



STELLAR MIXING LENGTH THEORY WITH ENTROPY RAIN

AXEL BRANDENBURG^{1,2,3,4}

¹ Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO 80303, USA

² JILA and Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO 80303, USA

³ Nordita, KTH Royal Institute of Technology and Stockholm University, Roslagstullsbacken 23, SE-10691 Stockholm, Sweden

⁴ Department of Astronomy, AlbaNova University Center, Stockholm University, SE-10691 Stockholm, Sweden

Received 2015 April 15; revised 2016 August 20; accepted 2016 August 30; published 2016 November 10

ABSTRACT

The effects of a non-gradient flux term originating from the motion of convective elements with entropy perturbations of either sign are investigated and incorporated into a modified version of stellar mixing length theory (MLT). Such a term, first studied by Deardorff in the meteorological context, might represent the effects of cold intense downdrafts caused by the rapid cooling in the granulation layer at the top of the convection zone of late-type stars. These intense downdrafts were first seen in the strongly stratified simulations of Stein & Nordlund in the late 1980s. These downdrafts transport heat nonlocally, a phenomenon referred to as *entropy rain*. Moreover, the Deardorff term can cause upward enthalpy transport even in a weakly Schwarzschild-stably stratified layer. In that case, no giant cell convection would be excited. This is interesting in view of recent observations, which could be explained if the dominant flow structures were of small scale even at larger depths. To study this possibility, three distinct flow structures are examined: one in which convective structures have similar size and mutual separation at all depths, one in which the separation increases with depth, but their size is still unchanged, and one in which both size and separation increase with depth, which is the standard flow structure. It is concluded that the third possibility with fewer and thicker downdrafts in deeper layers remains the most plausible, but it may be unable to explain the suspected absence of large-scale flows with speeds and scales expected from MLT.

Key words: convection – Sun: granulation – turbulence

1. INTRODUCTION

Late-type stars such as our Sun have outer convection zones. The observed solar granulation is a surface manifestation of their existence. Solar granulation and the first few megameters (Mm) of the convection zone have been modeled successfully using mixing length theory (MLT) and numerical simulations with realistic physics included (Stein & Nordlund 1989, 1998; Vögler et al. 2005; Gudiksen et al. 2011; Freytag et al. 2012). As a function of depth, the simulations reproduce some essential features predicted by MLT, in particular the depth dependence of the turbulent root mean square (rms) velocity, $u_{\text{rms}} \approx (F_{\text{enth}}/\rho)^{1/3}$, where F_{enth} is the enthalpy flux and ρ is the density. Simulations also seem to confirm an important *hypothesis* of MLT regarding the gradual increase of the typical convective time and length scales with depth. Given such agreements, there was never any reason to question our basic understanding of convection.

In recent years, local helioseismology has allowed us to determine subsurface flow velocities (Duvall et al. 1997; Gizon & Birch 2005). Helioseismic observations by Hanasoge et al. (2010, 2012) using the deep-focusing time–distance technique have not, however, detected large-scale convection velocities at the expected levels (Miesch et al. 2012); see also Gizon & Birch (2012) for a comparison with both global simulations and radiation hydrodynamics simulations with realistic radiation and ionization physics included. Greer et al. (2015) have suggested that the approach of Hanasoge et al. (2012) may remove too much signal over the time span of the measurements. Using ring-diagram analysis (Gizon & Birch 2005) with appropriately assembled averaging kernels to focus on the deeper layers, Greer et al. (2015) instead find values of the turbulent rms velocities that are consistent with conventional wisdom. Moreover, they find that at large length scales

corresponding to spherical harmonic degrees of 30 or below (corresponding to scales of about 140 Mm and larger), the rms velocity actually *increases* with depth. This in itself is remarkable, because velocity perturbations from deeper layers were expected to be transmitted to the surface almost unimpededly (Stix 1981), unlike perturbations in the heat flux, which are efficiently being screened by the convection (Spruit 1977). Thus, an increase in horizontal velocity power with depth is unexpected. However, van Ballegoijen (1986) pointed out that for the scale and depth of giant cells, the screening might be large enough to allow giant cell convection of 100 m s^{-1} to result in only 10 m s^{-1} at the surface, which would be compatible with observations (Hathaway et al. 2013). The presence of the near-surface shear layer of the Sun (Schou et al. 1998) might enhance the screening further.

Given that the deeply focused kernels used by Greer et al. (2015) can still have 5%–10% sensitivity to Doppler shifts arising from flows in the upper layers, such near-surface flows could still leave an imprint on the signal. This effect would be exaggerated further if the near-surface flows were stronger compared to those in deeper layers, as was theoretically expected. On the other hand, if convection in the surface layers is of smaller scale, the signal from those layers will to a large extent be averaged out despite its larger amplitude. It would obviously be important to examine this more thoroughly by applying the kernels of Greer et al. (2015) to realistic simulations. Unfortunately, this has not yet been attempted. Conversely, the deep-focusing technique of Hanasoge et al. (2012) could be extended to allow for imaging of the deeper flow structures and thereby a direct comparison of individual turbulent eddies with those detected by Greer et al. (2015).

It should be mentioned that small flow speeds of giant cell convection have been found by correlation tracking of supergranule proper motions (Hathaway et al. 2013). The

typical velocities are of the order of 20 m s^{-1} at spherical degrees of around 10. Those speeds are still an order of magnitude above the helioseismic upper limits of Hanasoge et al. (2012), but Hanasoge & Sreenivasan (2014) argue that these surface measurements translate to lower flow speeds in the deeper and denser layers when considering mass conservation. This argument may be too naive and would obviously be in conflict with the results of Greer et al. (2015), which show an *increase* with depth at low spherical harmonic degrees. On the other hand, Bogart et al. (2015) have argued that the flows reported by Hathaway et al. (2013) may be non-convective in nature and, in fact, magnetically driven and perhaps related to the torsional oscillations.

Quantitatively, the horizontal flow speed at different scales is characterized by the kinetic energy spectrum, $E(k)$, where k is the wavenumber or inverse length scale. Stein et al. (2007) have shown that the spectra of surface Dopplergrams and correlation tracking collapse with those of Stein & Nordlund onto a single graph such that the spectral velocity $[kE(k)]^{1/2}$ is proportional to k , i.e., $E(k) \propto k$; see also Nordlund et al. (2009). This is a remarkable agreement between simulations and observations. While the origin of such a spectrum is theoretically not understood, it should be emphasized that these statements concern the horizontal velocity at the surface and do not address the controversy regarding the spectrum at larger depths, where the flow speeds at large length scales may still be either larger (e.g., Greer et al. 2015) or smaller (e.g., Featherstone & Hindman 2016) than those at the surface. If most of the kinetic energy were to reside on small scales throughout the convection zone even at a depth of several tens of megameters and beneath, $E(k)$ is expected to decrease toward smaller wavenumbers k either like white noise $\propto k^2$ or maybe even with a Batchelor spectrum $\propto k^4$ (Davidson 2004). However, even in simulations of forced isothermally stratified turbulence, in which there is no larger-scale driving from thermal buoyancy, there is a shallower spectrum, $E(k) \propto k^{3/2}$ (Losada et al. 2013). By contrast, if one imagines the flow to be anelastic so that the mass flux $\rho \mathbf{u} = \nabla \times \psi$ can be written as a vectorial stream function ψ , and if ψ is given by white noise, one should expect a k^4 Batchelor spectrum. This is qualitatively similar to the results of granule tracking, which reveal an intermediate scaling proportional to k^3 (Rieutord et al. 2008; Roudier et al. 2012). While the steeper k^4 spectrum for $k < k_f$ might leave some hope that the results of Hanasoge et al. (2012) could be reconciled with these theoretical and observational constraints, this would be virtually impossible for the linear energy spectrum, $E(k) \propto k$.

The issue has recently been examined by Lord et al. (2014), who have shown that simulation results with the MURaM code (Vögler et al. 2005) can be reproduced with a model that is composed of a continuous hierarchy of layers, each with its own driving scale at a scale of four local density scale heights. At wavenumbers below that driving scale, the spectral power falls off in a certain way. The horizontal surface spectrum is a superposition of these contributions, each of which is assumed to decay with height, contributing therefore progressively less with depth. However, both their model and the MURaM results (Lord 2014) show an order of magnitude more power (factor 2.5 in rms velocity) than the Sun at small wavenumbers. To understand this, they show that the observations can be reproduced if below a depth of 10 Mm the convective energy flux is augmented by an artificially added flux term, which thus

significantly reduces the rms velocity of the resolved flow field. They speculate that the flow-temperature correlation entering the enthalpy flux may be larger in the Sun, possibly caused by a magnetic field that may maintain flow correlations and boost the transport of convective flux by smaller scales. This has been partially confirmed by Hotta et al. (2015), who have discussed the possible role of small-scale magnetic fields in suppressing the formation of large-scale flows. Furthermore, Featherstone & Hindman (2016) have found that with increasing Rayleigh numbers, there is more kinetic energy at small scales and less at large scales such that the total kinetic energy is unchanged by this rearrangement of energy.

Given the importance of this subject, it is worthwhile reviewing possible shortcomings in our theoretical understanding and numerical modeling of stellar convection. Both global and local surface simulations of solar-type convection would become numerically unstable with just the physical values of viscosity and radiative diffusivity. Therefore, the radiative flux (in the optically thick layers) is modified in one of two possible ways. (i) The contribution from temperature fluctuations is greatly enhanced, i.e.,

$$F_{\text{rad}} = -K \nabla T \rightarrow -K \nabla \bar{T} - K_{\text{SGS}} \nabla (T - \bar{T}), \quad (1)$$

where T and \bar{T} are the actual and horizontally averaged temperatures, respectively, K is the radiative conductivity, and K_{SGS} is a subgrid scale (SGS) conductivity. The latter is enhanced by many orders of magnitude relative to K . Furthermore, numerical diffusion operators often do not translate in any obvious way to the physical operators. (ii) Alternatively, in *direct* numerical simulations (DNS), one uses physical viscosity and diffusivity operators, i.e., the replacement in Equation (1) is not invoked, but the coefficients are enhanced ($K \rightarrow K_{\text{enh}}$) and exceed the physical ones by many orders of magnitude. Both approaches are problematic.

In many global DNS with enhanced coefficients (e.g., Käpylä et al. 2013), the lower boundary is closed, so at the bottom of the domain all the energy is carried by radiation alone. By choosing an enhanced radiative diffusivity, the radiative flux is increased by a corresponding amount, and therefore also the total flux. The luminosity in those simulations can exceed the solar value by several orders of magnitude. This has a series of consequences. Most notably, the convective velocities are too high in the upper parts where most of the flux is carried by convection (Käpylä et al. 2013). There are two ways to avoid this problem. One is to use simulations with a polytropic hydrostatic reference solution, whose polytropic index is close enough to the adiabatic one so that the convective flux is everywhere a small fraction of the radiative one (Brandenburg et al. 2005), which reduces the convective velocities. Alternatively, one applies the enhanced radiative conductivity only to the temperature fluctuations so as not to disturb the very small radiative energy flux, $-K \nabla T$, compared with $-K_{\text{enh}} \nabla T$, which it would have been in the DNS approach without subtracting \bar{T} . The temperature smoothing implied by invoking alternative (i) is also necessary in the simulations with realistic opacities, because the Péclet number (which is similar to the Reynolds number) based on the physical value of K would reach values above 10^{10} , which cannot be handled by a DNS (Barekat & Brandenburg 2014). In the global simulations, it is common to use the specific entropy gradient instead of $\nabla (T - \bar{T})$ (e.g., Käpylä et al. 2013). This is equivalent if \bar{T} is close to the adiabatic value. However, the diffusion coefficient in the SGS term can easily be five orders of magnitude

larger than the physical one acting on the mean stratification. This has the consequence of suppressing small-scale turbulent flows and entropy fluctuations, which may prematurely damp the low-entropy fluid parcels that originate within the strong cooling layer at the surface and which will be discussed as *entropy rain* in the bulk of this paper. If such low-entropy contributions are poorly captured by the simulation, this could be compensated for by a sufficiently strong contribution proportional to the superadiabatic gradient, which in turn would give rise to the excitation of flows on much larger scales than would be compatible with the observations discussed by Lord et al. (2014). The sensitivity of large-scale flow excitation to small changes in the superadiabatic gradient was also discussed by Cossette & Rast (2016), who studied models with different types of surface driving. It is further supported by the work of Featherstone & Hindman (2016), who found that there is more kinetic energy at small scales and less at large scales as the Rayleigh number increases.

There is yet another problem that concerns the global simulations in which a predetermined profile $K(z)$ is used instead of calculating its local value with a physical opacity. Such an approach has been used routinely in studies of compressible convection, especially when stably and unstably stratified layers are combined (Hurlburt et al. 1986; Brandenburg et al. 1996). The simulations of Käpylä et al. (2013) adopted a profile that yields a negative (unstable) radial entropy gradient through most of the layer (see the inset of their Figure 3). However, as will be pointed out below, it is only a tiny surface layer in which the non-convective model is unstable. The rest of the model is a priori stable to convection, but it becomes marginally stable (or perhaps slightly unstable) as a consequence of the resulting turbulence leading to bulk mixing across the deeper layers. This is quite contrary to the models with a predefined $K(z)$ profile, which would be unstable by construction over the entire depth of the convection zone.

We should emphasize that the non-convective solution is mainly of academic interest. It is used to compute, for example, the Rayleigh number, a measure of the degree of instability. However, there is no doubt that the actual stratification of the Sun is close to marginal stability down to a fractional radius of 0.71 before turning decidedly stable, as confirmed by helioseismology (Christensen-Dalsgaard et al. 1991; Basu 1997). Nevertheless, the concept of convection being driven by surface cooling rather than heating from below has been promoted in a number of papers by Stein & Nordlund (1989, 1998) and elaborated upon by Spruit (1997), who introduced the idea that flows in the deeper parts are being driven nonlocally through the entropy rain from the surface. The question is to what extent this affects our understanding of the speed and especially the typical scales of convective motions, how MLT models would need to be modified, and whether this might have any bearing on the interpretation of the observed flow amplitudes at large scales.

To include the nonlocal effects described by Spruit (1997), we must look for a contribution to the enthalpy flux that is not related to the local entropy gradient.⁵ Such a term has been identified in the meteorological context by Deardorff (1966), who describes it as a counter-gradient flux. In Deardorff

(1972), he derives an expression for this flux, which depends on the local temperature or entropy fluctuations which, in his case, come from measurements. In the present case, we assume that such fluctuations have their origin in what Spruit (1997, p. 406) refers to as *threads*, which are thin downdrafts on an almost perfectly isentropic passive fluid upflow between the threads. Spruit already emphasized back then the dynamical importance of the “forest of narrow cool threads that are produced at the solar surface.” He further suggested that their absence in simulations with a fixed top boundary at some depth below the actual surface might be “the reason why they would produce large-scale flows with amplitudes that are about two orders of magnitude larger than observed.”

In the scenario described by Spruit (1997), convection is driven solely within the surface layers. This is superficially reminiscent of convective overshoot, as modeled in some of the aforementioned papers (Hurlburt et al. 1986; Brandenburg et al. 1996), where convective flows are driven into stable layers. The extent of overshoot depends not only on the stiffness of the stable gradient that the convective plumes are flowing into (Hurlburt et al. 1994), but also on the flow speed. The depth of such a layer can therefore not be determined a priori from purely hydrostatic stability considerations. The lower boundary might then not be sharp. As explained above, in a hydrostatic non-convecting reference model, only a very thin surface layer is a priori unstable to convection. The rest is made unstable purely by bulk mixing. If entropy rain convection is a nonlocal phenomenon, the extent of convection should depend on surface properties and cannot be predicted from the local entropy gradient, similarly to convective overshoot. One might then not be able to understand the relatively sharp demarcation at the bottom of the convection zone, as found in global helioseismology. Of course, once convection has become fully developed, the low-entropy elements descend into buoyantly neutral layers, which is quite different from the usual overshoot. Furthermore, the usual overshoot layer is characterized by negative buoyancy and therefore a downward enthalpy flux (Hurlburt et al. 1986). An important purpose of the present paper is to produce a quantitative model and to demonstrate that, with a hypothetical nonlocal contribution to the flux, it is possible to obtain models that still have a sharply defined lower boundary.

A quantitative model of convection with entropy rain, even with the somewhat hypothetical Deardorff term included, would also be useful to illustrate the qualitative nature of the resulting stratification, and to show whether it is weakly super or subadiabatic. Indeed, as already argued by Spruit (1997), if the local mass fraction of entropy deficient material, descending from the cooling surface, decreases with depth and if the stratification outside the entropy rain were exactly isentropic, the resulting horizontally averaged entropy would *increase* with depth. This suggests that the entropy rain itself can make an otherwise vanishing radial entropy gradient negative and therefore Schwarzschild unstable, as seen in surface simulations. This raises two important questions. First, to what extent is such a stratification affected by radiative heating—especially toward the bottom of the convection zone, where it would tend to produce a positive (stable) mean entropy gradient in the upflows. Second, would such a negative (unstable) mean entropy gradient always lead to giant cell convection, as has been seen in global convection simulations (Miesch

⁵ As will be discussed in Section 6, Spruit (1997) argues that the low-entropy material from the top always leads to a negative mean entropy gradient. However, here we argue that the mean entropy gradient is only one contribution to a mean-field (here one-dimensional) parameterization of the enthalpy flux, and that there is another one that is not locally connected with it.

et al. 2008), or could the stratification still be stable to the excitation of large-scale flows?

