Title
The Effects of Small-Scale Mixing Processes on Supernova Progenitors

Author
Brown, Justin

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

THE EFFECTS OF SMALL-SCALE MIXING PROCESSES ON SUPERNOVA PROGENITORS

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

ASTRONOMY & ASTROPHYSICS

by

Justin M. Brown

June 2016

The Dissertation of Justin M. Brown
is approved:

Professor Pascale Garaud, Chair

Professor Stan E. Woosley

Professor Nic Brummell

Dean Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

The Effects of Small-Scale Mixing Processes on Supernova Progenitors

by

Justin M. Brown

Minor mixing processes—any fluid processes that mix material or transport heat other than convection or other large-scale flows—play a critical role in stellar evolution. These processes have been invoked through phenomenological models in order to explain away many issues in stellar evolution, such as the blue-red supergiant ratio problem and the progenitor problem of SN 1987A. We discuss our incompressible numerical simulations of one such process, thermohaline convection. We developed a semi-analytic formulation of the thermal and compositional transport of this process, which has been tested by other groups and implemented into MESA and the Toulouse–Geneva Evolution code. We also discuss compressible simulations of overshooting convection by Brummell *et al.* (2002) and the incompressible simulations of semi-convection by Wood *et al.* (2013), from which we developed one-dimensional models. We present some recent results of the effects of these models on a 15 solar mass star using the stellar evolution code KEPLER. These simulations put stricter limits on the extent of overshooting convection and the efficiency of semi-convection.

Acknowledgments

I would like to thank the tireless efforts of my advisers, Stan Woosley and Pascale Garaud. Pascale's expertise and attention to detail have been invaluable to me during my graduate career, and I'm grateful for the freedom I've been granted to pursue topics of my own interest. Working with her has opened up entirely new avenues for me to apply my expertise of fluid dynamics to a broad range of topics.

Stan has one of the best intuitions for science that I have ever encountered, and working with him has been a pleasure. I've learned so much in how to tackle problems in different ways than I initially anticipated and in how to defend my scientific ideas.

Thank you both. I'm looking forward to collaborating with you in the future!

I would also like to thank my collaborators, Nic Brummell, Stephan Stellmach, Timour Radko, and Pawel Artymowytcz for their discussions and insight upon a wide breadth of topics.

I am grateful for my fellow graduate students have supported me through graduate school, and I am grateful to have had the opportunities to work with the brightest young minds I've known. Katie Hamren, Alex Rudy, Camille Leibler, Elizabeth Lovegrove, Caroline Morley, Jennifer Burt, Claire Dorman, Rosalie McGurk, Emily Cunningham, and Zach Jennings have all been invaluable in reminding me—occasionally on a daily basis—to stay sane.

The text of this dissertation includes reprints of the following previously published material: Brown, J. M.; Garaud, P. and Stellmach, S. Chemical Transport and Spon-

taneous Layer Formation in Fingering Convection in Astrophysics. *The Astrophysical Journal* **768**, 16 (2013); Brown, J. M. and Woosley, S. E. NUCLEOSYNTHETIC CONSTRAINTS ON THE MASS OF THE HEAVIEST SUPERNOVAE. *The Astrophysical Journal* **769**, 99 (2013); Sukhbold, T.; Ertl, T.; Woosley, S. E.; Brown, J. M. and Janka, H. T. CORE-COLLAPSE SUPERNOVAE FROM 9 TO 120 SOLAR MASSES BASED ON NEUTRINO-POWERED EXPLOSIONS. *The Astrophysical Journal* **821**, 38 (2016). This work has been included with the consent of the relevant authors and co-authors. The co-authors Pascale Garaud and Stan Woosley directed and supervised the research which forms the basis for the dissertation.

This research would not be possible without the support of a number of grants and fellowships. The author has received support from a UCSC Regent's Fellowship and several NSF grants, including AST-0807672 and AST-1412951. This research is supported in part by the Department of Energy Office of Science Graduate Fellowship Program (DOE SCGF), made possible in part by the American Recovery and Reinvestment Act of 2009, administered by ORISE–ORAU under contract no. DE-AC05-06OR23100.

Chapter 1

Introduction

Mixing processes have consequences throughout astronomy through their effects on stellar evolution, including the duration of each phase of stellar evolution, the observable qualities of supernova progenitors, and the nucleosynthesis taking place within these objects. Minor mixing processes, which occur on scales much smaller than the size of the star, can play a critical role in stellar evolution and remain a major source of uncertainty. This thesis work focuses on three such processes: fingering convection, semi-convection, and overshooting convection.

