simulation for a few hundred freefall times. Our accuracy in measuring results from this simulation is limited by the long evolutionary timescales of the simulation (see Fig. 4). Even accounting for our accelerated evolutionary procedure, we can only be confident that the PZ heights of this simulation are converged to within a few percent. Future work should aim to better understand the trend of PZ height with turbulence. However, the displayed relationships between δ_{p} and f, f and ℓ_ν, and ℓ_ν and $\mathcal{R}$ — all of which are effects we largely understand — suggest that PZ heights should saturate at high $\mathcal{R}$.

In summary, we find that δ_{p} decreases as $\mathcal{R}$ increases. We find that these changes are caused by increases in f. In our simulations, these increases in f seem to be caused by changes in the size of the viscous boundary layer ℓ_ν. By measuring f and ℓ_ν in a simulation, the value of f_∞ can be found from Eqn. 35. Stellar convection zones are not adjacent to hard walls and so should be represented by the $\ell_\nu \to 0$ limited. Core convection zones geometrically have no lower boundary, and convective shells should be bounded both above and below by penetrative regions.

While we have examine a Case I simulation with $\mathcal{P} = 4$ here, we expect the simulation with $\mathcal{P}_L = 1$ (a linear radiative conductivity profile) to be the most representative of conditions near a stellar convective boundary. In this simulation, we measure $\xi \approx 0.6$, $f \approx 0.785$, $\ell_\nu \approx 0.08$, and $L_s = 1$. We therefore estimate that

$$f_\infty = 0.86 \qquad \text{and} \qquad \xi = 0.6 \tag{37}$$

are good first estimates for f and ξ when applying our theory of penetrative convection to stellar models.

6. A MODIFIED SOLAR MODEL

Our simulation results present a strong case for a flux- and dissipation-based model of convective penetration, similar to those considered by Zahn (1991) and Rox-

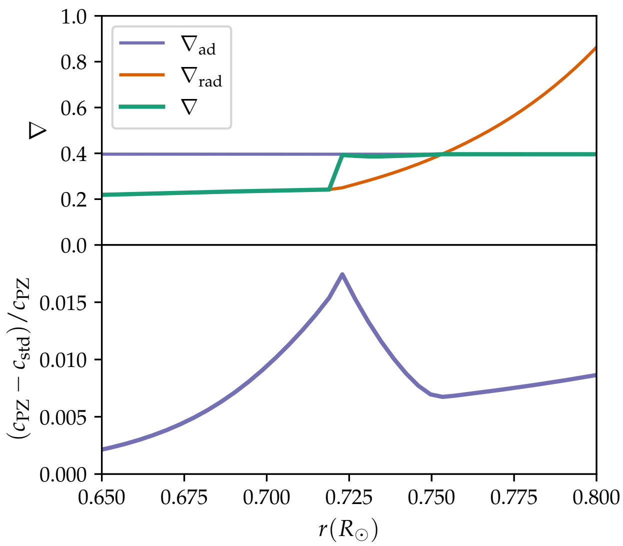

Figure 8. (top) Profiles of ∇ (green), $\nabla_{\rm ad}$ (purple), and $\nabla_{\rm rad}$ (orange) in a MESA solar model with a penetration zone. (bottom) The difference in the sound speed in the MESA model with a PZ that is presented in the top panel compared to a model run at identical parameters but without a PZ. We see that the addition of a PZ increases the sound speed by $\mathcal{O}$(1-2%) below the convection zone.

burgh (1989). In order to implement our theory into MESA, we need to extend Eqn. 14 to spherical geometry. To do so, we replace horizontal averages in Eqn. 9 with integrals over latitude and longitude, and find that the relevant integral constraint contains the convective luminosity,

$$\int |\alpha| g L_{\rm conv}\, dr = \int_V \rho_0 \Phi\, dV, \tag{38}$$

where $L_{\rm conv} = 4\pi\rho_0 r^2 \overline{F_{\rm conv}}$, r is the radial coordinate, and we write the RHS as a volume integral. We next define f in the same way as in Eqn. 11 and define ξ similarly to Eqn. 13,

$$\int_{\rm PZ} \rho_0 \Phi\, dV = \xi \frac{V_{\rm PZ}}{V_{\rm CZ}} \int_{\rm CZ} \rho_0 \Phi\, dV, \tag{39}$$

where $V_{\rm PZ}$ and $V_{\rm CZ}$ are the volumes of the CZ and PZ respectively. Eqn. 39 generalizes Eqn. 13 outside of the assumption of a plane-parallel atmosphere. Thus Eqn. 14 in spherical geometry is

$$-\frac{\int_{\rm PZ} L_{\rm conv}\, dr}{\int_{\rm CZ} L_{\rm conv}\, dr} + f\xi \frac{V_{\rm PZ}}{V_{\rm CZ}} = (1-f), \tag{40}$$

We implemented Eqn. 40 in MESA (see Appendix B for details) and computed a standard solar model with $f = 0.86$ and $\xi = 0.6$ (Eqn. 37) to determine an estimate of how the convective overshoot parameterization presented in this work might adjust a familiar stellar model. We note that the simulations that we have presented here do not include many of the complications of stellar convection like density stratification, sphericity, rotation, magnetism, etc. We present this model as a proof of concept and to inspire further work.