We postpone addressing the two aforementioned issues connected with the qualitative reasoning of Spruit (1997) to the end of the paper (Section 6), and begin by highlighting the less commonly known fact that in the non-convecting reference model, only the layer in the top 1 Mm is convectively unstable (Section 2). We then explain the nature of the Deardorff flux (Section 3), present a correspondingly modified mixing length model (Section 4), and give some illustrative numerical solutions (Section 5). We discuss alternative explanations for the lack of giant cell convection in Section 6 and present our conclusions in Section 7.

2. A HIGHLY UNSTABLE SURFACE LAYER

The main argument for the existence of a highly unstable surface layer comes from the consideration of the associated non-convective reference solution, where the flux is forced to be transported by radiation only, i.e., $F = F_{\text{rad}}$. This is something that is not normally considered in stellar physics, because we know that such a solution would be unstable to convection and would therefore never be realized. However, as mentioned in the introduction, this solution is of certain academic interest. We postpone the discussion of an explicit numerical solution to Section 5 and present in this section the basic argument only at a qualitative level, making reference to earlier numerical calculations by Barekat & Brandenburg (2014) for a simple model. They considered an opacity law of the form

$$\kappa = \kappa_0 (\rho/\rho_0)^a (T/T_0)^b, \quad (2)$$

where a and b are adjustable parameters, ρ_0 and T_0 are reference values for density and temperature, respectively, and κ_0 gives the overall magnitude of the opacity. The essential point to note here is that the exponents a and b determine the gradient of specific entropy in a purely non-convecting reference model. In thermodynamic equilibrium, the radiative flux must be constant, i.e.,

$$F_{\text{rad}} = -K dT/dz = \text{const}, \quad (3)$$

where $K = 16\sigma_{\text{SB}}T^3/(3\kappa\rho)$ is the radiative conductivity with σ_{SB} being the Stefan–Boltzmann constant, and z is the vertical coordinate in a Cartesian coordinate system. For the simple opacity law (2), but with

$$b < 4 + a, \quad (4)$$

the optically thick regime is characterized by a constant temperature gradient and therefore $K = \text{const}$ (see Appendix A). We have then a polytropic stratification with $\rho \propto T^n$, where

$$n = (3 - b)/(1 + a) \quad (5)$$

is the polytropic index. It is larger than -1 when Equation (4) is obeyed, i.e., when the pressure decreases with height, as is required for a physically meaningful solution. For a ratio of specific heats of $\gamma = 5/3$, the value of n for marginal Schwarzschild stability is $n_{\text{crit}} = 1/(\gamma - 1) = 3/2$. In most of the solar convection zone the dominant opacity is the bound-free absorption owing to the absorption of light during ionization of a bound electron, which is well described by

the Kramers-type opacity law with $a = 1$ and $b = -7/2$, and so $n = 3.25$, which corresponds to a Schwarzschild-stable solution. Only near the surface, at temperatures typically below 15,000 K, is the dominant opacity the H^- opacity. It can no longer be approximated by a simple power law of the type given by Equation (2). However, in limited density and temperature ranges, certain values of a and b can tentatively be specified, e.g., $a = 0.5$ and $b = 7 \dots 18$. Clearly, for all these values the constraint (4) is violated, so the hydrostatic stratification is no longer polytropic, but solutions can still be constructed numerically (Barekat & Brandenburg 2014) and they demonstrate, not surprisingly, that the stratification is highly unstable. What is more surprising is the fact that even with a combined opacity law of the form

$$\kappa^{-1} = \kappa_{\text{Kr}}^{-1} + \kappa_{\text{H}^-}^{-1}, \quad (6)$$

where κ_{Kr} and κ_{H^-} are given by Equation (2) with suitable exponents a and b , the solutions of the non-convective reference state is unstable only over a depth of approximately 1 Mm. We return to this at the end of Section 5, where we present numerical solutions.

Of course, as stated earlier, the non-convective reference state is only of academic interest. Already with standard MLT (Biermann 1932; Vitense 1953), which allows for a non-vanishing enthalpy flux, one finds a vastly extended convection zone with a depth of the order of 100 Mm (Biermann 1938). However, the question now is how this can be affected by the presence of the Deardorff flux. This will be the subject of the rest of this paper. If the Deardorff flux were to become dominant and the stratification subadiabatic, it would become locally stable to the onset of convection. One might further speculate that the typical scale would no longer be controlled by the local pressure scale height, but it might be imprinted from the downdraft pattern just beneath the surface and be therefore comparable to the granulation scale or at least the supergranulation scale (Cossette & Rast 2016). This can have other important consequences that will also be addressed in this paper. It should be pointed out, however, that the surface simulations have so far not produced evidence for subadiabatic stratification. On the other hand, the models presented below predict subadiabatic stratification only a certain distance below the surface, depending on ill-known input parameters.

3. THE DEARDORFF FLUX

3.1. Derivation

In the meteorological context, counter-gradient heat flux terms have been noticed for a long time (Ertel 1942; Priestley & Swinbank 1947; Deardorff 1966). They appear naturally when calculating the enthalpy flux F_{enth} using the τ approximation in its minimalistic form (e.g., Blackman & Field 2003). In this approach, one computes the time derivative of F_{enth} . In the absence of ionization effects, F_{enth} can be written as

$$F_{\text{enth}} = \overline{\rho u_z c_p T}, \quad (7)$$

where the overbar denotes horizontal averaging. In standard MLT, the enthalpy flux is usually referred to as the convective flux and the kinetic energy flux vanishes because of the assumed perfect symmetry between up- and downflows. In

deeper layers especially, however, this is not justified (Stein et al. 2009), so this restriction will later be relaxed.

In the case of strongly stratified layers, it is convenient to use pressure P and specific entropy S as thermodynamic variables, because we later want to ignore pressure fluctuations on the grounds that pressure disturbances are quickly equilibrated by sound waves. Furthermore, we will restrict ourselves to second order correlations in Equation (7) and thus to fluctuations only in the correlation of specific entropy and velocity. Up to some reference value, we have

$$S/c_P = \ln T - \nabla_{\text{ad}} \ln P, \quad (8)$$

where $\nabla_{\text{ad}} = 1 - 1/\gamma$ and $\gamma = c_P/c_V$ is the ratio of specific heats at constant pressure and constant volume, respectively, all constants for a perfect gas with a fixed degree of ionization. Ignoring pressure variations and linearizing $\delta \ln T = \delta T/T$, we can replace $c_P \delta T$ by $T \delta S$. In the following, we denote fluctuations by lower case characters, i.e., $S = \bar{S} + s$, where $s \equiv \delta S$ are used interchangeably. As argued above, other fluctuations are omitted. We also ignore mean flows ($\bar{\mathbf{U}} = \mathbf{0}$), so there are only velocity fluctuations ($\mathbf{U} = \mathbf{u}$). Focusing thus on the dominant contribution proportional to $\overline{u_z s}$, we have

$$F_{\text{enth}} = \bar{\rho} \bar{T} \overline{u_z s}. \quad (9)$$

Next, we write the time derivative of F_{enth} as

$$\partial F_{\text{enth}} / \partial t = \bar{\rho} \bar{T} (\overline{\dot{u}_z s} + \overline{u_z \dot{s}}), \quad (10)$$

where dots denote partial time derivatives and changes of the background state have been neglected. Using the governing equations for u_i and s , we have (see Appendix B)

$$\dot{s} = -u_j \nabla_j \bar{S} - s / \tau_{\text{cool}} \dots, \quad (11)$$

$$\dot{u}_i = -g_i s / c_P + \dots, \quad (12)$$

where g_i is the i th component of the gravitational acceleration in Cartesian coordinates, namely $\mathbf{g} = (0, 0, -g)$, the ellipses refer to terms that are nonlinear in the fluctuations,

$$\tau_{\text{cool}}^{-1} = \iota c_\gamma k_f \quad \text{with} \quad \iota = \ell k_f / (3 + \ell^2 k_f^2) \quad (13)$$

is the inverse heating and cooling time owing to radiation (Unno & Spiegel 1966; Edwards 1990), $c_\gamma = 16\sigma_{\text{SB}} \bar{T}^3 / \bar{\rho} c_P$ is the photon diffusion speed (Barekat & Brandenburg 2014), $\ell = 1/\kappa_\rho$ is the photon mean-free path, and k_f is the typical wavenumber of the fluctuations, which may be associated with the inverse mixing length used routinely in MLT. We mention at this point that the nonlocal nature of the entropy rain must lead to yet another contribution in Equation (11). We expect this to be the result of the nonlinear term (indicated by the ellipses), of which a part later gives rise to the ∇_{D} term. This will be motivated further at the end of this section and in Section 3.3, where it will be included in the final expression for ∇_{D} .

Using Equations (11) and (12) in Equation (10), we have

$$\frac{\partial F_{\text{enth}}}{\partial t} = \bar{\rho} \bar{T} (-\overline{u_i u_j} \nabla_j \bar{S} - g_i \overline{s^2} / c_P) - \frac{F_{\text{enth}}}{\tau_{\text{cool}}} + \mathcal{T}_i, \quad (14)$$

where the $\overline{s^2}$ term is the Deardorff flux and the \mathcal{T}_i refer to triple correlations that will be approximated by the quadratic

correlation F_{enth} as

$$\mathcal{T}_i = -F_{\text{enth}}^{\text{ent}} / \tau, \quad (15)$$

where τ is a relaxation time due to turbulence (Blackman & Field 2003), which will later be identified with the turnover time. This procedure, in which correlations with pressure fluctuations are also neglected, is called the minimal τ approximation. Important aspects of the closure assumption (15) have been verified numerically for passive scalar transport (Brandenburg et al. 2004). Among other things, they found that the time derivative on the left-hand side of Equation (14) restores causality by turning the otherwise parabolic heat equation into a hyperbolic wave equation. Here, however, we are interested in slow variations such that the time derivative of F_{enth} can be neglected. (Note, however, that this does not imply that the cooling term will be neglected.)

Since the relaxation term in Equation (15) is similar to the cooling term in Equation (14), we can combine the two by introducing the reduced relaxation time τ_{red} defined through

$$\tau_{\text{red}}^{-1} = \tau_{\text{cool}}^{-1} + \tau^{-1}. \quad (16)$$

We can then solve for F_{enth} , which appears on the right-hand sides (rhs) of Equations (14) and (15), and find $F_{\text{enth}} = F_{\text{G}} + F_{\text{D}}$, where

$$F_{\text{G}} = -\frac{1}{3} \tau_{\text{red}} u_{\text{rms}}^2 \bar{\rho} \bar{T} \nabla \bar{S}, \quad (17)$$

$$F_{\text{D}} = -\tau_{\text{red}} \overline{s^2} \mathbf{g} \bar{\rho} \bar{T} / c_P \quad (18)$$

are the ordinary gradient and the new Deardorff fluxes, respectively, and anisotropies have been ignored for the benefit of simpler notation. Thus, we write $\overline{u_i u_j} \approx \frac{1}{3} \delta_{ij} u_{\text{rms}}^2$, where u_{rms} is the rms velocity of the turbulence. Equations (17) and (18) are written in vectorial forms to highlight the directions of the fluxes: F_{G} is counter-gradient and F_{D} is countergravity. In Equation (17), the term $\frac{1}{3} \tau_{\text{red}} u_{\text{rms}}^2 \equiv \chi_{\text{t}}$ is the turbulent thermal diffusivity.

In view of astrophysical applications, we replace the specific entropy gradient by the commonly defined superadiabatic gradient, i.e.,

$$-d(\bar{S}/c_P)/dz = (\nabla - \nabla_{\text{ad}})/H_P, \quad (19)$$

where $\nabla = d \ln \bar{T} / d \ln \bar{P}$ is the double-logarithmic temperature gradient and $H_P = -(d \ln \bar{P} / dz)^{-1}$ is the pressure scale height. Thus, we arrive at

$$F_{\text{enth}} = \frac{1}{3} \bar{\rho} c_P \bar{T} (\tau_{\text{red}} u_{\text{rms}}^2 / H_P) (\nabla - \nabla_{\text{ad}} + \nabla_{\text{D}}), \quad (20)$$

where ∇_{D} is a new contribution to standard MLT, which results from F_{D} . Using the terms on the rhs of Equation (18), we can write it explicitly as

$$\nabla_{\text{D}} = (3/\gamma) (\overline{s^2} / c_P^2) \text{Ma}^{-2}, \quad (21)$$

where $\text{Ma} = u_{\text{rms}} / c_s$ is the Mach number of the turbulence and c_s is the sound speed with $c_s^2 = \gamma g H_P$. Note that this Mach number dependence arises in order to cancel the corresponding u_{rms}^2 factor in the definition (20). The Deardorff term is a contribution to the flux that is always directed outward and results from the transport of fluid elements with entropy fluctuations of either sign, as is illustrated in Figure 1.

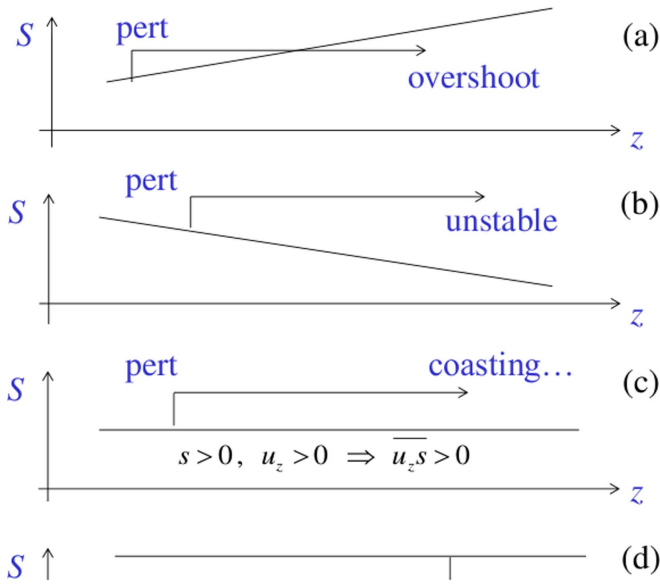


Figure 1. Sketch illustrating overshoot in a stably stratified layer (a), growth of perturbations in an unstably stratified layer (b), buoyant rise of a blob with positive entropy perturbation (c), and descent of a blob with negative entropy perturbation and hence negative buoyancy (d). In cases (b)–(d), the turbulent flux is upward (increasing z). Case (d) is relevant to entropy rain.

3.2. Physical Interpretation

We recall that in a stably (unstably) stratified layer, shown in Figures 1(a) and (b), and under the assumption of pressure equilibrium, the entropy perturbation of a blob with respect to its current surroundings decreases (increases) after a small ascent. By contrast, in the marginally or nearly marginally stratified cases, the perturbation remains constant if one ignores mixing with the surroundings; see panels (c) and (d). Therefore, a positive entropy perturbation ($s > 0$) always corresponds to a positive temperature perturbation and a negative density perturbation. Thus, the fluid parcel is buoyant and moves upward ($u_z > 0$), so $u_z s > 0$, giving a positive contribution to F_{enth} . Likewise, for $s < 0$, the parcel is cooler and heavier and sinks ($u_z < 0$), and again, $u_z s > 0$, giving a positive contribution to F_{enth} .

In laboratory convection, strong positive entropy perturbations can be driven at the lower boundary, and strong negative ones at the top boundary. In the Sun, however, only the strong radiative cooling in the photosphere provides a significant source of entropy perturbations. This became clear with the emergence of the realistic solar convection simulations of Stein & Nordlund (1989). These simulations showed for the first time the strong vertical asymmetry of solar convection resulting from the physics of ionization and strong cooling via the H^- opacity. This led to the notion of *entropy rain* and Spruit’s description of solar convection as a nonlocal phenomenon that is driven solely by surface cooling. Our considerations suggest that the resulting (one-dimensional) mean-field energy flux is parameterized not just in terms of the local superadiabatic gradient.

In analogy with laboratory convection (Heslot et al. 1987; Castaing et al. 1989), Spruit refers to the downdrafts as threads. Interestingly, the laboratory experiments show that these threads persist even at large Rayleigh numbers. For such threads to persist, one could imagine them to be stabilized by their intrinsic vorticity, akin to a Hill vortex (Hill 1894). These

are vortex rings that can stay concentrated over large distances. Vortex rings have also been reported by Stein et al. (2009), but many of those are associated with mushroom shapes, indicating that they widen and soon break up and mix with their surroundings. Numerical studies (see Appendix C) suggest that Hill vortices also survive in a highly stratified isothermal layer, and that their persistence increases significantly with resolution and decreasing viscosity. On the other hand, the presence of background turbulence provides an effective turbulent viscosity, which contributes to a premature break up of small-scale vortices. However, those studies were done with an isothermal equation of state and thus ignore the effects of continued driving by a persistent negative entropy perturbation between the vortex and its surroundings.