The application and properties of massive star evolution is briefly discussed in Section 1.1. The nature and significance of semi-convection is discussed in Section 1.2, that of fingering convection in Section 1.3, and that of overshooting convection in Section 1.4.

1.1 Introduction to Evolution of Massive Stars

Though massive stars[1] are intrinsically rare in stellar populations, they can critically affect both nearby stars through interactions with the supernova ejecta and future populations through the metallicity enrichment caused by nucleosynthesis. This thesis will focus mostly on understanding the effects of mixing on stellar evolution and nucleosynthesis and will not address supernovae directly. For a thorough review of the field of massive star evolution and supernovae, the reader is referred to Woosley *et al.* (2002).

The internal properties of massive stars late in their evolution distinguishes them from more common stars. The larger masses of these objects require higher luminosities for support against the force of gravity, which in turn requires more rapid exhaustion of fuel. Upon exhausting helium, massive stars proceed to fuse heavier elements and thus provide an avenue for producing many of the intermediate-mass elements (i.e., sodium through calcium) necessary to life on Earth. The nucleosynthesis of massive stars and supernovae is a rich topic, and is covered in greater detail in Chapter 2.

One of the major difficulties in theoretical studies of massive star evolution is the treatment of hydrodynamical mixing processes, including convection, semi-convection, fingering convection, convective overshoot, and rotationally-induced mixing. This thesis work will focus on numerical studies of a few of these processes with the inten-

[1] In the context of this thesis, massive stars are those stars expected to undergo a supernova.

tion of applying them to massive star interiors. Fully resolving the hydrodynamics of stellar interiors—even for domains much smaller than the entire star—remains impossible for the astropyhysical parameter regime; however, partial progress has recently been made and is summarized for semi-convection, fingering convection, and convective overshoot, respectively, in the next three sections.

1.2 Introduction to Semi-Convection

Semi-convection is a form of "double-diffusive convection," which occurs when a fluid is stable in density but unstable in temperature and has two components (e.g., temperature and composition) that diffuse at different rates. The behavior in such a region was first studied by Schwarzschild and Härm (1958), and the stability was later analyzed by Walin (1964) for the oceanographic case and by Kato (1966) for the astrophysical case. Over the years, the understanding of this process has improved drastically; however, the astrophysical and fluids communities have now diverged substantially, with most astrophysicists being satisfied with modeling using the analytical semi-convection prescription from Langer *et al.* (1983) while fluid dynamicists continue to expand their understanding through numerical simulations (e.g., the works of Mirouh *et al.*, 2012; Zaussinger and Spruit, 2013; Wood *et al.*, 2013). For further progress to be made, the next generation of semi-convection models in stars should use the current understanding of the physics of semi-convection from hydrodynamical simulations.

1.2.1 Physics

Like convection, semi-convection can be viewed as a fluid instability: in the case of convection, the instability is exponential in nature; however, as described in Kato (1966), the fundamental instability of semi-convection is both exponential and oscillatory. When examining a very small semi-convective region of the star, the temperature and compositional profiles are well approximated as being monotonically decreasing functions of radius. When a parcel of high entropy, high metallicity fluid is displaced upward into a low entropy, lower metallicity environment, the heat diffuses and equalizes the temperature between the parcel and the surrounding substance. The parcel, now denser, returns to its original location but can potentially sink below it thanks to the potential energy acquired. As the parcel is now of lower entropy and lower metallicity than its surroundings, it absorbs surrounding heat and becomes lighter, returning to its original position. The additional buoyancy provided by heating propels it even further than before, and the amplitudes of these oscillations grow exponentially with time. What has just been described is the fundamental or primary instability, which occurs on small scales; however, knowing the nature of the fundamental linear instability does not imply knowledge of how semi-convection behaves nonlinearly and/or on large scales.