In the top panel of Fig. 8 we display $\nabla \equiv d\ln T/d\ln P$ from a 1 $M_\odot$ MESA model which has been evolved to an age of 4.56 Gyr and which includes convective penetration. Note that ∇ (green) remains close to $\nabla_{\rm ad}$ (purple) below the Schwarzschild convective boundary ($\nabla_{\rm ad} = \nabla_{\rm rad}$) in a penetration zone. After some depth $\nabla \to \nabla_{\rm rad}$ (orange) in the star's interior. We additionally evolved a standard solar MESA model to a 4.56 Gyr age without the inclusion of a PZ. We compare the sound speed profiles of the PZ and standard (std) model in the bottom panel of Fig. 8. We see that the inclusion of a PZ leads to an $\mathcal{O}$(1-2%) increase in the sound speed profile near the base of the solar convection zone, similar to what is observed between helioseismic observations and standard models (Bergemann & Serenelli 2014).

At the Schwarzschild base of the CZ, we find $H_p \approx 0.082 R_\odot$, and the depth of the penetration zone in Fig. 8 is $0.030 R_\odot$, such that this PZ extends $\sim 36\%$ of a scale height below the convection zone. Helioseismic studies have suggested that a PZ depth should be constrained to 5-10% of H_P (Basu et al. 1994; Basu 2016); these constraints are significantly smaller than our crude model. Some recent simulations (Käpylä 2019) have suggested that solar convection could exhibit penetration depths of $0.2 H_P$, which are closer to these models but still smaller than what we measure. Despite these discrepancies, Fig. 8 shows the promise of a convective penetration parameterization such as the one presented here. Pathways for improving our parameterization will be discussed in the following section.

7. DISCUSSION

In this work, we presented dynamical simulations of convective penetration, in which convection flattens $\nabla \to \nabla_{\rm ad}$ beyond the Schwarzschild boundary. To understand these simulations, we used an integral constraint (reminiscent of Roxburgh 1989) and flux-based arguments (similar to Zahn 1991) to derive a parameterization of convective penetration according to the convective flux and viscous dissipation. In doing so, we have laid down the first steps (Eqns. 14 & 40) towards incorporating convective penetration into stellar structure codes. We parameterized the viscous dissipation into a bulk-CZ portion (f) and a portion in the extended penetrative region (ξ), and derived predictions for how the height of a penetrative region, $\delta_{\rm p}$, should scale with these measurable parameters and a new flux-based "penetration parameter" $\mathcal{P}$. We designed and analyzed two sets of simulations which showed good agreement with these theoretical predictions. We briefly examined what the impliciations of this theory could be for a simple solar model.

Our simulation results suggest that stellar convection zones should be bounded by sizeable penetration zones. In extreme simulations, we observe penetration zones which are as large as the convection zones they accompany; however, for realistic stellar values ($\mathcal{P} \approx 1$), we find that they may be as large as 20-30% of the convective zone length scale (∼the mixing length).

The simulations we presented in this work use a simplified setup to test the basic tenets of our theory. In particular, they demonstrate that the shape of the flux near the convective boundary and the viscous dissipation together determine the height of the penetration zone. The precise values of the parameters f and ξ achieved in natural, turbulent, fully compressible, spherical stellar convection may be different from those presented in e.g., Fig. 5 here. Future work should aim to understand how these parameters and the theory presented in e.g., Eqn. 40 change when these realistic effects are taken into account.

Furthermore, it is important to note that stellar opacities, and thus stellar conductivities, are functions of thermodynamic variables rather than radial location. As a result, the formation of a penetration zone will in turn affect the conductivity profile and ∇_{rad}, which will in turn affect the location of the Schwarzschild boundary and the estimate of how deep the penetration zone should be. Future studies should follow e.g., Käpylä et al. (2017) and implement realistic opacity profiles which evolve self-consistently with the thermodynamic state in order to understand how these effects feedback into one another.

An additional complication is that stellar fluid dynamics exist in the regime of $\mathrm{Pr} \ll 1$ (Garaud 2021). Dynamics in this regime may be different from those in the regime of $\mathrm{Pr} \lesssim 1$ that we studied here, which in theory could affect f and ξ. Recently, Käpylä (2021) found that convective flows exhibited more penetration at low Pr than high Pr. Future work should aim to understand whether f and/or ξ depend strongly on Pr in the turbulent regime.

Two other interesting complications in stellar contexts are rotation and magnetism. In the rapidly rotating limit, rotation creates quasi-two-dimensional flows, which could affect the length scales on which dissipation acts and thus modify f. Furthermore, magnetism adds an additional ohmic dissipation term, which could in theory drastically change our hydrodynamical measurement of f.