Studies of the Deardorff flux in the meteorological context suggest that the flux-carrying plumes are associated with so-called coherent structures (De Roode et al. 2004) and those may not simply be the vortex-like structures envisaged above. It is clear, however, that the structures seen in the Earth’s atmosphere are driven by the boundary layer on the ground, and sometimes also by the upper inversion layer (De Roode et al. 2004). Pleim (2007) discusses the Deardorff flux in connection with other parameterizations of nonlocal fluxes such as the transilient matrix approach (Stull 1984, 1993). The close connection between counter-gradient fluxes and nonlocal transport has been elaborated upon by van Dop & Verver (2001) and Buske et al. (2007). Thus, while the approach presented in Section 3.1 may capture the physical phenomenon of the Deardorff flux qualitatively correctly, it is quite possible that the nature of the underlying coherent structures may require additional refinements.

3.3. Depth Dependence of the Deardorff Term

To estimate the resulting depth dependence of the Deardorff term in Equation (21), we must know how the s^2 associated with the Deardorff term varies with depth. If the entropy rain was purely of the form of Hill vortex-like structures, as discussed above, their filling factor f_s would decrease with increasing depth and density like

$$f_s \propto \bar{\rho}^{-\zeta}, \quad (22)$$

where $\zeta = 0.8$ has been found for spherical vortex structures descending in an isothermally stratified layer; see Appendix C. However, Equation (23) is not contingent on Hill vortices, but is a consequence of downdrafts along a density gradient. For purely spherical compression one would expect $\zeta = 2/3$, while for horizontal compression one has $\zeta = 1$.

If we neglect non-ideal (radiative or viscous) effects, as well as entrainment between up- and downflows, the difference $\Delta S = S_{\uparrow} - S_{\downarrow}$ between the entropy of the slowly rising surroundings, S_{\uparrow} , and that in the downward propagating vortex, S_{\downarrow} , would remain constant and equal to the entropy deficit ΔS_0 suffered by overturning motions at the surface, which is the only location where radiative losses are significant. On sufficiently short length scales, however, radiative heating from the surroundings would erode this entropy difference and lead to a decrease of the effective ΔS . We model this by considering ζ an adjustable parameter that is *increased* relative to the value 0.8 that we expect in the absence of radiation, i.e., f_s decreases more strongly. The resulting fractional entropy

difference, $f_s \Delta S_0$, therefore decreases faster with depth than for $\zeta = 0.8$.

As will become clear from the considerations below, if there is no entrainment from the upflows into the downdrafts and all of the downflows were confined to the small surface area of the vortex structures, the resulting downward-directed *kinetic* energy flux would increase with depth and eventually exceed the enthalpy flux. This would be unphysical, because convection should lead to outward energy transport. However, if there is entrainment with associated mixing, the actual filling factor (fractional area) of downflows, which we denote by f , would be larger and the mean entropy difference $\Delta \bar{S}$ would be diluted correspondingly such that

$$f \Delta \bar{S} = f_s \Delta S_0 \quad (\propto \bar{\rho}^{-\zeta} \Delta S_0). \quad (23)$$

We recall that non-ideal effects can be captured by choosing $\zeta > 0.8$ in the prescription (23). The physics of entrainment has been modeled in the stellar context by Rieutord & Zahn (1995) and the extent of entrainment has been quantified in realistic surface simulations by Trampedach & Stein (2011), who determined the entrainment length scale as the typical scale height of the mass flux in the up- and downflows separately. They call this scale the mass mixing length and find its value to be comparable to H_P .

To estimate the depth dependence of various horizontal averages, and to compute enthalpy and kinetic energy fluxes, we now compute various averages in the two-stream approximation, in which all horizontal averages are the sum of the fraction $1 - f$ of the value in upflows and the fraction f of the value in downflows. Thus, for the mean specific entropy stratification we have

$$\bar{S} = (1 - f)\bar{S}_\uparrow + f\bar{S}_\downarrow = \bar{S}_\uparrow - f \Delta \bar{S}, \quad (24)$$

where \bar{S}_\uparrow and \bar{S}_\downarrow are the mean specific entropies in up- and downflows, and $\Delta \bar{S} = \bar{S}_\uparrow - \bar{S}_\downarrow$ is their difference. We expect that $\bar{S}_\uparrow \approx S_\uparrow$ will be constant if there is no heating in the upwellings, while $\bar{S}_\downarrow \approx (1 - f_s/f)S_\uparrow + (f_s/f)S_\downarrow$ will be dominated by contributions from S_\uparrow due to entrainment.

To calculate $\overline{s^2}$, we must first compute the fluctuating quantity $s = S - \bar{S}$ and then average the squared values in up- and downflows, which, using Equation (24), gives

$$\overline{s^2} = (1 - f)(\bar{S}_\uparrow - \bar{S})^2 + f(\bar{S}_\downarrow - \bar{S})^2 = \hat{f}(\Delta \bar{S})^2, \quad (25)$$

where $\hat{f} = (1 - f)f$ has been introduced as a shorthand. Analogously to the specific entropy, we find for the velocity

$$\bar{U}_z = (1 - f)\bar{U}_\uparrow + f\bar{U}_\downarrow = \bar{U}_\uparrow - f \Delta \bar{U}, \quad (26)$$

where $\bar{U}_\uparrow > 0$ and $\bar{U}_\downarrow < 0$ are the mean up- and downflow velocities with $\Delta \bar{U} = \bar{U}_\uparrow - \bar{U}_\downarrow$. Since the densities in up- and downflows are nearly the same, especially in deeper layers, we have $\bar{U}_z \approx 0$ from mass conservation. With this, we find

$$u_{\text{rms}}^2 \equiv \overline{u_z^2} = (1 - f)\bar{U}_\uparrow^2 + f\bar{U}_\downarrow^2 = \hat{f}(\Delta \bar{U})^2, \quad (27)$$

and thus $\bar{U}_\uparrow/u_{\text{rms}} = [f/(1 - f)]^{1/2}$. Similarly, we calculate

$$\overline{u_z^3} = (1 - f)\bar{U}_\uparrow^3 + f\bar{U}_\downarrow^3 = -\hat{f}(1 - 2f)(\Delta \bar{U})^3, \quad (28)$$

which is negative for $f < 1/2$, and finally

$$\overline{u_z s} = (1 - f)\bar{U}_\uparrow \bar{S}_\uparrow + f\bar{U}_\downarrow \bar{S}_\downarrow = \hat{f} \Delta \bar{U} \Delta \bar{S}. \quad (29)$$

Table 1
 $\phi_{\text{kin}}, -\bar{U}_\downarrow/u_{\text{rms}},$ and $\bar{U}_\uparrow/u_{\text{rms}}$ for Selected Values of f

f	1/2	1/3	0.14	0.015	0.0006
ϕ_{kin}	0	0.35	1	4	20
$-\bar{U}_\downarrow/u_{\text{rms}}$	1	1.4	2.5	8	40
$\bar{U}_\uparrow/u_{\text{rms}}$	1	0.7	0.4	0.12	0.025

Thus, the kinetic energy flux, $\overline{\rho u^2 u_z}/2$, which reduces to $\bar{\rho} \overline{u_z^3}/2$ in this two-stream approximation, is given by

$$F_{\text{kin}} = -\phi_{\text{kin}} \bar{\rho} u_{\text{rms}}^3, \quad (30)$$

where $\phi_{\text{kin}} = (1/2 - f)/\hat{f}^{1/2}$ is a positive prefactor (corresponding to downward kinetic energy flux) if $f < 1/2$. Stein et al. (2009) find $f \approx 1/3$, nearly independently of depth, which yields $\phi_{\text{kin}} \approx \sqrt{2}/4 \approx 0.35$; see Table 1, where we list ϕ_{kin} and $-\bar{U}_\downarrow/u_{\text{rms}} = [(1 - f)/f]^{1/2}$ for selected values of f . The enthalpy flux is proportional to $\overline{u_z s}$ and, using Equations (9) and (29) together with Equations (25) and (27), we find $F_{\text{enth}} = \bar{\rho} \bar{T} u_{\text{rms}} s_{\text{rms}}$. In Appendix D we show that $\bar{T} s_{\text{rms}} = u_{\text{rms}}^2 k_f H_P / (a_{\text{MLT}} \nabla_{\text{ad}})$, where $a_{\text{MLT}} = 1/8$ is a geometric factor in standard MLT.⁶ This leads to

$$F_{\text{enth}} = \phi_{\text{enth}} \bar{\rho} u_{\text{rms}}^3 \quad (31)$$

with $\phi_{\text{enth}} = k_f H_P / (a_{\text{MLT}} \nabla_{\text{ad}})$. This yields $\phi_{\text{enth}} \approx 20$, which is rather large. By contrast, Brandenburg et al. (2005) determined a quantity k_u such that $\phi_{\text{enth}} = k_u^{-3/2} \approx 4$.

We see that the presence of a kinetic energy flux just modifies the usual expression for the convective flux, which then becomes the sum of enthalpy and kinetic energy fluxes; see Section 3.1. This is compatible with recent simulations of R. F. Stein (2016, private communication), in which the fractional kinetic energy flux increases toward the deeper parts ($\gtrsim 40$ Mm) of the domain.

When f becomes small (< 0.14), ϕ_{kin} exceeds unity and for $f < 0.015$, ϕ_{kin} exceeds the estimate $\phi_{\text{enth}} \approx 4$ found by Brandenburg et al. (2005), so the sum of enthalpy and kinetic energy fluxes may become negative, which appears unphysical. In that case, the idea of reconciling the results of Hanasoge et al. (2012) with such convection transporting the solar luminosity could be problematic, unless both up- and downflows were to occur on sufficiently small scales. In the case $f = 0.14$, the resulting \bar{U}_\downarrow would still not be particularly fast and only 2.5 times larger than u_{rms} ; for $f = 0.015$ we have $-\bar{U}_\downarrow \approx 8 u_{\text{rms}}$; see Table 1. Conversely, if the cold entropy blobs were much smaller and were still to contribute significantly to the total energy flux, i.e., if they were much faster, this might not be compatible with an upward directed total energy transport. A possible alternative might be suggested by the compressible simulations of Cattaneo et al. (1991), where the downward-directed kinetic energy flux in the downdrafts was found to balance the upward directed enthalpy flux in the downdrafts. This would imply that convection would transport energy only in the upwellings. This result has however not been confirmed in realistic surface simulations (Stein et al. 1992).

⁶ Not to be confused with the parameter α_{mix} of Section 4.2 below.

Equation (23) shows how f_s decreases with depth, but its actual value and that of ∇_D remains undetermined. In the following, we propose a quantitative prescription for ∇_D by estimating its value within the top few hundred kilometers. In those top layers, we expect ∇_D to reach its maximum value, ∇_D^{\max} , and it should be a certain fraction of $\nabla - \nabla_{\text{ad}}$. In Appendix E we show that $\nabla_D^{\max}/(\nabla - \nabla_{\text{ad}})_{\max} \approx 3$.

In deeper layers, where the local value of $\nabla - \nabla_{\text{ad}}$ has become small, ∇_D should scale with $\bar{s}^2 = \hat{f}(\Delta\bar{s})^2$, and therefore, using Equation (23), it should be proportional to $f_s^2(z) \propto \bar{\rho}^{-2\zeta}$, in addition to an Ma^{-2} factor; see Equation (21). Thus, we have

$$\nabla_D = f_s^2 \nabla_D^{\max}, \quad (32)$$

where we have used for f_s the expression

$$f_s = f_{s0} (\bar{\rho}/\rho_*)^{-\tilde{\zeta}} \quad (\text{for } \bar{\rho} > \rho_*). \quad (33)$$

Here, f_{s0} is a prefactor determining the strength of the resulting Deardorff flux and ρ_* is the density at the point at which $\nabla - \nabla_{\text{ad}}$ is equal to its maximum value (which is just below the photosphere). The new exponent $\tilde{\zeta} = \zeta - \Delta\zeta$ takes the scaling of Mach number with density into account, i.e.,

$$\text{Ma} \propto \bar{\rho}^{-\Delta\zeta}, \quad (34)$$

where $\Delta\zeta (>0)$ will be computed in Section 4.3.

3.4. Kinetic Energy Flux

The importance of the kinetic energy flux has been stressed for some time as a property of compressible stratified convection (Hurlburt et al. 1984; Cattaneo et al. 1991). This flux is related to the asymmetry of up- and downflows (Stein & Nordlund 1989) and is non-vanishing when $f \neq 1/2$; see Section 3.3. It is neglected in standard MLT, as has been discussed by Arnett et al. (2015). Just like the Deardorff flux, the kinetic energy flux is also a non-gradient flux, but it is always directed downward for $f < 1/2$ and must therefore be overcome by the enthalpy flux so that energy can still be transported outwards. However, since the lowest order correlations in F_{kin} are triple correlations, there is no τ approximation treatment analogous to that of the Deardorff flux. However, by comparing $F_{\text{kin}} = -\phi_{\text{kin}} \bar{\rho} u_{\text{rms}}^3$ with the enthalpy flux in Equation (20), it is possible to define a corresponding nabla term via $F_{\text{kin}} = -\frac{1}{3} \bar{\rho} c_p \bar{T} (\tau_{\text{red}} u_{\text{rms}}^2 / H_P) \nabla_{\text{kin}}$. This yields

$$\nabla_{\text{kin}} = 3 \phi_{\text{kin}} u_{\text{rms}} H_P / \tau_{\text{red}} c_p \bar{T}, \quad (35)$$

which is obtained analogously to ∇_D in Equation (21). The prefactor is here defined with a positive sign, so the total non-radiative flux caused by the turbulence is proportional to $\nabla - \nabla_{\text{ad}} + \nabla_D - \nabla_{\text{kin}}$.

Obviously, if ∇_{kin} is strictly proportional to $\nabla - \nabla_{\text{ad}}$, the addition of the ∇_{kin} term does not modify standard MLT, provided that ∇_{kin} depends just on the *local* value of u_{rms} , which in turn is, again, assumed to depend just on the local value of $\nabla - \nabla_{\text{ad}}$. This is different for ∇_D , which depends on the entropy deficit produced by cooling near the surface and in this way on the value of $\nabla - \nabla_{\text{ad}}$ at the position where $\bar{\rho} = \rho_*$. The ∇_D term is therefore truly nonlocal. In the following, we

combine kinetic and enthalpy fluxes into a total contribution $F_{\text{conv}} = F_{\text{enth}} + F_{\text{kin}}$ which arises from convection.

4. MODIFIED MIXING LENGTH MODEL

4.1. Stratification and Flux Balance

To construct an equilibrium model, we begin by considering first the case without convection, so the flux F is carried by radiation alone. Hydrostatic and thermal equilibrium then imply $d\bar{P}/dz = -\bar{\rho}g$ and $Kd\bar{T}/dz = -F$, or, alternatively for the logarithmic gradients,

$$d \ln \bar{P} / dz = -\bar{\rho}g / \bar{P}, \quad (36)$$

$$d \ln \bar{T} / dz = -F / (K\bar{T}). \quad (37)$$

The double-logarithmic temperature gradient is obtained by dividing the two equations through each other, i.e.,

$$\nabla = \frac{d \ln \bar{T}}{d \ln \bar{P}} = \frac{F\bar{P}}{K\bar{T}\bar{\rho}g} = \frac{F c_p \nabla_{\text{ad}}}{Kg}, \quad (38)$$

where we have used the perfect gas equation of state in the form $\bar{P}/\bar{T} \bar{\rho} = c_p - c_v = c_p(1 - 1/\gamma) = c_p \nabla_{\text{ad}}$. If the energy is no longer carried by radiation alone, ∇ cannot be computed from Equation (38), but we have to invoke a suitable theory of convection. In standard MLT, one obtains ∇ as a solution of a cubic equation (Vitense 1953; Kippenhahn & Weigert 1990). In the following, we consider a modification that accounts for the possibility of a Deardorff flux.