Of great interest to fluid dynamicists, semi-convective regions can sometimes exhibit various behaviors on large spatial scales (Merryfield, 1995; Mirouh *et al.*, 2012; Wood *et al.*, 2013), including large-scale gravity waves and convective layers. Of these, the last has received much scientific attention. Turner and Stommel (1964) first observed these layers in salty water, when they showed through laboratory experiments that semi-convective regions could take the form of stacks of efficient convecting layers. These layers have seen much experimental study since (e.g., Turner, 1965; Shirtcliffe, 1973; Noguchi and Niino, 2010; Carpenter *et al.*, 2012). Radko (2003) discovered a mean-field instability that explained the formation of layers in another case of double-diffusive convection (thermohaline convection), which was successfully applied to the semi-convective case by Rosenblum *et al.* (2011). Mirouh *et al.* (2012) showed that the conditions for layer formation in semi-convection are only realized in systems near the standard convective instability threshold and were able to determine the conditions for layer formation in various astrophysical parameter regimes. Further, Wood *et al.* (2013) empirically derived a set of flux laws for composition and thermal energy as a function of the semi-convective layer height. The latter, however, remains an unknown quantity in semi-convection theory. In all periodic Boussinesq simulations of semi-convective layers run to date, the layers are seen to merge over time until only one remains in the simulated domain. It is unclear whether the layers would continue to merge indefinitely or settle at a finite height in a real astrophysical setting. This remains an active area of research.

Spruit (2013) addressed the merging layers problem via energy flux arguments, arguing that layers equilibrate when reaching a height where the flux through the layers matches that through the interfaces. In this work, the flux through the interfaces is assumed to be diffusive in nature, which is consistent with the simulations of Zaussinger and Spruit (2013); however, the flux is primarily turbulent in the simulations of Wood *et al.* (2013). The discrepancy between the two sets of simulations could be due to boundary effects or parameter selection.

In cases where layers do not form, simulations (e.g., Mirouh *et al.*, 2012; Moll *et al.*, 2016) have also shown a tendency to form large-scale gravity waves, which mix somewhat more efficiently than the basic semi-convective state but less so than layered semi-convection.

1.2.2 Significance in Stellar Evolution

It has been known for some time that the inclusion of semi-convection can have substantial effects on the luminosities and temperatures of stars during core helium burning. In particular, Langer *et al.* (1985) showed that the blue loops—phases when a red supergiant burning helium can briefly become a blue supergiant—can be quite sensitive to the treatment of semi-convection. The strength of semi-convection controls the shape of these loops on the color–magnitude diagram and the time spent as either a red or blue supergiant. The properties of the loops are also sensitive to other processes aside from semiconvection, such as mass loss and rotation. Since the prevalence of blue loops partially determines the red-to-blue supergiant ratio in galaxies, the latter is a priori dependent on all of the aforementioned physical processes. However, as shown by Langer and Maeder (1995) using mixing prescriptions available at the time, even with a combination of all these processes, stellar evolution models were unable to reproduce observations of this ratio.

In addition to determining the final structure of a massive star, semi-convection can also substantially change the susceptibility of massive stars to explosion. The work by Sukhbold and Woosley (2014) has shown that the "compactness," a parameter that indicates the likelihood of a given stellar model to explode in simulations, depends strongly on the parameterization of semi-convection (see their Figure 20). Thus, it is of prime interest to the supernova community to use stellar models with as accurate a treatment of semi-convection as possible.

1.3 Introduction to Fingering Convection

1.3.1 Physics

The phenomenon of fingering convection (otherwise known as thermohaline convection) occurs in regions of stellar and planetary interiors where the mean molecular weight, μ, increases upwards while remaining stable to the Ledoux criterion. If not for the effects of diffusion, such a system would be stable to small perturbations. In the presence of diffusion, however, instability can occur. Indeed, when a high μ, high entropy parcel is displaced downward, heat rapidly leaks via thermal diffusion. Since compositional diffusion is much weaker, the parcel remains denser than its surroundings and proceeds to sink. The instability takes the form of finger-like structures and can significantly increase the flux of heat and composition above that of molecular or radiative diffusion.

Although the applications of interest are astrophysical, much of the formalism of fingering convection was first developed to address the phenomenon of "salt fingers" in the ocean. There, temperature and salt play the roles of the entropy and metallicity, and salt diffuses about 100 times more slowly than temperature. Stern (1960) first proposed the theory of a "salt-fountain" to explain the persistent, small-scale motions observed in warm, salty water lying above cool, fresh water even though density decreases upwards. For a recent review of the field, see Kunze (2003). By contrast with astrophysical fingering, it is possible to run laboratory experiments of fingering convection in salt water and measure its turbulent transport efficiency (e.g., Turner, 1967; Stern, 1969; Schmitt, 1979). Unfortunately, none of these laboratory experiments are applicable to the parameter regime relevant to astrophysics. Because of this, theoretical models of fingering convection in astrophysics have remained, until recently, mostly phenomenological.