Our work here assumes a uniform composition through the convective and radiative region. Convective boundaries often coincide with discontinuities in composition profiles (Salaris & Cassisi 2017). Future work should determine if stabilizing composition gradients can prevent the formation of the penetration zones seen here (as in Fig. 4).

In summary, we have unified Roxburgh (1989)'s integral constraint with Zahn (1991)'s theory of flux-dependent penetration into a parameterized theory of convective penetration. We tested this theory with simulations and found good agreement between the theory and our simulations. In future work, we will aim to more robustly implement this theory into MESA and use simulations to test some of the complicating factors we discussed here.

We thank Keaton Burns, Matt Browning, Matteo Cantiello and Kyle Augustson for useful discussions and questions which improved the content of this manuscript. EHA is funded as a CIERA Postdoctoral fellow and would like to thank CIERA and Northwestern University. We acknowledge the hospitality of Nordita during the program "The Shifting Paradigm of Stellar Convection: From Mixing Length Concepts to Realistic Turbulence Modelling," where the groundwork for this paper was set. This work was supported by NASA HTMS grant 80NSSC20K1280. Computations were conducted with support from the NASA High End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center on Pleiades with allocation GID s2276. The Flatiron Institute is supported by the Simons Foundation.

APPENDIX

A. ACCELERATED EVOLUTION

As demonstrated in Fig. 4, the time evolution of simulations which start from a state based on the Schwarzschild criterion can be prohibitively long. In Anders et al. (2018), we explored the long time evolution of simple convective simulations and found that fast-forwarding the evolution of a convective simulation's internal energy and thermal structure can be done accurately. This can be done because the convective dynamics converge rapidly even if the thermal profile converges slowly. This same separation of scales is observed in the penetrative dynamics in this work, and so similar techniques should be applicable.

To more quickly determine the final size of the evolved penetration zones we use the following algorithm.

1. Once a simulation has a volume-averaged Reynolds number greater than 1, we wait 10 freefall times to allow dynamical transients to pass.

2. We measure the departure points ($\delta_{0.1}$, $\delta_{0.5}$, $\delta_{0.9}$) every freefall time, and store this information for 30 freefall times.

3. We linearly fit each of the departure points' evolution against time using NumPy's `polyfit` function. We assume that convective motions influence $\delta_{0.1}$ and $\delta_{0.5}$ more strongly than $\delta_{0.9}$. We measure the time-evolution of the convective front $\frac{d\delta_{\rm p}}{dt}$ by averaging the slope of the linear fits for $\delta_{0.1}$ and $\delta_{0.5}$.

4. We take a large "time step" of size $\tau_{\rm AE}$ forward. We calculate $\Delta\delta_p = \tau_{\rm AE}\frac{d\delta_{\rm p}}{dt}$.

   - If $\Delta\delta_p < 0.005$, we erase the first 15 time units worth of departure point measures and return to step 2 for 15 time units.
   - If $\Delta\delta_p$ is large, we adjust the top of the PZ by setting $\delta_{0.5,\rm new} = \langle\delta_{0.5}\rangle_t + \Delta\delta_p$ (angles represent a time average). If $|\Delta\delta_p| > 0.05$, we limit its value to 0.05. We calculate the width of the thermal adjustment layer d_w as the minimum of $\langle\delta_{0.9} - \delta_{0.5}\rangle_t$ and $\langle\delta_{0.5} - \delta_{0.1}\rangle_t$. We adjust the mean temperature gradient to

     $$\partial_z\overline{T} = -\nabla_{\rm ad} - H(z; \delta_{0.5,\rm new}, d_w)\Delta\nabla, \tag{A1}$$

     where H is defined in Eqn. 30 and $\Delta\nabla = \nabla_{\rm rad} - \nabla_{\rm ad}$. We also multiply the temperature perturbations and full convective velocity field by $(1 - H(z; 1, 0.05))$. This sets all fluctuations above the nominal Schwarzschild convection zone to zero, thereby avoiding any strange dynamical transients caused by the old dynamics at the radiative-convective boundary (which has moved as a result of this process).

5. Return to step 1.

In general, the initial profile of $\overline{T}$ that we use when we start our simulations is given by Eqn. A1 with a value $\delta_{0.5,\rm new} \geq 0$. We then evolve $\overline{T}$ towards a statistically stationary state using the above algorithm and standard timestepping. If a simulation returns to step 2 from step 4 ten times over the course of its evolution, we assume that it has converged near its answer, stop this iterative loop, and allow the simulation to timestep normally. Additionally, in some simulations, we ensure that this process occurs no more than 25 times. This process effectively removes the long diffusive thermal evolution on display in the upper left panel of Fig. 4 by immediately setting the mean temperature profile to the radiative profile above the PZ.