Flux balance implies that the sum of the radiative, enthalpy, and kinetic energy fluxes equals the total flux, i.e.,

$$F_{\text{tot}} = F_{\text{rad}} + F_{\text{conv}}. \quad (39)$$

where $F_{\text{conv}} = F_{\text{enth}} + F_{\text{kin}}$ is the flux that arises from convection. In MLT, it is customary to express these fluxes in terms of nablas. For a given double-logarithmic temperature gradient ∇ , the radiative flux is evidently

$$F_{\text{rad}} = \frac{Kg}{c_p \nabla_{\text{ad}}} \nabla, \quad (40)$$

where ∇ characterizes the *actual* temperature gradient. We can also define a *hypothetical* radiative temperature gradient ∇_{rad} that would result if all the energy were carried by radiation, so we can write

$$F_{\text{tot}} = \frac{Kg}{c_p \nabla_{\text{ad}}} \nabla_{\text{rad}}, \quad (41)$$

which follows from Equation (38). We note in passing that the ratio between F_{tot} and the radiative flux carried by the adiabatic temperature gradient, Kg/c_p , is known in laboratory and theoretical studies of convection as the Nusselt number (Hurlburt et al. 1984), whose local value is thus equal to

$$\text{Nu} = \nabla_{\text{rad}} / \nabla_{\text{ad}}. \quad (42)$$

Finally, as explained in Section 3, we have from Equation (20)

$$F_{\text{enth}} = F_0 (\nabla - \nabla_{\text{ad}} + \nabla_D), \quad (43)$$

with $F_0 = \frac{1}{3} \bar{\rho} c_p \bar{T} \tau_{\text{red}} u_{\text{rms}}^2 / H_P$, and F_{kin} being essentially proportional to F_{enth} ; see Equations (30) and (31). Flux equilibrium

then implies

$$\nabla_{\text{rad}} = \nabla + \epsilon(\nabla - \nabla_{\text{ad}} + \nabla_{\text{D}}), \quad (44)$$

where $\epsilon = F_0 c_P \nabla_{\text{ad}} (1 - \phi_{\text{kin}}/\phi_{\text{enth}})/Kg$. However, the expression for F_0 involves the still unknown values of u_{rms} and τ_{red} , which will be discussed in Section 4.2.

4.2. Relaxation Time and Mixing Length

In convection the relevant timescale is the turnover time, which we write as $\tau = 1/u_{\text{rms}} k_f$. We argue that k_f should be estimated via the separation between the entropy rain structures and not their thickness. This becomes important in the picture in which the downdraft threads of the entropy rain merge with neighboring ones to form a tree-like structure, as seen in the surface simulations (Stein & Nordlund 1989; Spruit 1997), leading therefore to different scalings of the two length scales (separation and thickness of structures) with depth. It is then not obvious which of these scales are more relevant for determining the k_f that is relevant for mixing. In view of the uncertainty regarding the choice of k_f , as well as for comparison with standard MLT, we consider models with a fixed value of k_f as well as the more conventional case in which $k_f H_P \approx \text{const}$. To capture the various cases in one expression, we assume in the following

$$k_f = \alpha_{\text{mix}}^\beta H_P^{-1} (k_{f0} H_P)^{1-\beta}, \quad (45)$$

where $\beta = 0$ corresponds to $k_f = k_{f0}$ with a fixed value k_{f0} of the wavenumber of the entropy rain and $\beta = 1$ corresponds to $k_f H_P = \alpha_{\text{mix}}$, where we have allowed for the possibility of a free mixing length parameter, α_{mix} , as is commonly done in standard MLT. It is not to be confused with the parameter a_{MLT} that was introduced in Section 3.3 and Appendix D. Negative values of β would correspond to shrinking scales, but will not be considered here. Likewise, non-integer values of β are conceivable, but will also not be considered here.

Returning now to the discussion of the relaxation time τ in the beginning of this subsection, instead of associating it with the turnover time $(u_{\text{rms}} k_f)^{-1}$, we will allow for the possibility of an additional dilution factor $\phi(z)$ and write $\tau = (u_{\text{rms}} k_f \phi)^{-1}$. Here, $\phi(z)$ increases with depth in a similar fashion as the scale height, so we assume, in analogy with our treatment of k_f in Equation (45), the expression $\phi = (k_{f0} H_P / \alpha_{\text{mix}})^{\beta - \tilde{\beta}}$. This dilution factor only enters in expressions involving the turbulent diffusivity, such as in Equation (17), and hence terms involving τ or τ_{red} . For $\beta = 1$, the case $\tilde{\beta} = 1$ corresponds to the usual MLT concept, while $\tilde{\beta} = 0$ corresponds to a value of ϕ that increases with depth.

In Figure 2 we present illustrative flow structures as well as their depth-dependent horizontal power spectra associated with the three combinations of β and $\tilde{\beta}$ considered here. At the top of the domain, the size and separation of flow structures are the same in all three cases. They remain constant with depth in the case $\beta = \tilde{\beta} = 0$ (Case I), while for $\beta = 0$ and $\tilde{\beta} = 1$ (Case II), the separation increases with depth, but the thickness is still constant. Finally, for $\beta = \tilde{\beta} = 1$ (Case III), both thickness and separation increase with depth. The filling factor decreases with depth in the case $\beta = 0$ and $\tilde{\beta} = 1$, which could potentially make the kinetic energy flux divergent with depth, while for both $\beta = \tilde{\beta} = 0$ and $\beta = \tilde{\beta} = 1$, the filling factor is independent of height.

With these preparations in place, we can write F_0 in Equation (43) in the form

$$F_0 = (\sigma/3\gamma \nabla_{\text{ad}}) \bar{\rho} u_{\text{rms}} c_s^2 \alpha_{\text{mix}}^{-\tilde{\beta}} (k_{f0} H_P)^{-(1-\tilde{\beta})}, \quad (46)$$

where

$$\sigma \equiv \tau_{\text{red}} u_{\text{rms}} k_f \phi = u_{\text{rms}} / (u_{\text{rms}} + \iota c_\gamma / \phi) \quad (47)$$

quantifies the radiative heat exchange between convective elements and the surroundings. The ι term defined in Equation (13) has a maximum at $\ell k_f = \sqrt{3}$, which typically occurs near the surface (e.g., Barekat & Brandenburg 2014).

In Equation (46), the local value of the turbulent rms velocity always depends on the actual flux transported, and therefore it must also depend on $\nabla - \nabla_{\text{ad}} + \nabla_{\text{D}}$. On dimensional grounds, since F_{enth} is proportional to u_{rms}^3 , we have

$$u_{\text{rms}} = c_0 (\nabla - \nabla_{\text{ad}} + \nabla_{\text{D}})^{1/2}, \quad (48)$$

so F_{enth} and F_{kin} are proportional to $(\nabla - \nabla_{\text{ad}} + \nabla_{\text{D}})^{3/2}$. Standard mixing length arguments can be used to show that the prefactor in Equation (48) is a fraction of c_s . As shown in Appendix F, we have

$$c_0/c_s = \sqrt{\sigma a_{\text{MLT}}/3\gamma} \alpha_{\text{mix}}^{-\beta'} (k_{f0} H_P)^{-(1-\beta')}, \quad (49)$$

where $\beta' = (\beta + \tilde{\beta})/2$. It is then clear that F_{kin} scales with H_P like $(k_{f0} H_P)^{-3(1-\beta')}$. This is not the case for F_{enth} , however, because of the factor $(k_{f0} H_P)^{1-\tilde{\beta}}$ which, together with the $(k_{f0} H_P)^{1-\beta'}$ factor, gives the scaling proportional to $(k_{f0} H_P)^{2(1-\beta'')}$, where $\beta'' = (\beta' + \tilde{\beta})/2$.

4.3. Equation for the Superadiabatic Gradient

Now that we know F_0 , we can solve the equation for the superadiabatic gradient. This leads to an equation that is similar to the cubic equation for ∇ , which is familiar from standard MLT (Kippenhahn & Weigert 1990),

$$\nabla_{\text{rad}} = \nabla + \epsilon_*(\nabla - \nabla_{\text{ad}} + \nabla_{\text{D}})^\xi, \quad (50)$$

where $\xi = 3/2$ and $\epsilon_* = (1 - \phi_{\text{kin}}/\phi_{\text{enth}})(\epsilon_{\text{enth}} + \epsilon_{\text{kin}})$ has contributions from the enthalpy and kinetic energy fluxes. These expressions are similar to ϵ in Equation (44), except that ϵ_{enth} is evaluated with c_0 in place of u_{rms} , i.e.,

$$\epsilon_{\text{enth}} = (\sigma/3\gamma)^{3/2} a_{\text{MLT}}^{1/2} (c_s^3/\chi g) \alpha_{\text{mix}}^{-2\beta''} (k_{f0} H_P)^{-2(1-\beta'')}, \quad (51)$$

where $\chi = K/\bar{\rho} c_P$ is the radiative diffusivity. This shows that ϵ_* is essentially a Péclet number based on c_s . The contribution from the kinetic energy flux is given by

$$\epsilon_{\text{kin}}/\epsilon_{\text{enth}} = -\phi_{\text{kin}} a_{\text{MLT}} \nabla_{\text{ad}} \alpha_{\text{mix}}^{-\beta} (k_{f0} H_P)^{-(1-\beta)}. \quad (52)$$

We note in passing that ϵ_* is also related to the Rayleigh number, which is commonly defined in laboratory and numerical studies of convection; see Appendix G. Furthermore, because of convection and the resulting bulk mixing, \bar{S} is now approximately constant, and therefore, unlike in the non-convecting reference solution with $K = \text{const}$ (Section 2), K can no longer be constant, but it reaches a minimum at the point where κ is maximum, which turns out to be at a depth of about 1 Mm in the convection models presented below. Since ϵ_* is inversely proportional to K , it reaches a maximum at that depth and falls off both toward the top and the bottom of the convection zone.

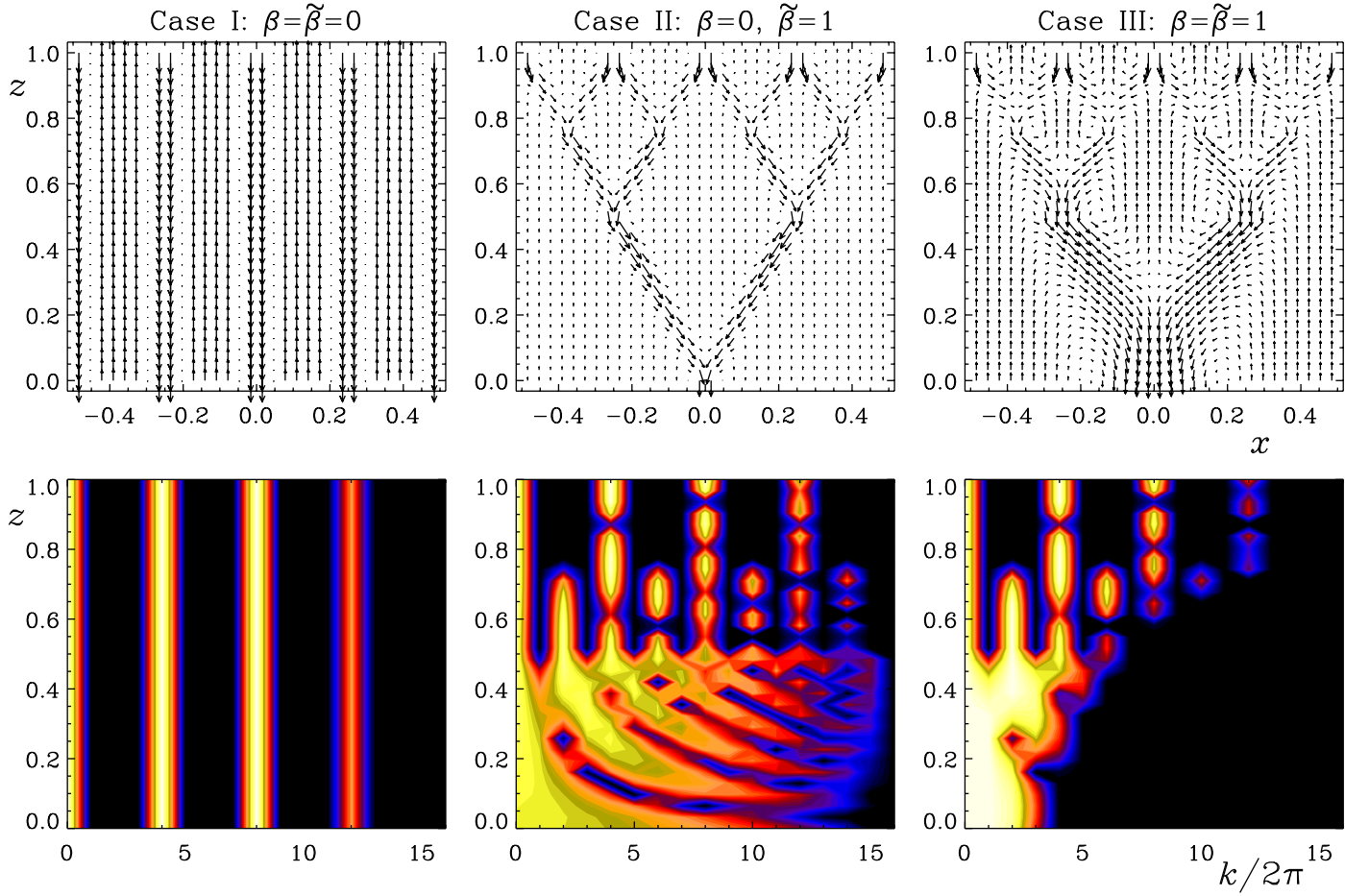


Figure 2. Illustrative flow structures (upper row) and corresponding horizontal power spectra (lower row) associated with the three combinations of β and $\tilde{\beta}$ considered in this paper. Light shades correspond to large logarithmic power, which is seen to extend over large values of k for $\beta = 0$ (Cases I and II) and is confined to progressively smaller k at larger depths when $\beta = 1$ (Case III).

An essential difference between Equation (50) and the usual one in MLT is the presence of ∇_D arising from the Deardorff flux. Within the usual MLT, where $\nabla_D = 0$, one finds that ∇ is slightly above ∇_{ad} , but now it might instead be slightly above $\nabla_{ad} - \nabla_D$. There are indeed two possibilities for convecting solutions ($F_{conv} > 0$), one corresponding to a Schwarzschild-stable solution,

$$\nabla_{ad} - \nabla_D < \nabla < \nabla_{ad} \quad (\text{stable}), \quad (53)$$

and one that is Schwarzschild unstable,

$$\nabla_{ad} - \nabla_D < \nabla_{ad} < \nabla \quad (\text{unstable}). \quad (54)$$

Which of the two possibilities is attained depends on the value of ∇_D and also on details of the solution. As will be discussed in Section 6 below, entropy rain convection may actually still be Schwarzschild unstable without exciting giant cell convection if small-scale turbulent viscosity and diffusivity are strong enough so that the local turbulent Rayleigh number for the deeper layers is subcritical.

In this connection, we note that in standard MLT, one includes the effects of radiative cooling of the convective elements in a different manner than here. Instead of $\nabla - \nabla_{ad}$, the effective buoyancy force is written as $\nabla - \nabla'$, where ∇' always lies between ∇ and ∇_{ad} (Vitense 1953). Thus, one has $\nabla_{ad} < \nabla' < \nabla$, which resembles Equation (54) with a negative value of ∇_D .

To understand the nature of the solutions of Equation (50), it is instructive to treat ξ as an adjustable parameter. For given values of ∇_{rad} and $\tilde{\nabla}_{ad} \equiv \nabla_{ad} - \nabla_D$, the case $\xi = 1$ yields

$$\nabla(\epsilon_*) = \frac{\nabla_{rad} + \epsilon_* \tilde{\nabla}_{ad}}{1 + \epsilon_*}. \quad (55)$$

It shows that $\nabla \rightarrow \nabla_{rad}$ for $\epsilon_* \rightarrow 0$ (stable surface layers) and $\nabla \rightarrow \tilde{\nabla}_{ad}$ for $\epsilon_* \gg 1$ (deeper layers). Next, to discuss the general case $\xi \neq 1$, we define $\Delta\nabla = \nabla - \tilde{\nabla}_{ad}$ and $\Delta\nabla_{rad} = \nabla_{rad} - \tilde{\nabla}_{ad}$. For $\Delta\nabla_{rad} < 0$ we have $\Delta\nabla = \Delta\nabla_{rad}$, while for $\Delta\nabla_{rad} > 0$ and $\Delta\nabla \ll 1$, a useful approximation is

$$\Delta\nabla \approx \Delta\nabla_{rad}^{1/\xi} / (q \Delta\nabla_{rad}^{1/\xi-1} + \epsilon_*^{1/\xi}), \quad (56)$$

where $q = \xi^{-1}$. It agrees with Equation (55) in the special case $\xi = 1$, where $\Delta\nabla = \Delta\nabla_{rad} / (1 + \epsilon_*)$. In Figure 3 we plot ∇ versus ϵ_* for $\nabla_{rad} = 10^5$. The approximation yields $\nabla > \nabla_{rad}$ for $\epsilon_* < 10^{-3}$, which is unphysical. This can be mitigated by choosing $q = 1$; see Figure 3.