Furthermore, like semi-convection, fingering convection has been shown to exhibit interesting large-scale behavior in the oceanographic literature. Not long after the initial linear instability theory was proposed, Stern (1969) realized that fingering convection had the peculiar property of driving coherent, large-scale gravity waves by an instability he termed the "collective-instability." Another possible large-scale outcome of fingering convection in the ocean is the generation of thermohaline staircases (c.f., observations in the tropical Atlantic by Schmitt, 2005). These staircases are stacks of distinct, well-mixed, convective layers separated by thin fingering interfaces. Currently, the most promising explanation for staircase formation is the γ-instability, proposed by Radko (2003) and supported by numerical simulations (Stellmach *et al.*, 2011).

Both large-scale instabilities mix material much more efficiently than "homogeneous" fingering convection in the ocean (Schmitt, 2005). The γ-instability and collective-instability theories can be applied to the astrophysical case with minor modifications (for details, see Traxler *et al.*, 2011a). However, because the viscous and compositional diffusion scales are much smaller in stars than in the ocean, it is currently infeasible to perform full 3D numerical simulations in the astrophysical regime that can resolve both the fingers and these large-scale instabilities. From a theoretical point of view, Traxler *et al.* (2011a) predicted that layers could not form by the γ-instability in astrophysics. Brown *et al.* (2013) investigated astrophysical thermohaline layers more thoroughly and posed the question of whether layers may form by the collective-instability instead. Medrano *et al.* (2014) found layer formation to be unlikely in normal stars but potentially possible in the case of white dwarfs.

1.3.2 Significance in Stellar Evolution

Fingering convection appears in several key problems in astrophysics in which an inverse μ-gradient is observed, the most notable of which are planet infall and ^{3}He fusion.

Take for instance the problem of planet infall. Stars with detected planets have higher metallicities than those without (e.g., Santos *et al.*, 2001; Fischer and Valenti, 2005). The relevant question is then whether this observation results from a superior ability of high-metallicity gas to form planets or from pollution by planet infall. In the latter scenario, it is thought that stars accrete planets onto their outer convective zones, raising the metallicity at the surface. Vauclair (2004) was the first to argue that fingering convection may play an important role in this problem. Because the mean molecular weight in the convective zone rises upon absorbing a planet, the radiative region just beneath experiences an inverse μ-gradient and becomes unstable to the fingering instability. This drains the excess metallicity from the convective zone into the radiative interior. However, the time scale for mixing by fingering convection was until recently unknown and is needed to determine how long the post-infall excess metallicity of the convective zone remains observable. Thus, in order to determine whether the planet-metallicity connection is an effect of planet formation or infall, the mixing of thermohaline convection must be quantified (see Garaud, 2011, for more detail).

During the process of ^{3}He fusion, $^3\text{He}+^3\text{He} \rightarrow {}^4\text{He}+2^1\text{H}$. This reaction causes the total number of particles to increase, so the mean molecular weight decreases (Ulrich, 1972). Because fusion occurs more frequently at higher temperatures and densities, this reaction preferentially decreases the molecular weight deeper in the star, and builds up an inverse μ-gradient. This is thought to occur in red giant branch stars. Indeed, these stars are observed to undergo a process known as dredge-up, in which particular isotopes (such as ^{3}He) produced in the deep stellar interior are advected outwards (e.g., Gilroy, 1989; Charbonnel, 1994). Simulations by Charbonnel (1994) produced surface abundances that agree well with observations for the so-called "first dredge-up." However, she noted that some additional mixing in low-mass red giants is needed to account for the observed carbon isotope ratios. Eggleton *et al.* (2006) initially suggested that a Rayleigh-Taylor instability caused by the inverse μ-gradient could provide this additional mixing. However, Charbonnel and Zahn (2007) later proposed that a fingering instability caused by ^{3}He burning was more likely.

Fingering convection is also likely to appear in massive star evolution. The nuclear processes that govern the cores of these stars are complex, and scenarios with increasing mean molecular weight do arise and can be prominent, as is seen in Figure 3.1.

1.3.3 Implementation in Stellar Codes

Several prescriptions have been suggested for the implementation of fingering convection in stellar codes over the past four decades. Ulrich (1972) and Kippenhahn *et al.* (1980) attempted to constrain the dimensions of fingers using stability arguments and determined the resulting thermal and compositional mixing timescales using dimensional analysis. Both these prescriptions have a free parameter: in the former, the free parameter is the "effective inertia of the flow;" whereas in the latter, it is the aspect ratio of the fingers. Later, Schmitt (1983) used linear theory to predict the ratio of the heat to the compositional turbulent fluxes in astrophysical fingering but could not determine the absolute fluxes directly.