In Fig. 9, we display the time evolution of $\delta_{\rm p}$ and f in Case I simulations with $\mathcal{S} = 10^3$, $\mathcal{R} = 400$, and $\mathcal{P}_D = [1, 2, 4]$ simulations using black lines. We overplot the evolution of simulations which use this accelerated evolution (AE) procedure using orange and green lines. Time units on the x-axis are normalized in terms of the total simulation run time in order to more thoroughly demonstrate the evolutionary differences between standard timestepping and AE. However, the AE simulations are much shorter: the vertical green-and-yellow lines demonstrate how long the AE simulation ran compared to the standard timestepping simulation (so for $\mathcal{P}_D = 1$, the AE simulations only took $\sim 1/4$ as long; for $\mathcal{P}_D = 2$, they took $\sim 1/8$ as long; for $\mathcal{P}_D = 4$, they took $\sim 1/20$ as long). AE simulations with orange lines start with PZ heights which are much larger than the final height, while green line solution start with initial PZ heights which are smaller than the expected height. Regardless of our choice of initial condition, we find that this AE procedure quickly evolves our simulations to within a few percent of the final value. After converging to within a few percent of the proper penetration zone height, this AE procedure continues to iteratively "jitter" around the right answer until the convergence criterion we described above are met. These jitters can be seen in the top panels of Fig. 9, where the solution jumps away from the proper answer in one AE iteration before jumping back towards it in the next iteration. If the PZ height continues to noticeably vary on timescales of a few hundred freefall times, we continue to timestep the simulations until the changes of $\delta_{\rm p}$ have diminished.

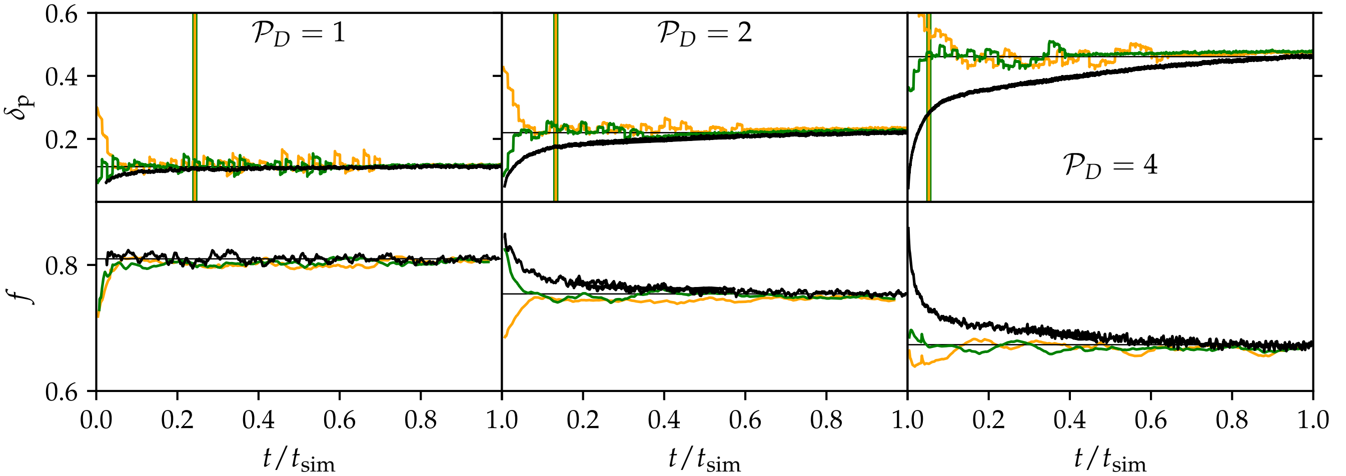

Figure 9. (top row) Time traces of $\delta_{0.5}$ for simulations using standard timestepping (black lines), accelerated evolution with large initial values of δ_p (orange lines), and accelerated evolution with small initial values of δ_p (green lines). The thin horizontal lines represent the final converged values from the standard timestepping simulations. Accelerated evolution timesteps can be seen as jumps in the δ_p trace. Once the procedure gets to within a few percent of the appropriate value, any large steps away from the "correct" solution are quickly reversed. Time units are normalized by the total run time of the simulation. All accelerated evolution simulations were run for $t_{\rm sim} = 3000$ freefall times. The standard timestepping (black line) simulations were run for $t_{\rm sim} = 1.2 \times 10^4$ ($\mathcal{P}_D = 1$), $t_{\rm sim} = 2.2 \times 10^4$ ($\mathcal{P}_D = 2$), and $t_{\rm sim} = 5.7 \times 10^4$ ($\mathcal{P}_D = 4$) freefall times. The vertical green-and-yellow bars show the total simulation time of the accelerated simulation in terms of the direct simulation time; i.e., the accelerated simulation converged in only $\sim 5\%$ of the simulation time of the direct simulation for $\mathcal{P} = 4$. (Bottom row) Rolling average of f over 200 freefall times, plotted in the same way as $\delta_{0.5}$.

B. MESA IMPLEMENTATION

Here we describe a first implementation of Eqn. 40 in MESA (Paxton et al. 2011, 2013, 2015, 2018, 2019). We note that this impelementation is likely not universal or robust enough to be used in any stellar model, but it is robust enough to timestep stably and produce the results displayed in Sct. 6. Future work should improve upon this model.