For the relevant case of large values of ϵ_* , we have

$$\Delta\nabla \approx (\nabla_{rad} / \epsilon_*)^{1/\xi}. \quad (57)$$

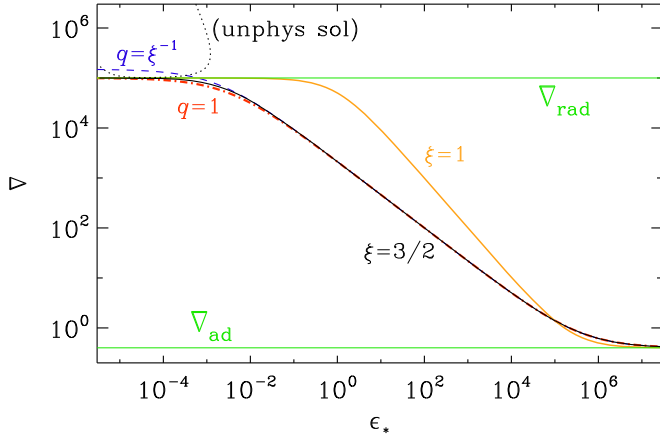


Figure 3. Solution of Equation (50) for $\xi = 3/2$ and $\nabla_{\text{rad}} = 10^5$ (solid black), compared with the approximation (56) for $q = \xi^{-1}$ (dashed blue) and $q = 1$ (dashed-dotted red), and $\xi = 1$ (thin orange). The dotted line gives an additional unphysical solution of Equation (50) for $\xi = 3/2$. The limiting cases $\nabla = \nabla_{\text{ad}}$ and ∇_{rad} are shown as thin horizontal green lines.

This relation is useful because, even though both ∇_{rad} and ϵ_* depend on K , their ratio does not and is given by

$$\frac{\nabla_{\text{rad}}}{\epsilon_*} = \frac{3\nabla_{\text{ad}}}{\sigma} \frac{F_{\text{tot}}/\bar{p}}{c_0 g/k_f} \propto \frac{(k_{f0} H_P)^{2-2\beta''}}{\bar{\rho}_s^3} \propto \bar{T}^{-(1+2\beta'')}, \quad (58)$$

where $\beta'' = (\beta + 3\tilde{\beta})/4$. Thus, $\Delta\nabla \propto \bar{T}^{-(1+2\beta'')/\xi}$ gives the scaling of $\Delta\nabla$ for the bulk of the convection zone. Therefore, looking at Equation (48), we find $\text{Ma} \propto \bar{T}^{-m}$ with

$$m = 1 - \beta' + (1/2 + \beta'')/\xi. \quad (59)$$

For $\xi = 3/2$, using the relation $3\beta' - 2\beta'' = \beta$, we find that $m = (4 - \beta)/3$ is independent of the value of $\tilde{\beta}$. Thus, we have $m = 4/3$ with $\beta = 0$ and $m = 1$ with $\beta = 1$. For isentropic stratification, this implies for the Mach number, given by Equation (34), the following scaling: $\Delta\zeta = 8/9$ with $\beta = 0$ and $\Delta\zeta = 2/3$ with $\beta = 1$.

4.4. Relative Importance of the Deardorff Term

In Section 3.3 we have considered the depth dependence of the Deardorff term via Equations (32) and (33). However, for ∇_D to be important at increasing depths, it must exceed the subadiabatic gradient $\nabla_{\text{ad}} - \nabla$, because otherwise it would not be possible for the Deardorff term to make $\nabla - \nabla_{\text{ad}} + \nabla_D$ positive. From Equations (57) and (58) we see that $\Delta\nabla$ depends on \bar{T} in a power-law fashion. We are particularly interested in the conditions under which $\Delta\nabla$ falls off faster than ∇_D , because that would ensure that the ∇_D term remains important even at larger depths.

For $\xi = 3/2$, and since the convection zone is nearly isentropically stratified ($\bar{T} \propto \bar{p}^{2/3}$), we have

$$\Delta\nabla \propto \bar{p}^{-\zeta_{\Delta\nabla}} \quad \text{with} \quad \zeta_{\Delta\nabla} = \frac{4}{9}(1 + 2\beta''). \quad (60)$$

In Table 2 we compare for various combinations of β and $\tilde{\beta}$ the exponents for $\Delta\nabla$ and ∇_D , i.e., $2\tilde{\zeta} = 2\zeta$ and $2\Delta\zeta$, where $\Delta\zeta = 2m/3$. In all cases we have chosen $\zeta = 1$, i.e., we allow for moderately non-ideal (radiative) effects relative to the ideal case with $\zeta = 0.8$. We see that the difference is positive and non-vanishing in all cases, except for $\beta = 1$ and $\tilde{\beta} = 0$. In the

Table 2
Comparison of $\zeta_{\Delta\nabla}$ and $2\tilde{\zeta} = 2\zeta - 2\Delta\zeta$, and Their Difference for Various Combinations of β and $\tilde{\beta}$ Using $\zeta = 1$

	β	$\tilde{\beta}$	β'	β''	$\zeta_{\Delta\nabla}$	$2\tilde{\zeta}$	$\zeta_{\Delta\nabla} - 2\tilde{\zeta}$
Case I	0	0	0	0	4/9	2/9	2/9
Case II	0	1	1/2	3/4	10/9	2/9	8/9
	1	0	1/2	1/4	2/3	2/3	0
Case III	1	1	1	1	4/3	2/3	2/3

Table 3
Illustration of Scaling Relationships with $k_{f0} H_P$

$F_{\text{enth}} \propto u_{\text{rms}}^3 (k_{f0} H_P)^{1-\beta}$	buoyancy force
$F_{\text{enth}} \propto u_{\text{rms}} \Delta\nabla / (k_{f0} H_P)^{1-\tilde{\beta}}$	mean-field expression
$u_{\text{rms}} \propto (\Delta\nabla)^{1/2} / (k_{f0} H_P)^{1-\beta'}$	$\beta' = (\beta + \tilde{\beta})/2$
$F_{\text{enth}} \propto (\Delta\nabla)^{3/2} / (k_{f0} H_P)^{2(1-\beta'')}$	$\beta'' = (\beta' + \tilde{\beta})/2 = (\beta + 3\tilde{\beta})/4$
$F_{\text{kin}} \propto (\Delta\nabla)^{3/2} / (k_{f0} H_P)^{3(1-\beta')}$	kinetic energy flux

following, we present solutions for all of the remaining three cases. Before doing this, let us recapitulate what led to the threefold dominance of $\tilde{\beta}$ over β . For better illustration, we summarize in Table 3 the various relationships that led to the scaling of $\Delta\nabla$ with $k_{f0} H_P$.

In the first expression for F_{enth} , β enters because it characterizes the relation between the buoyancy force proportional to $\delta T/\bar{T}$ and advection proportional to $k_f u_{\text{rms}}^2$; see Appendix D. For thin threads, we expect the relevant k_f to be large, i.e., $\beta = 0$ (Cases I and II in Figure 2). Next, in the second expression for F_{enth} , we have used a mean-field expression to relate F_{enth} to $\Delta\nabla$ via a turbulent diffusivity proportional to $\tau u_{\text{rms}}^2 \approx u_{\text{rms}}/k_f \phi$, where, as argued above, the dilution factor ϕ has entered. It is this expression that is closest to conventional MLT, because here we expect $\tilde{\beta} = 1$.

The remaining two relationships in Table 3 explain why the $\tilde{\beta}$ term appears three times more dominantly than the β term. By equating the first two expressions for F_{enth} in Table 3, we find first of all the relation between u_{rms} and $\Delta\nabla$, where β and $\tilde{\beta}$ contribute equal shares through $\beta' = (\beta + \tilde{\beta})/2$. However, to see the scaling of $\Delta\nabla$, we need to go back to the second expression for F_{enth} , because it changes only weakly with depth. Now, β' and $\tilde{\beta}$ contribute equal shares, and this means that $\tilde{\beta}$ has now become three times more dominant than β through the expression $\beta'' = (\beta + 3\tilde{\beta})/4$. Thus, for $\beta = 0$ and $\tilde{\beta} = 1$, $\Delta\nabla$ shows nearly the standard scaling with H_P . Furthermore, looking again at the first expression for F_{enth} in Table 3, we see that only β enters, so the scaling of u_{rms} is fully characterized by that of small blobs with negative buoyancy.

5. NUMERICAL SOLUTIONS

In this section we present numerical solutions to demonstrate the effect of the ∇_D term on the resulting stratification. At the end of this section, we also compare with the non-convecting reference solution mentioned in the introduction. We should emphasize that, although we use solar parameters, our models cannot represent the Sun, because ionization effects have been ignored (we take $\mu = 0.6$ for the mean molecular weight in $c_p - c_v = \mathcal{R}/\mu$, where \mathcal{R} is the universal gas constant). A rather simple opacity law of the form of Equation (6), with $\kappa_0 = 10^4 \text{ cm}^2 \text{ g}^{-1}$, $\rho_0 = 10^{-5} \text{ g cm}^{-3}$, $T_0 = 13,000 \text{ K}$, and $a = 0.5, b = 18$, for the exponents in the power-law expression

for κ_H^- has been used. We also neglect the departure from plane-parallel geometry, so our model can only give qualitative indications.

The system of two differential Equations (36) and (37) decouples by using $\ln \bar{P}$ as the independent variable. We thus integrate

$$d \ln \bar{T} / d \ln \bar{P} = \nabla, \quad (61)$$

using Equations (32), (33), and (50) to compute ∇ . As an initial condition we use $\bar{T}_{\text{top}} = T_{\text{eff}}/2^{1/4}$ at a sufficiently low pressure (here $\bar{P}_{\text{top}} = 10^5 \text{ dyn cm}^{-2}$) so as to capture the initially isothermal part of the atmosphere; see, e.g., Böhm-Vitense (1958) or Mihalas (1978). Here, $T_{\text{eff}} = (F_{\text{tot}}/\sigma_{\text{SB}})^{1/4}$ is the effective temperature. We approximate the u_{rms} term in Equation (47) by using the value from the previous step. The Deardorff term is characterized by the assumed value of f_{s0} and the value of $\tilde{\zeta} = \zeta - \Delta\zeta$ (Section 3.3), where $\zeta = 1$ and $\Delta\zeta$ is a function of β and $\tilde{\beta}$ (Section 4.3). Since we integrate from the top downward, no prior knowledge of ρ_* is needed, because the Deardorff term is invoked only after $\nabla - \nabla_{\text{ad}}$ has reached its peak value.

Geometrical and optical depths are obtained respectively as

$$-z = \int (\bar{P}/\bar{\rho}g) d \ln \bar{P} \quad \text{and} \quad \tau = \int (\kappa \bar{P}/g) d \ln \bar{P}, \quad (62)$$

where the integration of τ starts at $\ln \bar{P}_{\text{top}}$ and that of $-z$ at the position where $\tau = 1$, which is referred to as the surface. The factor $\kappa \bar{P}/g$, which is the same as H_P , is retained because of the similarity with that in the expression for τ . The z coordinate is used in some of our plots.

In the following, we present solutions for the three combinations of β and $\tilde{\beta}$ sketched in Figure 2. We recall that only Cases I and II ($\beta = 0$ with $\tilde{\beta} = 0$ or 1) correspond to small length scales in the deeper layers (see also the power spectra in Figure 2) and are therefore of interest when trying to reconcile the non-detection of convective motions by Hanasoge et al. (2012, 2016) at the theoretically expected levels. We did already emphasize that Case II with $\beta = 0$ and $\tilde{\beta} = 1$ is likely to lead to large kinetic energy fluxes. However, based on the results presented below, it turns out that in the case $\beta = \tilde{\beta} = 0$ (Case I), which was favored by these two requirements (small length scale and non-divergent kinetic energy flux), the Deardorff term is unlikely to have a significant effect, because for $\tilde{\beta} = 0$, the gradient term in the enthalpy flux becomes rather inefficient and must therefore be compensated for by a correspondingly larger superadiabatic gradient, and thus, only a rather large Deardorff flux ($f_{s0} \geq 0.5$) can make the resulting stratification sufficiently subadiabatic. This is the case shown in Figure 4, where we present profiles of \bar{S}/c_P , $\nabla - \nabla_{\text{ad}}$, and u_{rms} for $f_{s0} = 0$ (no Deardorff term), as well as $f_{s0} = 0.2, 0.3$, and 0.5 . The u_{rms} profiles are basically the same for all values of f_{s0} and fall off like $\bar{P}^{-1/3}$. This is expected, because $u_{\text{rms}} \propto \bar{T}^{1/2-m}$ and $m = 4/3$; see Equation (59), where we have used $c_s \propto \bar{T}^{1/2}$. Thus, for isentropic stratification we have $u_{\text{rms}} \propto \bar{P}^{(1-2m)/5} = \bar{P}^{-1/3}$.

Next, we show in Figures 5 and 6 Cases II and III with $\tilde{\beta} = 1$ and $f_{s0} = 0, 0.1$, and 0.2 . For both $\beta = 0$ and $\beta = 1$, the Deardorff flux now has a stronger effect and is able to make the deeper parts of the domain Schwarzschild stable even for $f_{s0} = 0.1$. Those layers would then be Schwarzschild stable and no longer a source of giant cells. Again, the profiles of u_{rms}

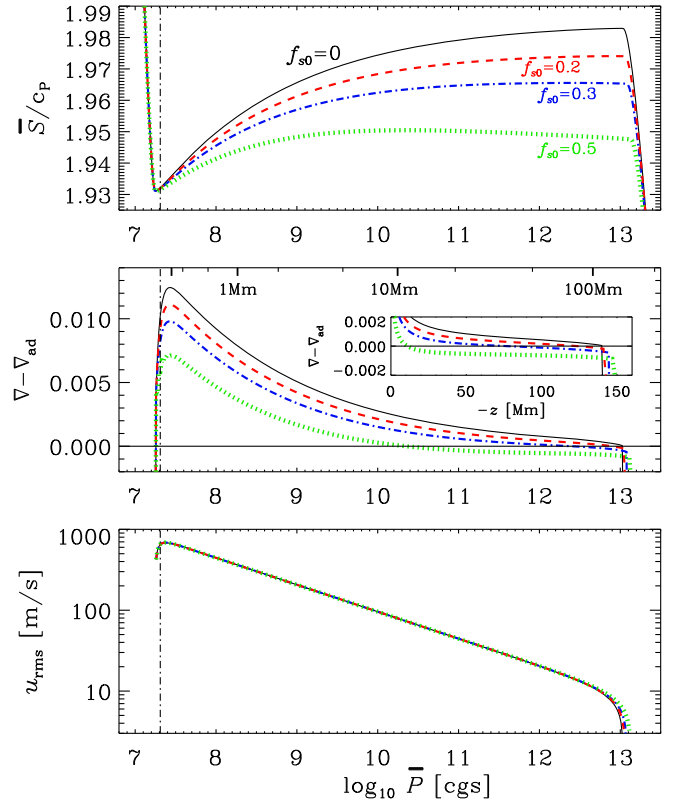


Figure 4. Profiles of \bar{S}/c_P , $\nabla - \nabla_{\text{ad}}$, and u_{rms} for $f_{s0} = 0$ ($\nabla_D = 0$), as well as $f_{s0} = 0.2, 0.3$, and 0.5 for $\beta = \tilde{\beta} = 0$ (Case I) with $\tilde{\zeta} = 1/9$. The location of the surface ($\tau = 1$) is indicated by vertical dashed-dotted lines and geometric depths below the surface are indicated in the middle panel, starting with tick marks at 100, 200, and 500 km, and continuing with 1, 2, and 5 Mm, etc. The inset in the middle panel shows $\nabla - \nabla_{\text{ad}}$ over a narrower range as a function of $-z$.

are similar regardless of the value of f_{s0} , but fall off more slowly for $\beta = 1$ ($u_{\text{rms}} \propto \bar{P}^{-1/5}$), compared to $\beta = 0$ ($u_{\text{rms}} \propto \bar{P}^{-1/3}$). This agrees with our theory, because for $m = 1$ we have $u_{\text{rms}} \propto \bar{P}^{(1-2m)/5} = \bar{P}^{-1/5}$.

Given that the stratification is stable in the deeper parts, we can calculate the Brunt–Väisälä frequency of buoyancy oscillations, N_{BV} , which is given by $N_{\text{BV}}^2 = -(\nabla - \nabla_{\text{ad}})g/H_P$. At intermediate and larger depths of the convection zone, we have $g/H_P \lesssim 10^{-2} \text{ s}^{-2}$ and $\nabla_{\text{ad}} - \nabla \lesssim 10^{-4}$, so the period of buoyancy oscillations would be of the order of days. This is comparable to or less than the turnover time τ . Indeed, one finds that $\tau N_{\text{BV}} \approx \sqrt{\nabla_D/\Delta\nabla - 1}$ exceeds unity in the deeper parts, which is a consequence of ∇_D falling off with \bar{p} with a smaller power than $\Delta\nabla$, as discussed in Section 4.4. Nevertheless, the resulting decrease in \bar{S}/c_P with depth remains always small (10^{-3} or less) compared with the value of $\Delta S_0/c_P$ produced at the surface ($\approx u_{\text{rms}}^2/a_{\text{MLT}} \nabla_{\text{ad}} c_s^2 \approx 0.01$). Thus, based on this, the descending low-entropy blobs would reach the bottom of the convection zone before their negative buoyancy is neutralized by the decreasing average entropy. In other words, they will never perform any actual oscillations before reaching the bottom of the convection zone, i.e., where $F_{\text{conv}} = 0$.