Only in the last few years have numerical simulations of fingering convection approaching the astrophysical parameter regime become more readily available (e.g., Denissenkov, 2010; Traxler *et al.*, 2011a). Denissenkov (2010) ran 2D simulations to measure the aspect ratio of fingers and used the previous literature and a dimensional argument to find a prescription for mixing by fingering convection in red giant branch stars. He concluded that while this process could provide some of the required mixing discussed by Charbonnel and Zahn (2007), it alone was insufficient to account for the observed abundances of low-mass red giant branch stars. Traxler *et al.* (2011a) presented 3D numerical simulations of fingering convection and generated an empirical fit of their results to propose transport laws for fingering convection in astrophysics. Using their mixing model, Garaud (2011) studied the evolution of the surface metallicity of solar-type stars after planet infall and found that fingering convection would transport the material out of the convective zone too quickly to explain the planet-metallicity connection. This then suggests that planets can form more easily in high-metallicity proto-planetary disks. However, the results of Traxler *et al.* (2011a) also confirmed the findings of Denissenkov (2010) on the red giant branch star abundances, implying that the post-dredge-up carbon isotope ratio remains a mystery.

Finally, Medrano *et al.* (2014) showed recently that most large-scale instabilities arising from fingering convection are suppressed at astrophysical parameters. Thus, astrophysical fingering convection really is a small-scale mixing process.

1.4 Introduction to Overshooting Convection

Overshooting convection has been an important question in massive star evolution since it was originally posed by Saslaw and Schwarzschild (1965); however, the nature of this process has remained a topic of intense debate since then. Overshooting convection is a phenomenon that occurs at any radiative–convective boundary within a star, and extends mixing of both entropy and composition beyond the convective boundary. However, the amplitude, nature, and extent of this mixing are not well understood.

1.4.1 Physics

While the exact nature of overshooting convection is currently a hot subject of debate, the general notion is that, in the vicinity of a convective region, the adjacent radiative zone undergoes some additional mixing. Simulations (e.g., Hurlburt *et al.*, 1994; Brummell *et al.*, 2002; Meakin and Arnett, 2007) have shown that there are two different realizations of this process, which have become known as "overshooting" and "penetrative" convection by the fluids community (though the terms have been used interchangeably by the astronomical community). "Overshooting" convection describes a ballistic process, in which plumes pierce the boundary and mix the composition of the surrounding material slightly. "Penetrative" convection is more efficient and affects the thermal structure of the star as well, gradually entraining material into the convection zone and extending the convective boundary over time into the stable region.

Veronis (1963) analyzed the behavior of overshooting convection by comparing the results of a finite-amplitude stability analysis with laboratory experiments on water, finding that convective motion can penetrate well into the stable layer. Analyses of the nonlinear problem in both Boussinesq (as in Musman, 1968) and anelastic (as in Latour *et al.*, 1981) cases showed that convective motions do penetrate substantially into the stable region. Further work (e.g., Hurlburt *et al.*, 1986; Freytag *et al.*, 1996) simulated these scenarios in two dimensions. Hurlburt *et al.* (1986) found, in simulations with a convection zone in between two radiative zones, that plumes from the convection zone into the lower radiative zone were stronger than those into the upper one, suggesting that the extent of overshoot may depend on whether the radiation zone is above or below a convection zone. Much work has been devoted to three-dimensional simulations in the Boussinesq approximation (e.g., Singh *et al.*, 1994; Muthsam *et al.*, 1995); however, some of these, such as Singh *et al.* (1994), use a subgrid scale model, which generally is a parameterization of physics that occurs on a scale smaller than the simulation resolution. The results of such calculations can be quite sensitive to the subgrid scale model used. It should be noted that despite the great progress that has been achieved, the true nature of overshooting in stars remains unclear since many of these simulations seem to be only moderately turbulent.