To find the extent of the penetrative region, we override MESA's TODO routine for writing ∇ and instead have the solver follow these steps:

1. At each timestep, we find the location of the Schwarzschild convective boundary, r_s. We measure the value of the mixing length λ there, and choose a convection zone length scale $\ell_{\rm CZ} = \min(\lambda, L_{\rm CZ})$, where $L_{\rm CZ}$ is the full depth of the Schwarzschild-unstable convective region. We take the "CZ" integrals in Eqn. 40 to span radially from the Schwarzschild boundary to a depth $\ell_{\rm CZ}$ within the convection zone.

2. We calculate the buoyancy integral $\int_{\rm CZ} L_{\rm conv} dr$ and the convection zone volume $V_{\rm CZ}$.

3. Starting at the Schwarzschild boundary, we iteratively expand the size of the PZ until Eqn. 40 is satisfied. For each radial bin i below the convection zone, we calculate $\int_{\rm PZ,i} L_{\rm conv} dr$ and $V_{\rm PZ,i}$. We evaluate the LHS of Eqn. 40 using those values. If LHS $\geq$ RHS, we assume that the radial coordinate of bin i is the boundary of the penetrative region.

4. We enforce $\nabla \approx \nabla_{\rm ad}$ in the PZ by doing TODO.

5. We ensure that the transition from the PZ to the RZ is smooth. Outside of the PZ, we model ∇ as TODO.

Using this procedure with $f = 0.8$ and $\xi = 1$, and timestepping a solar model to the age of the current Sun (~ 4.5 Gyr), we find the profile displayed in Sec. 6.

Table 1. Table of simulation information.

Type	$\mathcal{P}$	$\mathcal{S}$	$\mathcal{R}$	$nz \times (nx \times ny)$	$t_{\rm sim}$	$(\delta_{0.1}, \delta_{0.5}, \delta_{0.9})$	f	ξ	$\langle u \rangle$
"Standard timestepping" simulations									
D	1	10^3	4×10^2	$256{\times}64^2$	12347	$(0.078, 0.112, 0.136)$	0.81	0.70	0.61
D	2	10^3	4×10^2	$256{\times}64^2$	22593	$(0.191, 0.220, 0.242)$	0.75	0.62	0.62
D	4	10^3	4×10^2	$256{\times}64^2$	57170	$(0.434, 0.461, 0.483)$	0.67	0.58	0.62
"Accelerated Evolution" simulations									
D	0.1,	10^3	4×10^2	$256{\times}64^2$	4561	$(0.017, 0.042, 0.069)$	0.83	0.77	0.59
D	0.3,	10^3	4×10^2	$256{\times}64^2$	4681	$(0.030, 0.064, 0.092)$	0.81	0.80	0.62
D	1,	10^3	4×10^2	$256{\times}64^2$	3000	$(0.082, 0.116, 0.140)$	0.80	0.69	0.62
D	2,	10^3	4×10^2	$256{\times}64^2$	5000	$(0.199, 0.228, 0.252)$	0.75	0.60	0.64
D	4,	10^3	4×10^2	$256{\times}64^2$	5000	$(0.452, 0.480, 0.505)$	0.67	0.55	0.62
D	6,	10^3	4×10^2	$256{\times}64^2$	6000	$(0.620, 0.647, 0.667)$	0.64	0.53	0.60
D	8,	10^3	4×10^2	$512{\times}128^2$	4357	$(0.731, 0.757, 0.778)$	0.64	0.48	0.59
D	10,	10^3	4×10^2	$512{\times}128^2$	4226	$(0.857, 0.884, 0.903)$	0.63	0.45	0.59
D	4	10^3	25	$256{\times}16^2$	3000	(0.321, 0.379, 0.437)	0.77	0.27	0.34
D	4	10^3	50	$256{\times}32^2$	3000	(0.398, 0.442, 0.487)	0.73	0.36	0.42
D	4	10^3	1×10^2	$256{\times}32^2$	3000	(0.469, 0.503, 0.534)	0.67	0.46	0.48
D	4	10^3	2×10^2	$256{\times}64^2$	3000	(0.485, 0.515, 0.542)	0.65	0.55	0.55
D	4	10^3	8×10^2	$256{\times}128^2$	3000	(0.407, 0.434, 0.455)	0.69	0.57	0.68
D	4	10^3	1.6×10^3	$256{\times}128^2$	3000	(0.366, 0.397, 0.419)	0.71	0.57	0.72
D	4	10^3	3.2×10^3	256^3	2469	(0.313, 0.350, 0.372)	0.73	0.60	0.75
D	4	10^3	6.4×10^3	384^3	414	(0.277, 0.314, 0.335)	0.74	0.62	0.76
D/SF	4	10^3	50	$256{\times}16^2$	5000	(0.435, 0.477, 0.516)	0.68	0.42	0.50
D/SF	4	10^3	1×10^2	$256{\times}32^2$	5000	(0.482, 0.516, 0.547)	0.64	0.54	0.57
D/SF	4	10^3	2×10^2	$256{\times}32^2$	5000	(0.490, 0.520, 0.547)	0.63	0.59	0.64
D/SF	4	10^3	4×10^2	$256{\times}64^2$	8000	(0.474, 0.502, 0.531)	0.65	0.59	0.69
D/SF	4	10^3	8×10^2	$256{\times}128^2$	5000	(0.410, 0.437, 0.461)	0.68	0.59	0.73
D/SF	4	10^3	1.6×10^3	$256{\times}128^2$	5710	(0.368, 0.400, 0.426)	0.70	0.59	0.76
D/SF	4	10^3	3.2×10^3	256^3	2933	(0.315, 0.352, 0.381)	0.73	0.59	0.77
D	4	10^2	4×10^2	$256{\times}128^2$	5000	(0.377, 0.505, 0.581)	0.65	0.53	0.62
D	4	3×10^2	4×10^2	$256{\times}128^2$	5000	(0.420, 0.477, 0.514)	0.66	0.55	0.62
D	4	3×10^3	4×10^2	$256{\times}128^2$	1170	(0.455, 0.472, 0.487)	0.66	0.58	0.62
L	0.01	10^3	8×10^2	$256{\times}128^2$	1139	(0.017, 0.030, 0.052)	0.87	0.79	0.45
L	0.03	10^3	8×10^2	$256{\times}128^2$	929	(0.021, 0.045, 0.070)	0.86	0.79	0.45
L	0.1	10^3	8×10^2	$256{\times}128^2$	1142	(0.084, 0.076, 0.102)	0.85	0.72	0.45
L	0.3	10^3	8×10^2	$256{\times}128^2$	1109	(0.077, 0.130, 0.158)	0.82	0.66	0.45
L	1	10^3	8×10^2	$256{\times}128^2$	3000	(0.182, 0.225, 0.251)	0.79	0.60	0.44
L	2	10^3	8×10^2	$256{\times}128^2$	3000	(0.278, 0.315, 0.340)	0.76	0.57	0.44
L	4	10^3	8×10^2	$256{\times}128^2$	5000	(0.396, 0.427, 0.449)	0.73	0.53	0.43
L	8	10^3	8×10^2	$256{\times}128^2$	5000	(0.519, 0.545, 0.562)	0.72	0.48	0.42
L	16	10^3	8×10^2	$256{\times}128^2$	8000	(0.687, 0.709, 0.723)	0.70	0.44	0.42