It may be interesting to note that the depth of the convection zone increases slightly with increasing values of f_{s0} . Of course, our model is idealized and represents the Sun at best only approximately. Furthermore, as emphasized in Section 1, the

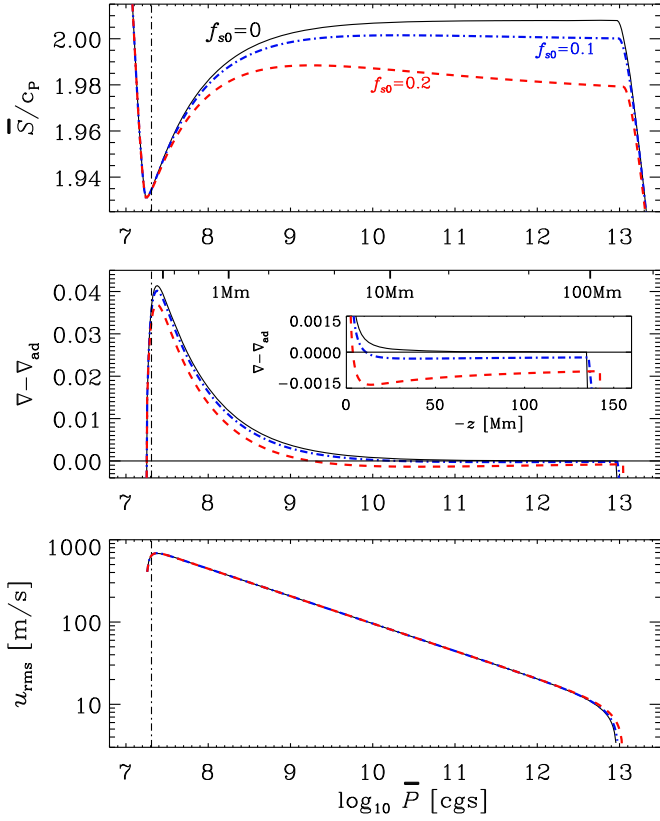


Figure 5. Same as Figure 4, but for $\beta = 0$, $\tilde{\beta} = 1$ (Case II) with $\tilde{\zeta} = 1/9$, and $f_{s0} = 0, 0.1$, and 0.2 .

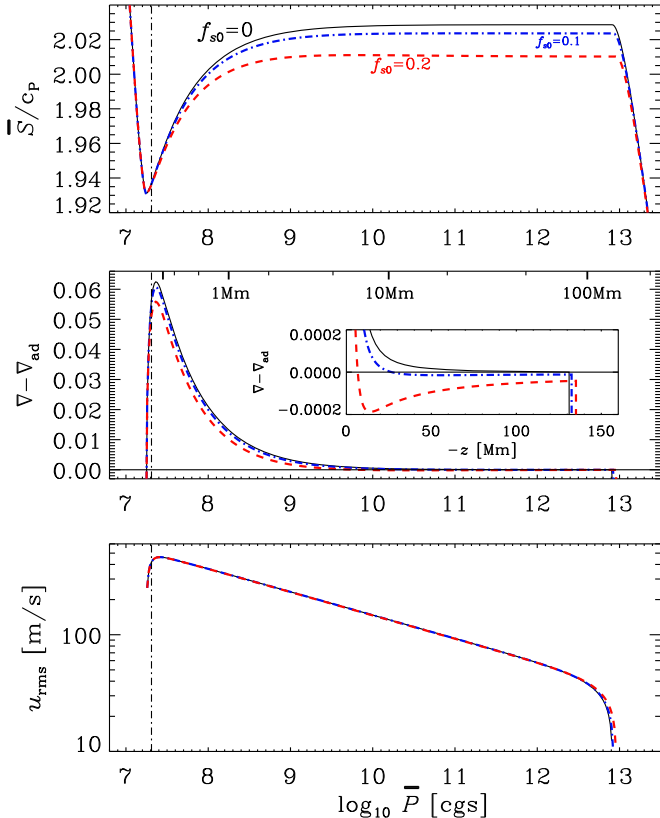


Figure 6. Same as Figure 4, but for $\beta = \tilde{\beta} = 1$ (Case III) $\tilde{\zeta} = 1/3$, and $f_{s0} = 0, 0.1$, and 0.2 .

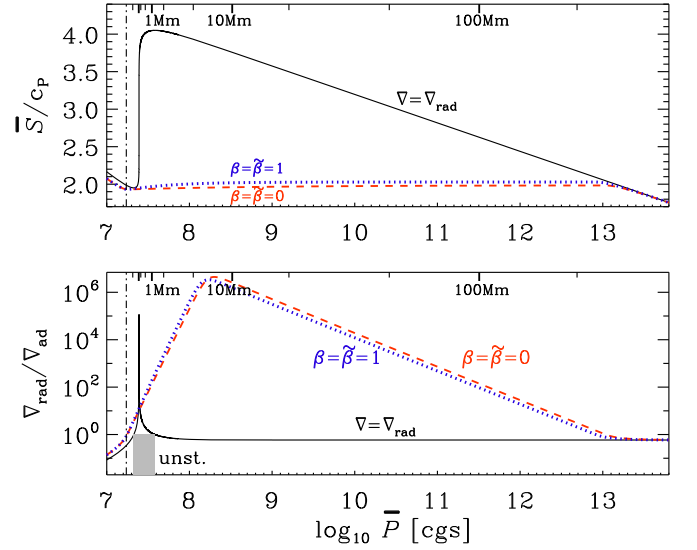


Figure 7. Comparison of \bar{S}/c_P (top) and $\nabla_{\text{rad}}/\nabla_{\text{ad}}$ (bottom) between the non-convective radiative reference solution ($\nabla = \nabla_{\text{rad}}$) and standard convective solutions ($\nabla_D = 0$) with $\beta = \tilde{\beta} = 0$ (red, dashed) and $\beta = \tilde{\beta} = 1$ (blue, dotted). In both panels, the location of the $\tau = 1$ surface is indicated by vertical dashed-dotted lines and geometric depths below the surface are indicated for the non-convective solution, starting with tick marks at 50, 100, and 200 km, etc. The location of the a priori unstable layer ($\nabla_{\text{rad}} > \nabla_{\text{ad}}$) is marked in the lower panel by a small gray strip. For the other solutions, the depths are different; see those in Figure 6 for the solution with $f_{s0} = 0$, which is the same as the one here for $\beta = \tilde{\beta} = 1$.

depth of the convection zone is well determined seismically, and this should be reproduced by a solar model with realistic atomic physics and appropriately chosen adjustable parameters. However, it is known from the work of Serenelli et al. (2009) that a slight expansion of the solar convection zone would actually be required to compensate for the shrinking that follows from the downward revision of solar abundances (Asplund et al. 2004) which is based on three-dimensional convective atmosphere simulations, compared to previous analysis based on one-dimensional semi-empirical models (Grevesse & Sauval 1998).

Finally, we compare in Figure 7 the standard convective solution ($\beta = \tilde{\beta} = 1$; same as the case with $f_{s0} = 0$ in Figure 6) with the non-convective radiative reference solution. Not surprisingly, owing to the absence of convection, the same flux can now only be transported with a greatly enhanced negative (unstable) entropy gradient near the surface. However, this layer is now extremely thin (1.15 Mm) and the peak in $\nabla_{\text{rad}}/\nabla_{\text{ad}}$ is about 100 km below the $\tau = 1$ surface. Note also that its peak value ($\approx 10^5$) is below that in the presence of convection, where it reaches a maximum of $\approx 4 \times 10^6$ at a depth of ≈ 1 Mm. As shown in Equation (42), this value can be interpreted as the local Nusselt number. Note also that the result for $\beta = \tilde{\beta} = 0$ (when ∇_D is weak) is rather similar to that for $\beta = \tilde{\beta} = 1$; see the dashed and dotted lines in Figure 7.

Our calculations have demonstrated that for $\tilde{\beta} = 1$, regardless of the value of β , bulk mixing changes the non-convecting reference state to a nearly isentropic one. However, whether the mean entropy gradient is slightly stably or slightly unstably stratified depends on the presence of the Deardorff flux. In the model with $\tilde{\beta} = 0$, however, bulk mixing is rather inefficient and the stratification would be Schwarzschild unstable unless an unrealistically large Deardorff flux is invoked.

At the end of the introduction, we discussed that the entropy rain itself might create an unstable stratification. Let us now return to this question with more detailed estimates. This will be done in the following section.

6. ALTERNATIVE CONSIDERATIONS

Assuming that the upflows are perfectly isentropic, Spruit (1997) argues that the low-entropy material from the top with its decreasing entropy filling factor f_s in deeper layers necessarily leads to a negative mean entropy gradient. Specifically, using Equation (24), one obtains

$$-H_P \frac{d\bar{S}/c_P}{dz} = \frac{(\Delta S)_0}{c_P} \frac{df_s}{d \ln \bar{P}} = \frac{2}{5} \frac{(\Delta S)_0}{c_P} f_s > 0, \quad (63)$$

where we have used $d \ln f_s / d \ln \bar{p} = 2/3$ and $d \ln \bar{p} / d \ln \bar{P} = 3/5$ for an isentropic layer with $\gamma = 5/3$. Thus, the stratification would be Schwarzschild unstable. This is also borne out by the solar simulations with realistic physics, although the computational domains are sufficiently shallow so that the radiative flux is still small in the deeper parts and usually even neglected altogether. Toward the bottom of the convection zone, however, radiation becomes progressively more important and the mean entropy gradient in the upflows may no longer be vanishing.

To estimate the mean entropy gradient in the upflows, we may balance the steady state entropy advection with the negative radiative flux divergence, i.e.,

$$\bar{p} \bar{T} \bar{U}_\uparrow d\bar{S}_\uparrow/dz \approx -(dF_{\text{rad}}/dz)_\uparrow. \quad (64)$$

The sign of dF_{rad}/dz is negative and thus compatible with a positive $d\bar{S}_\uparrow/dz$ in the upflows, but it would only be large enough near the bottom of the convection zone. Higher up, dF_{rad}/dz becomes smaller and eventually unimportant. On the other hand, $\bar{U}_\uparrow/u_{\text{rms}}$ can be rather small (see Table 1). Furthermore, our considerations neglect the fact that the gas in the upflows expands, so only a fraction of the gas can ascend before it begins to occupy the available surface area. Therefore, the rest of the gas would have to remain stagnant and continue to heat up. In reality, of course, there would be continuous entrainment, resulting in a finite, but still low, effective upward velocity.

We now use selected solutions obtained in Section 5 to estimate the effective fractional upward velocity for radiative heating/cooling to dominate over advection, defined as

$$\bar{U}_{\text{eff}}^\uparrow/u_{\text{rms}} = (-dF_{\text{rad}}/dz)/(u_{\text{rms}} \bar{p} \bar{T} d\bar{S}/dz). \quad (65)$$

Here we have used $\bar{T} d\bar{S}/dz = g \Delta \nabla / \nabla_{\text{ad}}$. Figure 8 shows this quantity for Cases I–III, which are shown in Figures 4–6. It turns out that for Case I with $f_{s0} = 0.5$, the effective fractional upward velocity is around 0.01, while for Case III with $f_{s0} = 0.1$ it can be even larger. The value 0.01 is compatible with the $\bar{U}^\uparrow/u_{\text{rms}}$ values in Table 1 if we assume $f = 0.01$ (so $\bar{U}^\uparrow/u_{\text{rms}} = 0.1$) and $\bar{U}_{\text{eff}}^\uparrow/\bar{U}^\uparrow = 0.1$. For Case II with $f_{s0} = 0.1$, the effective fractional upward velocity is around 10^{-3} , which may be unrealistically small.

Based on the considerations above, we may conclude that the case for a Schwarzschild-stable entropy gradient depends on model assumptions for the implementation of the Deardorff flux that are potentially in conflict with the mean entropy

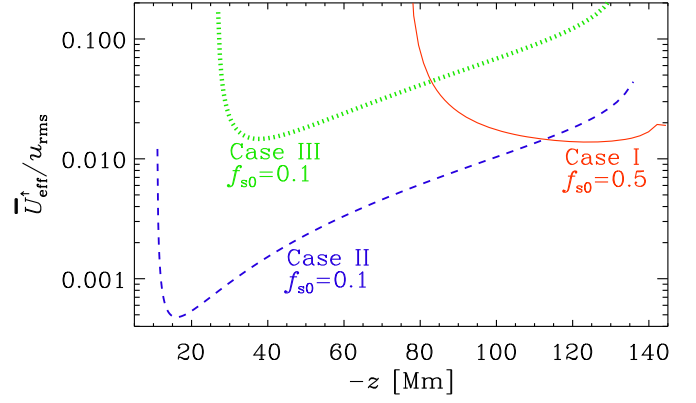


Figure 8. Dependence of $\bar{U}_{\text{eff}}^\uparrow/u_{\text{rms}}$ on depths $-z$ for Case I with $f_{s0} = 0.5$ (solid red line), as well as Cases II and III with $f_{s0} = 0.1$ (dashed blue and dotted green lines, respectively).

gradient expected from the two-stream model. Whether or not there is really a potential conflict depends not only on the model parameters, but also on Equation (64) itself. This is because the \bar{T} factor in the entropy equation is outside the derivative, making it impossible to derive the total flux balance assumed in Equation (39); see also Rogachevskii & Kleeorin (2015) for related details.

In view of these complications, it is worthwhile to discuss alternative ways of avoiding giant cell convection. For this purpose, we have to address the global problem of using some coarse-grained form of effective mean-field equations. When considering the equations of mean-field hydrodynamics, in which the small-scale enthalpy and momentum fluxes are parameterized in terms of negative mean entropy and mean velocity gradients, respectively, one finds solar differential rotation as a result of non-diffusive contributions to the Reynolds stress, but in certain parameter regimes a new instability was found to develop (Rüdiger 1989; Tuominen & Rüdiger 1989; Rüdiger & Spahn 1992). This instability was later identified as one that is analogous to Rayleigh–Bénard convection, but now for an already convecting mean state (Tuominen et al. 1994). This instability would lead to giant cell convection.

The existence of giant cell convection is under debate, but assuming that it does *not* exist in the Sun, one might either hypothesize that the mean-field equations are too simplified or that mean-field convection could be suppressed by sufficiently strong turbulent viscosity and turbulent thermal diffusivity coefficients. These turbulent coefficients define a turbulent Rayleigh number for a layer of thickness d ,

$$\text{Ra}_t = \frac{g d^4}{\nu_t \chi_t} \left(\frac{\Delta \nabla}{H_P} \right)_{z_*}. \quad (66)$$

This number would then have to be still below the critical value for convection. Equation (66) differs from the usual one defined through Equation (86) in Appendix G in that, first, $\nu \rightarrow \nu_t$ and $\chi \rightarrow \chi_t$ have been substituted, and second, the superadiabatic gradient of the non-convecting reference solution is now replaced by the actual one. This might be a plausible alternative to explaining the absence of giant cell convection if the idea of turning the stratification from Schwarzschild unstable to Schwarzschild stable through the ∇_D term were to turn out untenable.

We recapitulate that in this alternate explanation, the stratification is Schwarzschild unstable, i.e., $Ra_t > 0$, corresponding to Equation (54), but $Ra_t < Ra_t^{\text{crit}}$ is still below a certain critical value, so it would be stable by a turbulent version of the Rayleigh–Bénard criterion. Tuominen et al. (1994) found $Ra_t^{\text{crit}} \approx 300$ for a vanishing rotation rate. However, they also estimated that $Ra_t > Ra_t^{\text{crit}}$ for plausible solar parameters, so the direct adoption of this idea would be problematic, too. However, one might speculate that a more accurate treatment could lead to stability when allowing, for example, for spatial nonlocality of the turbulent transport (Brandenburg et al. 2008; Rheinhardt & Brandenburg 2012).

7. CONCLUSIONS

In the present work we have suggested that the enthalpy flux in stellar mixing length models should contain an extra nonlocal contribution so that the enthalpy flux is no longer proportional to the local superadiabatic gradient, $\nabla - \nabla_{\text{ad}}$, but to $\nabla - \nabla_{\text{ad}} + \nabla_{\text{D}}$, where ∇_{D} is a new nonlocal contribution that was first identified by Deardorff (1966) in the meteorological context. The significance of this term lies in the fact that it provides an alternative to the usual local entropy gradient term and can transport enthalpy flux outwards—even in a slightly stably stratified layer.

We have presented a modified formulation of stellar MLT that includes the ∇_{D} term, in addition to a ∇_{kin} term resulting from the kinetic energy flux. The formalism and the final results are similar to those of conventional MLT in that one also arrives here at a cubic equation for ∇ , but the term $\nabla - \nabla_{\text{ad}}$ is now replaced by $\nabla - \nabla_{\text{ad}} + \nabla_{\text{D}} - \nabla_{\text{kin}}$. This new formulation implies that convection can carry a finite flux while $\nabla - \nabla_{\text{ad}}$ is still negative and therefore the stratification is Schwarzschild stable, i.e., Equation (53) is obeyed. Consequently, if confirmed, no large length scales are being excited.