NOTE— Simulation type is specified as "D" for discontinuous/Case I or "L" for linear/Case II. "D/SF" simulations have stress-free boundary conditions. Input control parameters are listed for each simulation: the penetration parameter $\mathcal{P}$, stiffness $\mathcal{S}$, and freefall Reynolds number $\mathcal{R}$. We also note the coefficient resolution (Chebyshev coefficients nz and Fourier coeficients nx, ny). We report the number of freefall time units each simulation was run for $t_{\rm sim}$. Time-averaged values of the departure heights ($\delta_{0.1}$, $\delta_{0.5}$, $\delta_{0.9}$), the dissipation fraction f, and the dissipation fall-off ξ, as well as the average convection zone velocity $\langle u \rangle$ are reported. We take these time averages over the final 1000 freefall times of the simulation. For short simulations ($t_{\rm sim} < 1000$), we average over most of the simulation run time, but not early dynamical transients.

C. TABLE OF SIMULATION PARAMETERS

Input parameters and summary statistics of the simulations presented in this work are shown in Table 1.

REFERENCES

Aerts, C., Christensen-Dalsgaard, J., & Kurtz, D. W. 2010, Asteroseismology (Springer)

Ahlers, G., Grossmann, S., & Lohse, D. 2009, Reviews of Modern Physics, 81, 503, doi: 10.1103/RevModPhys.81.503

Anders, E. H., Brown, B. P., & Oishi, J. S. 2018, Physical Review Fluids, 3, 083502, doi: 10.1103/PhysRevFluids.3.083502

Asplund, M., Amarsi, A. M., & Grevesse, N. 2021, arXiv e-prints, arXiv:2105.01661. https://arxiv.org/abs/2105.01661

Bahcall, J. N., Serenelli, A. M., & Basu, S. 2005, ApJL, 621, L85, doi: 10.1086/428929

Basu, S. 2016, Living Reviews in Solar Physics, 13, 2, doi: 10.1007/s41116-016-0003-4

Basu, S., & Antia, H. M. 2004, ApJL, 606, L85, doi: 10.1086/421110

Basu, S., Antia, H. M., & Narasimha, D. 1994, MNRAS, 267, 209, doi: 10.1093/mnras/267.1.209

Bergemann, M., & Serenelli, A. 2014, Solar Abundance Problem, 245–258, doi: 10.1007/978-3-319-06956-2_21

Böhm-Vitense, E. 1958, ZA, 46, 108

Browning, M. K., Brun, A. S., & Toomre, J. 2004, ApJ, 601, 512, doi: 10.1086/380198

Brummell, N. H., Clune, T. L., & Toomre, J. 2002, ApJ, 570, 825, doi: 10.1086/339626

Brun, A. S., & Browning, M. K. 2017, Living Reviews in Solar Physics, 14, 4, doi: 10.1007/s41116-017-0007-8

Brun, A. S., Strugarek, A., Varela, J., et al. 2017, ApJ, 836, 192, doi: 10.3847/1538-4357/aa5c40

Burns, K. J., Vasil, G. M., Oishi, J. S., Lecoanet, D., & Brown, B. P. 2020, Physical Review Research, 2, 023068, doi: 10.1103/PhysRevResearch.2.023068

Carlos, M., Meléndez, J., Spina, L., et al. 2019, MNRAS, 485, 4052, doi: 10.1093/mnras/stz681

Caswell, T. A., Droettboom, M., Lee, A., et al. 2021, matplotlib/matplotlib: REL: v3.3.4, v3.3.4, Zenodo, doi: 10.5281/zenodo.4475376

Christensen-Dalsgaard, J. 2021, Living Reviews in Solar Physics, 18, 2, doi: 10.1007/s41116-020-00028-3

Christensen-Dalsgaard, J., Monteiro, M. J. P. F. G., Rempel, M., & Thompson, M. J. 2011, MNRAS, 414, 1158, doi: 10.1111/j.1365-2966.2011.18460.x

Claret, A., & Torres, G. 2018, ApJ, 859, 100, doi: 10.3847/1538-4357/aabd35

Couston, L. A., Lecoanet, D., Favier, B., & Le Bars, M. 2017, Physical Review Fluids, 2, 094804, doi: 10.1103/PhysRevFluids.2.094804

Cristini, A., Meakin, C., Hirschi, R., et al. 2017, MNRAS, 471, 279, doi: 10.1093/mnras/stx1535

Currie, L. K., & Browning, M. K. 2017, ApJL, 845, L17, doi: 10.3847/2041-8213/aa8301

Deardorff, J. W., Willis, G. E., & Lilly, D. K. 1969, Journal of Fluid Mechanics, 35, 7, doi: 10.1017/S0022112069000942

Dumont, T., Palacios, A., Charbonnel, C., et al. 2021, A&A, 646, A48, doi: 10.1051/0004-6361/202039515

Farmer, R., Renzo, M., de Mink, S. E., Marchant, P., & Justham, S. 2019, ApJ, 887, 53, doi: 10.3847/1538-4357/ab518b

Fuentes, J. R., & Cumming, A. 2020, Physical Review Fluids, 5, 124501, doi: 10.1103/PhysRevFluids.5.124501

Garaud, P. 2021, Physical Review Fluids, 6, 030501, doi: 10.1103/PhysRevFluids.6.030501

Goluskin, D. 2016, Internally heated convection and Rayleigh-Bénard convection (Springer International Publishing), arXiv:1506.01656, doi: 10.1007/978-3-319-23941-5

Hansen, C. J., Kawaler, S. D., & Trimble, V. 2004, Stellar interiors : physical principles, structure, and evolution (Springer)

Herwig, F. 2000, A&A, 360, 952. https://arxiv.org/abs/astro-ph/0007139

Higgins, E. R., & Vink, J. S. 2020, A&A, 635, A175, doi: 10.1051/0004-6361/201937374

Higl, J., Müller, E., & Weiss, A. 2021, A&A, 646, A133, doi: 10.1051/0004-6361/202039532

Horst, L., Hirschi, R., Edelmann, P. V. F., Andrassy, R., & Roepke, F. K. 2021, arXiv e-prints, arXiv:2107.02199. https://arxiv.org/abs/2107.02199

Hotta, H. 2017, ApJ, 843, 52, doi: 10.3847/1538-4357/aa784b

Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90, doi: 10.1109/MCSE.2007.55

Hurlburt, N. E., Toomre, J., & Massaguer, J. M. 1986, ApJ, 311, 563, doi: 10.1086/164796

Hurlburt, N. E., Toomre, J., Massaguer, J. M., & Zahn, J.-P. 1994, ApJ, 421, 245, doi: 10.1086/173642

Jermyn, A. S., Tout, C. A., & Chitre, S. M. 2018, MNRAS, 480, 5427, doi: 10.1093/mnras/sty1831

Jørgensen, A. C. S., & Weiss, A. 2019, MNRAS, 488, 3463, doi: 10.1093/mnras/stz1980

Käpylä, P. J. 2019, A&A, 631, A122, doi: 10.1051/0004-6361/201834921

—. 2021, arXiv e-prints, arXiv:2105.08453. https://arxiv.org/abs/2105.08453

Käpylä, P. J., Korpi, M. J., Stix, M., & Tuominen, I. 2007, in Convection in Astrophysics, ed. F. Kupka, I. Roxburgh, & K. L. Chan, Vol. 239, 437–442, doi: 10.1017/S1743921307000865