The present formulation allows for different treatments of the length scales governing buoyant elements on the one hand and the time and length scales associated with mixing on the other. When both are independent of depth ($\beta = \tilde{\beta} = 0$), mixing becomes inefficient at larger depths. Thus, to carry a certain fraction of the enthalpy flux, the superadiabatic gradient needs to be larger than otherwise, making it harder for the Deardorff term to revert the sign of $\nabla - \nabla_{\text{ad}}$. On the other hand, for a tree-like hierarchy of many downdrafts merging into fewer thin ones at greater depths ($\beta = 0$, $\tilde{\beta} = 1$), the increasing length scale associated with increasing separation enhances vertical mixing, making the stratification nearly isentropic without the Deardorff term, and slightly subadiabatic with a weak Deardorff term. In that case, however, the filling factor of the downflows decreases with depth as $\rho^{-\zeta}$, which may imply an unrealistically large downward kinetic energy flux. This leaves us with the standard flow topology ($\beta = \tilde{\beta} = 1$), where both size and separation of structures increase with depth. The Deardorff flux can still cause the stratification to have a subadiabatic gradient, so no giant cell convection would be excited locally, but the flow structures would be large and should be helioseismically detectable, as has been found by Greer et al. (2015) using ring-diagram local helioseismology.

It would be useful to explore the thermodynamic aspects of the present model more thoroughly and to connect with related approaches. An example is the work by Rempel (2004), who studied a semianalytic overshoot model that was driven nonlocally by downdraft plumes, similar to what was suggested

by Spruit (1997). Rempel also finds an extended subadiabatic layer in large parts of the model. Furthermore, there are similarities to the nonlocal mixing length model by Xiong & Deng (2001), who, again, find an extended deeper layer that is subadiabatically stratified in the deeper parts of their model. In this connection we emphasize the main difference between nonlocal turbulence owing to the Deardorff flux and usual overshoot: in the latter case the enthalpy flux would go inward as a consequence of the reversed entropy gradient, while in the present model the Deardorff flux goes outward.

Future work could proceed along two separate paths. On the one hand, one must establish the detailed physics leading to the ∇_{D} term using models with reduced opacity, in which reliable DNS are still possible, i.e., no SGS terms are added and the primitive equations are solved as stated, without invoking Equation (1). On the other hand, one could study suitably parameterized large eddy simulations that either include a nonlocal Deardorff term of the form given by Equation (32), as discussed in Section 3, or that explicitly release entropy rain at the surface such that the resulting stratification is still slightly stable. This would be particularly useful in global simulations that would otherwise have no entropy rain.

As stimulating as the results of Hanasoge et al. (2012) are, they do require further scrutiny and call for the resolution of the existing conflicts with other helioseismic studies such as those of Greer et al. (2015). Alternatively, global helioseismic techniques for detecting giant cell convection (Lavelly & Ritzwoller 1993; Chatterjee & Antia 2009; Woodard 2014) can provide another independent way of detecting deep larger-scale flows (Woodard 2016). Realistic simulations should eventually agree with helioseismic results of flows in deeper layers of the Sun, but at the moment it is still unclear whether a subgrid scale treatment as in Equation (1) adequately captures the small-scale flows that can be responsible for the Deardorff flux and whether they would in principle be able to predict the subtle departures from superadiabatic stratification on subthermal timescales.

I thank the referee for criticism that has led to many improvements, and Evan Anders, Jean-Francois Cossette, Ben Greer, Åke Nordlund, Mark Rast, Matthias Rempel, Matthias Rheinhardt, Bob Stein, Peter Sullivan, Regner Trampedach, and Jörn Warnecke for interesting discussions and comments. I also wish to acknowledge Juri Toomre for mentioning to me the need for modeling entropy rain during my time in Boulder some 24 years ago. This work was supported in part by the Swedish Research Council grant No. 2012-5797, and the Research Council of Norway under the FRINATEK grant 231444. This work utilized the Janus supercomputer, which is supported by the National Science Foundation (award number CNS-0821794), the University of Colorado Boulder, the University of Colorado Denver, and the National Center for Atmospheric Research. The Janus supercomputer is operated by the University of Colorado Boulder.

APPENDIX A POLYTROPIC STRATIFICATION FROM KRAMERS OPACITY

We show here that, for the non-convecting reference solution, using the Kramers-like opacity law of Equation (2), but not the combined opacity law of Equation (6), we have $K \rightarrow \text{const}$ in the deeper, optically thick layers. Dividing

Equation (3) by the equation for hydrostatic equilibrium, $dP/dz = -\rho g$, we have

$$\begin{aligned} \frac{dT}{dP} &= \frac{F_{\text{rad}}}{K\rho g} = \frac{F_{\text{rad}}}{K_0\rho_0 g} \frac{(\rho/\rho_0)^a}{(T/T_0)^{3-b}} \\ &= \frac{F_{\text{rad}}}{K_0\rho_0 g} \frac{(P/P_0)^a}{(T/T_0)^{3+a-b}}, \end{aligned} \quad (67)$$

where $K_0 = 16\sigma_{\text{SB}}T_0^3/(3K_0\rho_0)$ is a constant and $P/P_0 = (\rho/\rho_0)(T/T_0)$ is the ideal gas equation with a suitably defined constant $P_0 = (c_p - c_v)\rho_0 T_0$. Here, ρ_0 and T_0 are reference values that were defined in Equation (2). Equation (67) can be integrated to give

$$(T/T_0)^{4+a-b} = (n+1)\nabla_{\text{rad}}^{(0)}(P/P_0)^{1+a} + (T_{\text{top}}/T_0)^{4+a-b}, \quad (68)$$

where $\nabla_{\text{rad}}^{(0)} = F_{\text{rad}}P_0/(K_0T_0\rho_0g)$, which is defined analogously to the ∇_{rad} without superscript (0) in Equations (38) and (41), and T_{top} is an integration constant that is specified such that $T \rightarrow T_{\text{top}}$ as $P \rightarrow 0$. Note also that $4+a-b = (n+1)(1+a)$, where n was defined in Equation (5) as the polytropic index, so the ratio of $4+a-b$ to $1+a$ is just $n+1$, which enters in front of the $\nabla_{\text{rad}}^{(0)}$ term in Equation (68). Since $K \propto T^{3-b}/\rho^{1+a} \propto T^{4+a-b}/P^{1+a}$, we have $K \rightarrow \text{const} = K_0$ for $T \gg T_{\text{top}}$.

APPENDIX B DERIVATION OF EQUATIONS (10) AND (11)

To obtain Equations (11) and (12), which are used to derive the Deardorff flux term in the τ approximation, we start with the equations for specific entropy and velocity in the form (see, e.g., Barekat & Brandenburg 2014)

$$\rho T \frac{DS}{Dt} = -\nabla \cdot \mathbf{F}_{\text{rad}}, \quad (69)$$

$$\rho \frac{DU}{Dt} = -\nabla P + \rho g, \quad (70)$$

where $D/Dt = \partial/\partial t + \mathbf{U} \cdot \nabla$ is the advective derivative, \mathbf{F}_{rad} is the radiative flux, and viscosity has been omitted. Subtracting those equations from their averaged ones, we obtain the following set of equations

$$\frac{\partial s}{\partial t} + \bar{U}_j \frac{\partial s}{\partial x_j} + u_j \frac{\partial \bar{s}}{\partial x_j} = -\frac{1}{\bar{\rho} \bar{T}} \nabla \cdot \delta \mathbf{F}_{\text{rad}} + \mathcal{N}_s, \quad (71)$$

$$\begin{aligned} \frac{\partial u_i}{\partial t} + \bar{U}_j \frac{\partial u_i}{\partial x_j} + u_j \frac{\partial \bar{U}_i}{\partial x_j} \\ = -\frac{1}{\bar{\rho}} \nabla p + g_i(p/\gamma \bar{P} - s/c_p) + \mathcal{N}_{ui}, \end{aligned} \quad (72)$$

where we have used $\delta\rho/\bar{\rho} = p/\gamma \bar{P} - s/c_p$, which can be obtained from Equation (8) and the perfect gas law by linearization. Pressure fluctuations will again be neglected and $(\nabla \cdot \delta \mathbf{F}_{\text{rad}})/\bar{\rho} \bar{T}$ will be replaced by $-s/\tau_{\text{cool}}$, as explained in Section 3. Assuming $\bar{\mathbf{U}} = \mathbf{0}$ and omitting the nonlinear terms \mathcal{N}_s and \mathcal{N}_{ui} we arrive at Equations (11) and (12).

APPENDIX C FILLING FACTOR FOR A DESCENDING HILL VORTEX

The solution for a Hill vortex with radius a_H and propagation velocity u_H is given by a stream function Ψ in spherical coordinates (r, θ, ϕ) as (e.g., Moffatt & Moore 1978)

$$\Psi = \frac{1}{4} \begin{cases} -3u_H(1 - r^2/a_H^2)\varpi & (\text{for } r < a_H), \\ +2u_H(1 - a_H^3/r^3)\varpi & (\text{for } r > a_H). \end{cases} \quad (73)$$

We apply it in Cartesian coordinates as the initial condition for the mass flux as $\rho \mathbf{u} = \nabla \times (\Psi \hat{\phi})$, where $\hat{\phi} = (-y/\varpi, x/\varpi, 0)$ is the unit vector in the toroidal direction of the vortex, using $\varpi^2 = x^2 + y^2$ and $r^2 = \varpi^2 + z^2$ for cylindrical radius ϖ and spherical radius r . We adopt here an isothermal equation of state, i.e., there is no buoyancy force in this problem. For non-isothermal calculations, but in two dimensions, we refer to the work of Rast (1998). Density and pressure fall off exponentially with height and both the scale height and the sound speed are independent of height. No analytic solution exists in that case, so the Hill vortex solution is at best approximate. We consider a domain of size $L \times L \times 4L$ with $-L/2 < x, y/H_p < L/2$ and $-7L/2 < z < L/2$, where $L = 5H_p$. We choose $a_H = 0.5H_p$ and $u_H = 0.2c_s$. The viscosity is $\nu = 5 \times 10^{-5}$, so the Reynolds number is $a_H u_H / \nu = 2000$. We use the PENCIL CODE⁷ with a resolution of $1152 \times 1152 \times 4608$ meshpoints.

Figure 9 shows snapshots zoomed into the vortex as it traverses about five scale heights. The filling factor, which is proportional to the radius squared, decreases with depth and is found to scale with the surrounding density like Equation (23) with $\zeta = 0.8$; see the last panel of Figure 9.

APPENDIX D MLT RELATION BETWEEN u_{rms} AND s_{rms}

To find the relation between u_{rms} and the rms values of temperature or entropy fluctuations, it is customary in standard MLT to approximate the steady state momentum equation, $\mathbf{u} \cdot \nabla \mathbf{u} \approx -g \delta T / \bar{T}$, by $u_{\text{rms}}^2 k_f = a_{\text{MLT}} g \delta T / \bar{T}$, where $a_{\text{MLT}} \approx 1/8$ is a commonly adopted geometric factor (e.g., Spruit 1974). This leads to

$$s_{\text{rms}}/c_p = \delta T / T = u_{\text{rms}}^2 k_f / a_{\text{MLT}} g = \gamma \text{Ma}^2 k_f H_p / a_{\text{MLT}}, \quad (74)$$

where we have used $c_s^2 = \gamma g H_p$, and thus $(s_{\text{rms}}/c_p) \text{Ma}^{-2} = \gamma k_f H_p / a_{\text{MLT}}$. Furthermore, using $\nabla_{\text{ad}} c_p \bar{T} = g H_p$, we have $\bar{T} s_{\text{rms}} = u_{\text{rms}}^2 k_f H_p / (a_{\text{MLT}} \nabla_{\text{ad}})$, which is used to derive Equation (31).

APPENDIX E ESTIMATE FOR THE SURFACE VALUE OF ∇_D

The purpose of this section is to show that ∇_D is a certain fraction of $\nabla - \nabla_{\text{ad}}$ in the top layers, as stated in Equation (32). To have an estimate for \bar{s}^2 , we multiply Equation (11) by s and average, so we get

$$\frac{1}{2} \frac{\partial \bar{s}^2}{\partial t} = -\bar{u}_j \bar{s} \nabla_j \bar{s} - \bar{s}^2 / \tau_{\text{cool}} + \mathcal{T}_s. \quad (75)$$

⁷ <https://github.com/pencil-code>

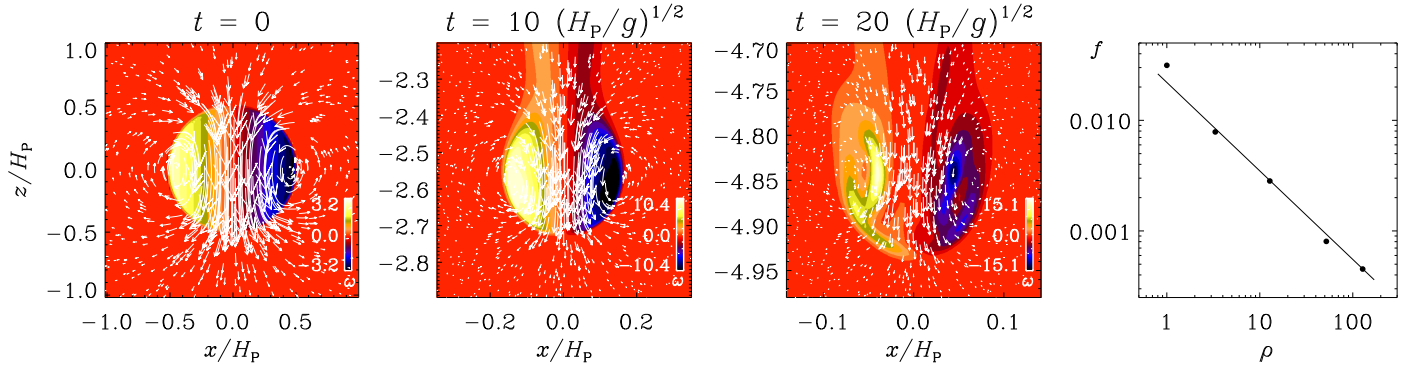


Figure 9. Velocity vectors superimposed on a color scale representation of the vorticity ω at times 0, 10, and 20 in units of $(H_p/g)^{1/2}$, as well as the resulting filling factor vs. density. The negative slope is $\zeta = 0.8$. Note that the frame of view changes as the vortex descends and shrinks.

As for F_{enth} in Equations (14) and (15), we have a triple correlation term, which is here $\mathcal{T}_s = -\overline{su} \cdot \nabla s$. Again, we adopt the τ approximation and replace \mathcal{T}_s by a damping term of the form $\mathcal{T}_s = -\overline{s^2}/\tau$, which, together with the τ_{cool} term, combines to give τ_{red} as the relevant timescale. Assuming a statistically steady state, $\partial \overline{s^2}/\partial t = 0$, we derive the following expression for $\overline{s^2}$:

$$\overline{s^2} = -\tau_{\text{red}} \overline{u_j s} \nabla_j \overline{s} \quad (\text{in the surface layers}). \quad (76)$$

This shows that fluctuations of specific entropy are produced when there is an outward flux ($\overline{u_z s} > 0$) and a locally negative (unstable) mean entropy gradient; see also Garaud et al. (2010) for a similar derivation. Inserting this into Equation (18) and using Equation (9), we obtain

$$F_D = \tau_{\text{red}}^2 \mathbf{g} \cdot (\mathbf{F}_{\text{enth}} \cdot \nabla \overline{s} / c_p) \quad (\text{in the surface layers}). \quad (77)$$

Here, both \mathbf{g} and $\nabla \overline{s}$ point downward, so F_D points upward and we can write $F_D = \lambda F_{\text{enth}}$, where $\lambda = (\tau_{\text{red}}/\tau_{\text{ff}})^2 (\nabla - \nabla_{\text{ad}})$ is a coefficient that itself is proportional to the superadiabatic gradient, and $\tau_{\text{ff}} = (H_p/g)^{1/2}$ is the free-fall time (after which a fluid parcel at rest has reached a depth of $H_p/2$). Since $F_{\text{enth}} = F_G + F_D$, this implies $F_G = (1 - \lambda)F_{\text{enth}}$ and therefore $\lambda < 1$ in the highly unstable layer at the top, where both F_G and F_{enth} are positive. We expect F_D to be largest just a few hundred kilometers below the photosphere, so λ should be maximum in the upper parts. To calculate the fraction $\nabla_D/(\nabla - \nabla_{\text{ad}})$, we use the fact that $F_{\text{enth}} = F_G + F_D$, together with the part of Equation (20) that relates to F_D , to write

$$\begin{aligned} F_D &= \frac{1}{3} \bar{\rho} c_p \bar{T} (\tau_{\text{red}} u_{\text{rms}}^2 / H_p) \nabla_D \\ &= (\tau_{\text{red}}/\tau_{\text{ff}})^2 (\nabla - \nabla_{\text{ad}}) F_{\text{enth}}. \end{aligned} \quad (78)$$

Starting with Equation (31) and expressing τ_{red} in terms of σ using Equation (47), we find

$$\nabla_D^{\text{max}} / (\nabla - \nabla_{\text{ad}})_{\text{max}} = 3\sigma \phi_{\text{enth}} \nabla_{\text{ad}} \alpha_{\text{mix}}^{-\tilde{\beta}} (k_{\text{f0}} H_p)^{-(1-\tilde{\beta})} \approx 3, \quad (79)$$

where we have used $\tilde{\beta} = 1$ (appropriate to the near-surface layers), ignored radiative cooling ($\sigma = 1$), and assumed $\alpha_{\text{mix}} = 1.6$ (Stix 2002), as well as $\phi_{\text{enth}} \approx 4$ that was found in simulations of Brandenburg et al. (2005), as discussed in Section 3.3.

We emphasize that these equations only characterize the *initiation* of entropy rain. They cannot be used to compute the Deardorff flux in the deeper layers where we have instead invoked Spruit's concept of nonlocal convection in the form of threads. The $\overline{s^2}$ associated with those threads is likely to result from a part of the triple correlation term $\mathcal{T}_s = -\overline{su} \cdot \nabla s$, which gives rise to a negative divergence of the flux of $\overline{s^2}$ of the form $\mathbf{H} \equiv \overline{us^2}$ on the rhs of Equation (75). This flux (which should not be confused with an energy flux) should point downward (s^2 is largest when $u_z < 0$) and be strongest in the upper layers, so $\nabla \cdot \mathbf{H} < 0$, leading to a positive contribution from $-\nabla \cdot \mathbf{H}$ on the rhs of Equation (75). This may explain the $\overline{s^2}$ associated with the ∇_D term.

To estimate the ratio F_D/F_G in the deeper layers, we use Equation (75) in the steady state with $1/\tau_{\text{cool}} \rightarrow 0$ and $\mathcal{T}_s = -\nabla \cdot \mathbf{H} \approx (2\zeta - \Delta\zeta) \overline{s^2} u_{\text{rms}} / \gamma H_p$ and obtain

$$0 = -\overline{u_j s} \nabla_j \overline{s} + (2\zeta - \Delta\zeta) \overline{s^2} u_{\text{rms}} / \gamma H_p \quad (\text{in deeper layers}). \quad (80)$$

Multiplying by $\frac{1}{3} \tau_{\text{red}} u_{\text{rms}}^2 (\bar{\rho} \bar{T})^2$, using Equations (9) and (17), and expanding the fraction by g/c_p , we have

$$0 = F_{\text{enth}} F_G + \frac{1}{3} \bar{\rho} u_{\text{rms}}^3 (2\zeta - \Delta\zeta) (\tau_{\text{red}} \overline{s^2} g \bar{\rho} \bar{T}) c_p \bar{T} / (\gamma g H_p). \quad (81)$$

This implies that $F_G < 0$ in the deeper layers. The second term in the parentheses is F_D (with a plus sign, because $g > 0$ is a scalar here); see Equation (18). Furthermore, we use $c_p \bar{T} / (\gamma g H_p) = 1/(\gamma - 1)$ and $\bar{\rho} u_{\text{rms}}^3 = F_{\text{enth}} / \phi_{\text{enth}}$; see Equation (31). The F_{enth} terms on both sides cancel, so we have

$$F_D / |F_G| = 3(\gamma - 1) / [\phi_{\text{enth}} (2\zeta - \Delta\zeta)] \approx 0.3 \dots 0.5, \quad (82)$$

where $2\zeta - \Delta\zeta = 2\tilde{\zeta} + \Delta\zeta = 10/9$ for Cases I and II, and $14/9$ for Case III (Section 4.3), and $\phi_{\text{enth}} = 4$ has been assumed.

APPENDIX F DERIVATION OF EXPRESSION FOR c_0

To find the coefficient c_0 given by Equation (49), we use Equation (74) in the form

$$c_p \delta T = u_{\text{rms}}^2 k_{\text{f}} H_p / \nabla_{\text{ad}} a_{\text{MLT}}. \quad (83)$$

Inserting this into Equation (7), and using Equation (45), yields

$$F_{\text{enth}} = \bar{\rho} u_{\text{rms}}^3 \alpha_{\text{mix}}^{\beta} (k_{f0} H_P)^{1-\beta} / \nabla_{\text{ad}} a_{\text{MLT}}. \quad (84)$$

Equating this expression with Equation (43), using Equation (46), we can derive the desired expression for u_{rms} in the form

$$u_{\text{rms}}^2 = \frac{\sigma}{3} \frac{c_s^2}{\gamma} \frac{a_{\text{MLT}}}{(k_{f0} H_P)^{2-2\beta'}} (\nabla - \nabla_{\text{ad}} + \nabla_{\text{D}}), \quad (85)$$

where $\beta' = (\beta + \tilde{\beta})/2$ and $c_s^2 = \gamma g H_P$ have been used. We thus find the coefficient c_0 as stated in Equation (49).

APPENDIX G RAYLEIGH NUMBER

The purpose of this section is to show that ϵ_* , as defined in Equation (51), is related to the Rayleigh number Ra. In laboratory and numerical studies of convection, it is customary to define Ra as (e.g., Käpylä et al. 2009)

$$\begin{aligned} \text{Ra} &= \frac{g d^4}{\nu \chi} \left(-\frac{d\bar{S}/c_P}{dz} \right)_{z_*}^{\text{non-conv}} \approx \frac{g d^4}{\nu \chi} \left(\frac{\nabla_{\text{rad}} - \nabla_{\text{ad}}}{H_P} \right)_{\text{max}}^{\text{non-conv}} \\ &\approx \frac{g d^4}{\nu \chi} \left(\frac{\nabla_{\text{rad}} - \nabla_{\text{ad}}}{H_P} \right)_{\text{max}}, \end{aligned} \quad (86)$$

which is usually evaluated in the middle of the layer at $z = z_*$. In the Sun, however, the maximum value is more relevant. The superscript “non-conv” indicates that the entropy gradient is taken for the non-convecting reference state, d is the thickness of the layer and ν is the viscosity. Introducing the Prandtl number $\text{Pr} = \nu/\chi$ and using the definition of Nu in Equation (42), we have

$$\frac{\text{Pr Ra}}{\text{Nu} - 1} = \frac{g d^4 \nabla_{\text{ad}}}{\chi^2 H_P}. \quad (87)$$

On the other hand, using Equation (51), we find

$$\begin{aligned} \epsilon_*^2 &= \frac{\sigma^3 a_{\text{MLT}}}{27 \gamma^3} \frac{(c_s^3 / \chi g)^2}{(k_{f0} H_P)^{4(1-\beta'')}} = \frac{\sigma^3 a_{\text{MLT}}}{27 (k_{f0} H_P)^{4(1-\beta'')}} \frac{H_P^3 g}{\chi^2} \\ &= \frac{\sigma^3 a_{\text{MLT}} (H_P/d)^4}{27 \nabla_{\text{ad}} (k_{f0} H_P)^{4(1-\beta'')}} \frac{g d^4 \nabla_{\text{ad}}}{\chi^2 H_P}, \end{aligned} \quad (88)$$

and therefore

$$\epsilon_*^2 = \frac{\sigma^3 a_{\text{MLT}} / 27 \nabla_{\text{ad}}}{(k_{f0} d)^4 (k_{f0} H_P)^{-4\beta''}} \frac{\text{Pr Ra}}{\text{Nu} - 1}, \quad (89)$$

which shows that Ra is proportional to ϵ_*^2 .

REFERENCES

- Arnett, W. D., Meakin, C., Viallet, M., et al. 2015, *ApJ*, **809**, 30
 Asplund, M., Grevesse, N., Sauval, A. J., Allende Prieto, C., & Kiselman, D. 2004, *A&A*, **417**, 751
 Barekat, A., & Brandenburg, A. 2014, *A&A*, **571**, A68
 Basu, S. 1997, *MNRAS*, **288**, 572
 Biermann, L. 1932, *ZA*, **5**, 117
 Biermann, L. 1938, *AN*, **264**, 395
 Blackman, E. G., & Field, G. B. 2003, *PhFI*, **15**, L73
 Bogart, R. S., Baldner, C. S., & Basu, S. 2015, *ApJ*, **807**, 125
 Böhm-Vitense, E. 1958, *ZA*, **46**, 108
 Brandenburg, A., Chan, K. L., Nordlund, Å., & Stein, R. F. 2005, *AN*, **326**, 681
 Brandenburg, A., Jennings, R. L., Nordlund, Å., et al. 1996, *JFM*, **306**, 325
 Brandenburg, A., Käpylä, P. J., & Mohammed, A. 2004, *PhFI*, **16**, 1020
 Brandenburg, A., Rädler, K.-H., & Schinner, M. 2008, *A&A*, **482**, 739
 Buske, D., Vilhena, M. T., Moreira, D. M., & Tirabassi, T. 2007, *EFM*, **7**, 43
 Castaing, B., Gunaratne, G., Heslot, F., et al. 1989, *JFM*, **204**, 1
 Cattaneo, F., Brummell, N. H., Toomre, J., Malagoli, A., & Hurlburt, N. E. 1991, *ApJ*, **370**, 282
 Chatterjee, P., & Antia, H. M. 2009, *ApJ*, **707**, 208
 Christensen-Dalsgaard, J., Gough, D. O., & Thompson, M. J. 1991, *SoPh*, **378**, 413
 Cossette, J.-F., & Rast, M. P. 2016, *ApJL*, **829**, L17
 Davidson, P. A. 2004, *Turbulence: An Introduction for Scientists and Engineers* (Oxford: Oxford Univ. Press)
 Deardorff, J. W. 1966, *JATIS*, **23**, 503
 Deardorff, J. W. 1972, *JGR*, **77**, 5900
 De Rooze, S. R., Jonker, H. J. J., Duynkerke, P. G., & Stevens, B. 2004, *BoLMe*, **112**, 179
 Duvall, T. L., Jr., Kosovichev, A. G., Scherrer, P. H., et al. 1997, *SoPh*, **170**, 63
 Edwards, J. M. 1990, *MNRAS*, **242**, 224
 Ertel, H. 1942, *MetZe*, **59**, 250
 Featherstone, N. A., & Hindman, B. W. 2016, *ApJ*, **818**, 32
 Freytag, B., Steffen, M., Ludwig, H.-G., et al. 2012, *JCoPh*, **231**, 919
 Garaud, P., Ogilvie, G. I., Miller, N., & Stellmach, S. 2010, *MNRAS*, **407**, 2451
 Gizon, L., & Birch, A. C. 2005, *LRSP*, **2**, 6
 Gizon, L., & Birch, A. C. 2012, *PNAS*, **109**, 11896
 Greer, B. J., Hindman, B. W., Featherstone, N. A., & Toomre, J. 2015, *ApJL*, **803**, L17
 Grevesse, N., & Sauval, A. J. 1998, *SSRv*, **85**, 161
 Gudiksen, B. V., Carlsson, M., Hansteen, V. H., et al. 2011, *A&A*, **531**, A154
 Hanasoge, S., Gizon, L., & Sreenivasan, K. R. 2016, *AnRFM*, **48**, 191
 Hanasoge, S. M., Duvall, T. L., & Sreenivasan, K. R. 2012, *PNAS*, **109**, 11928
 Hanasoge, S. M., Duvall, T. L., Jr., & DeRosa, M. L. 2010, *ApJL*, **712**, L98
 Hanasoge, S. M., & Sreenivasan, K. R. 2014, *SoPh*, **289**, 3403
 Hathaway, D. H., Upton, L., & Colegrove, O. 2013, *Sci*, **342**, 1217
 Heslot, F., Castaing, B., & Libchaber, A. 1987, *PhRvA*, **36**, 5870
 Hill, M. J. M. 1894, *RSPTA*, **185**, 213
 Hotta, H., Rempel, M., & Yokoyama, T. 2015, *ApJ*, **803**, 42
 Hurlburt, N. E., Toomre, J., & Massaguer, J. M. 1984, *ApJ*, **282**, 557
 Hurlburt, N. E., Toomre, J., & Massaguer, J. M. 1986, *ApJ*, **311**, 563
 Hurlburt, N. E., Toomre, J., Massaguer, J. M., & Zahn, J. P. 1994, *ApJ*, **241**, 245
 Käpylä, P. J., Korpi, M. J., & Brandenburg, A. 2009, *ApJ*, **697**, 1153
 Käpylä, P. J., Mantere, M. J., Cole, E., Warnecke, J., & Brandenburg, A. 2013, *ApJ*, **778**, 41
 Kippenhahn, R., & Weigert, A. 1990, *Stellar Structure and Evolution* (Berlin: Springer)
 Lavelly, E. M., & Ritzwoller, M. H. 1993, *SoPh*, **403**, 810
 Lord, J. W. 2014, PhD thesis, Univ. Colorado
 Lord, J. W., Cameron, R. H., Rast, M. P., Rempel, M., & Roudier, T. 2014, *ApJ*, **793**, 24
 Losada, I. R., Brandenburg, A., Kleeorin, N., & Rogachevskii, I. 2013, *A&A*, **556**, A83
 Miesch, M. S., Brun, A. S., De Rosa, M. L., & Toomre, J. 2008, *ApJ*, **673**, 557
 Miesch, M. S., Featherstone, N. A., Rempel, M., & Trampedach, R. 2012, *ApJ*, **757**, 128
 Mihalas, D. 1978, *Stellar Atmospheres* (San Francisco: Freeman)
 Moffatt, H. K., & Moore, D. W. 1978, *JFM*, **87**, 749
 Nordlund, Å., Stein, R. F., & Asplund, M. 2009, *LRSP*, **6**, 2
 Pleim, J. E. 2007, *JApMC*, **46**, 1383
 Priestley, C. H. B., & Swinbank, W. C. 1947, *RSPSA*, **189**, 543
 Rast, M. P. 1998, *JFM*, **369**, 125
 Rempel, M. 2004, *ApJ*, **607**, 1046
 Rheinhardt, M., & Brandenburg, A. 2012, *AN*, **333**, 71
 Rieutord, M., Meunier, N., Roudier, T., et al. 2008, *A&A*, **479**, L17
 Rieutord, M., & Zahn, J.-P. 1995, *A&A*, **296**, 127
 Rogachevskii, I., & Kleeorin, N. 2015, *JPIPh*, **81**, 395810504
 Roudier, T., Rieutord, M., Malherbe, J. M., et al. 2012, *A&A*, **540**, A88
 Rüdiger, G. 1989, *Differential Rotation and Stellar Convection: Sun and Solar-type Stars* (New York: Gordon and Breach)
 Rüdiger, G., & Spahn, F. 1992, *SoPh*, **138**, 1
 Schou, J., Antia, H. M., Basu, S., et al. 1998, *ApJ*, **505**, 390

- Serenelli, A. M., Basu, S., Ferguson, J. W., & Asplund, M. 2009, [ApJL](#), **705**, L123
- Spruit, H. 1997, *MmSAI*, **68**, 397
- Spruit, H. C. 1974, [SoPh](#), **34**, 277
- Spruit, H. C. 1977, *A&A*, **55**, 151
- Stein, R. F., Benson, D., & Nordlund, Å. 2007, in ASP Conf. Ser. 369, New Solar Physics with Solar-B Mission, ed. K. Shibata, S. Nagata, & T. Sakurai (San Francisco, CA: ASP), 87
- Stein, R. F., Brandenburg, A., & Nordlund, Å. 1992, in ASP Conf. Ser. 26, Cool Stars, Stellar Systems, and the Sun, ed. M. S. Giampapa & J. A. Bookbinder (San Francisco, CA: ASP), 148
- Stein, R. F., & Nordlund, Å. 1989, [ApJL](#), **342**, L95
- Stein, R. F., & Nordlund, Å. 1998, [ApJ](#), **499**, 914
- Stein, R. F., Nordlund, Å., Georgoviani, D., Benson, D., & Schaffenberger, W. 2009, in ASP Conf. Ser. 416, Solar-Stellar Dynamos as Revealed by Helio- and Asteroseismology: GONG 2008/SOHO 21, ed. M. Dikpati et al. (San Francisco, CA: ASP), 421
- Stix, M. 1981, *A&A*, **93**, 339
- Stix, M. 2002, *The Sun: An Introduction* (Berlin: Springer)
- Stull, R. B. 1984, [JAtS](#), **41**, 3351
- Stull, R. B. 1993, [BoLMe](#), **62**, 21
- Trampedach, R., & Stein, R. F. 2011, [ApJ](#), **731**, 78
- Tuominen, I., Brandenburg, A., Moss, D., & Rieutord, M. 1994, *A&A*, **284**, 259
- Tuominen, I., & Rüdiger, G. 1989, *A&A*, **217**, 217
- Unno, W., & Spiegel, E. A. 1966, *PASJ*, **18**, 85
- van Ballegoijen, A. A. 1986, [ApJ](#), **304**, 828
- van Dop, H., & Verver, G. 2001, [JAtS](#), **58**, 2240
- Vitense, E. 1953, *ZA*, **32**, 135
- Vögler, A., Shelyag, S., Schüssler, M., et al. 2005, [A&A](#), **429**, 335
- Woodard, M. 2014, [SoPh](#), **289**, 1085
- Woodard, M. 2016, [MNRAS](#), **460**, 3292
- Xiong, D. R., & Deng, L. 2001, [MNRAS](#), **327**, 1137