Käpylä, P. J., Rheinhardt, M., Brandenburg, A., et al. 2017, ApJL, 845, L23, doi: 10.3847/2041-8213/aa83ab

Kitiashvili, I. N., Kosovichev, A. G., Mansour, N. N., & Wray, A. A. 2016, ApJL, 821, L17, doi: 10.3847/2041-8205/821/1/L17

Korre, L., Garaud, P., & Brummell, N. H. 2019, MNRAS, 484, 1220, doi: 10.1093/mnras/stz047

Maeder, A. 1975, A&A, 40, 303

Marcus, P. S., Press, W. H., & Teukolsky, S. A. 1983, ApJ, 267, 795, doi: 10.1086/160915

Martinet, S., Meynet, G., Ekström, S., et al. 2021, A&A, 648, A126, doi: 10.1051/0004-6361/202039426

Meakin, C. A., & Arnett, D. 2007, ApJ, 667, 448, doi: 10.1086/520318

Moore, D. R., & Weiss, N. O. 1973, Journal of Fluid Mechanics, 61, 553, doi: 10.1017/S0022112073000868

Musman, S. 1968, Journal of Fluid Mechanics, 31, 343, doi: 10.1017/S0022112068000194

Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3, doi: 10.1088/0067-0049/192/1/3

Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208, 4, doi: 10.1088/0067-0049/208/1/4

Paxton, B., Marchant, P., Schwab, J., et al. 2015, ApJS, 220, 15, doi: 10.1088/0067-0049/220/1/15

Paxton, B., Schwab, J., Bauer, E. B., et al. 2018, ApJS, 234, 34, doi: 10.3847/1538-4365/aaa5a8

Paxton, B., Smolec, R., Schwab, J., et al. 2019, ApJS, 243, 10, doi: 10.3847/1538-4365/ab2241

Pedersen, M. G., Aerts, C., Pápics, P. I., et al. 2021, arXiv e-prints, arXiv:2105.04533. https://arxiv.org/abs/2105.04533

Pinsonneault, M. 1997, ARA&A, 35, 557, doi: 10.1146/annurev.astro.35.1.557

Pratt, J., Baraffe, I., Goffrey, T., et al. 2017, A&A, 604, A125, doi: 10.1051/0004-6361/201630362

Rempel, M. 2004, ApJ, 607, 1046, doi: 10.1086/383605

Rogers, T. M., & Glatzmaier, G. A. 2005, ApJ, 620, 432, doi: 10.1086/423415

Rogers, T. M., Glatzmaier, G. A., & Jones, C. A. 2006, ApJ, 653, 765, doi: 10.1086/508482

Roxburgh, I. W. 1978, A&A, 65, 281

—. 1989, A&A, 211, 361

—. 1992, A&A, 266, 291

—. 1998, Ap&SS, 261, 43, doi: 10.1023/A:1002029105802

Saikia, E., Singh, H. P., Chan, K. L., Roxburgh, I. W., & Srivastava, M. P. 2000, ApJ, 529, 402, doi: 10.1086/308249

Salaris, M., & Cassisi, S. 2017, Royal Society Open Science, 4, 170192, doi: 10.1098/rsos.170192

Shaviv, G., & Salpeter, E. E. 1973, ApJ, 184, 191, doi: 10.1086/152318

Singh, H. P., Roxburgh, I. W., & Chan, K. L. 1995, A&A, 295, 703

—. 1998, A&A, 340, 178

Spiegel, E. A., & Veronis, G. 1960, ApJ, 131, 442, doi: 10.1086/146849

Tian, C.-L., Deng, L.-C., & Chan, K.-L. 2009, MNRAS, 398, 1011, doi: 10.1111/j.1365-2966.2009.15178.x

Toppaladoddi, S., & Wettlaufer, J. S. 2018, Physical Review Fluids, 3, 043501, doi: 10.1103/PhysRevFluids.3.043501

Viallet, M., Meakin, C., Arnett, D., & Mocák, M. 2013, ApJ, 769, 1, doi: 10.1088/0004-637X/769/1/1

Viallet, M., Meakin, C., Prat, V., & Arnett, D. 2015, A&A, 580, A61, doi: 10.1051/0004-6361/201526294

Viani, L. S., & Basu, S. 2020, ApJ, 904, 22, doi: 10.3847/1538-4357/abba17

Vinyoles, N., Serenelli, A. M., Villante, F. L., et al. 2017, ApJ, 835, 202, doi: 10.3847/1538-4357/835/2/202

Wang, D., & Ruuth, S. J. 2008, Journal of Computational Mathematics, 26, 838. http://www.jstor.org/stable/43693484

Woosley, S. E., Heger, A., & Weaver, T. A. 2002, Reviews of Modern Physics, 74, 1015, doi: 10.1103/RevModPhys.74.1015

Zahn, J. P. 1991, A&A, 252, 179