The most up-to-date direct numerical simulations of overshoot in solar convection have been completed by Brummell *et al.* (2002) and Rogers *et al.* (2006) for a small, but very turbulent 3D domain, and a global, but largely laminar 2D domain, respectively. Both these groups confirm an overshoot length of approximately $0.05H_p$ beneath the solar convection zone. The work of Brummell *et al.* (2002), being one of the most resolved series of overshoot calculations, also provides a possible scaling of the overshoot length with the parameters of the system (discussed more in Section 3.2). However, both of these groups run their models in regimes far outside that of the true stellar parameter space. Alternately, Meakin and Arnett (2007) have performed 3D numerical simulations of convection in the correct stellar regime for massive stars (though this code does not include explicit viscous or chemical diffusion), finding that a mass entrainment model best represents overshoot.

The extent of overshooting convection in stars has been a topic of much debate for the past fifty years. Early investigations using mixing length theory found that the extent of overshooting was negligible. For example, Saslaw and Schwarzschild (1965) linearized the governing equations and let a quantity they term the "buoyancy factor" (a measure of the local superadiabaticity, which is quite small in the convection zones of stars) asymptote to zero. They solved for the eigenstates of the system and found that as the buoyancy factor asymptotes to zero, the overshoot extent becomes negligible; this was consistent with the findings of Roxburgh (1965), who performed a similar analysis.

However, Shaviv and Salpeter (1973) found through a nonlocal analysis that these regions might extend substantially. Shaviv and Salpeter (1973) explored the problem by adapting mixing length theory into a more realistic model. They modeled the buoyancy forces on the mixing length theory parcels accurately and let them diffuse when their velocity reaches zero or when it travels half a mixing length, regardless of whether that location is within the convection zone or not. They found that the overshoot extent was not negligible even as they approached the astrophysical regime, finding an average extent of 0.07 pressure scale heights. These results were consistent with other groups at the time (e.g., Maeder, 1975).

Some later groups attempted to model overshoot using alternate analytic theories. Schmitt *et al.* (1984) modeled the region below the solar convection zone using a simplified hydrodynamic model of stochastic plumes, and found only a small overshoot into the Sun. Xiong (1986); Xiong and Deng (2001) used a statistical non-local theory of convection and found that the overshooting was extensive, and that appropriate overshoot removed the need for semi-convection. More recently, Rempel (2004) constructed a semi-analytical entrainment theory for the behavior of overshooting convection which compared well with numerical simulations, such as those of Brummell *et al.* (2002).

1.4.2 Significance in Stellar Evolution

Overshooting convection, occurring at any radiative–convective boundary in a star, has substantial significance in stellar evolution. Around core convecting regions, it changes the amount of fuel available to the core and hence the time spent on that stage of evolution, particularly for hydrogen and helium core burning.

Many of the main observable consequences for overshooting convection occur on the main sequence. Strong overshooting convection implies older ages at the main-sequence turn-off and hence a larger ratio of main sequence stars to post-main sequence stars (Maeder and Meynet, 1989). Several groups (e.g., Bertelli *et al.*, 1990) have produced isochrones illustrating this behavior and have reported good agreement with observations of clusters. Buonanno *et al.* (1985) applied the same methodology to later stages to examine the mixing during the He burning lifetime. However, all these models use an arbitrary parameter to determine the strength of overshoot.

Below envelope convection, overshoot extends the depth at which material is dredged up to the surface for observation or how far delicate isotopes are dredged down for destruction. In red supergiants, the outer convection zone can bring unburned hydrogen into the helium layer (Stothers and Chin, 1991). This additional source of fuel can determine whether the star transitions back into a blue supergiant star. The relevance of overshooting convection to the lithium depletion in Sun-like stars has been discussed extensively in the literature (Straus *et al.*, 1976; Xiong *et al.*, 1991). More recently, Xiong and Deng (2009) concluded that overshooting is a dominant contributor in stars less than $1M_\odot$ to the lithium depletion problem.

Of most interest to people working on supernovae and nucleosynthesis, these small changes in the lifetime of the star can substantially affect the final nucleosynthesis. In particular, the s-process can be altered by up to a factor of four by adding a few percent of a pressure scale height to the overshooting extent for solar metallicity models (Pumo *et al.*, 2011); this is less extreme but still substantial at low metallicity. It can also change the final structure and compactness of the star, which has consequences not only for supernova nucleosynthesis, but also for the supernova lightcurve and even whether the star will become a supernova at all (Sukhbold and Woosley, 2014).

1.4.3 Observations

The advent of helioseismology has made it possible to attempt to directly measure the overshooting convection extent for the Sun. Most helioseismic studies (e.g., Basu, 1997) have found no substantial detection of penetrative convection beneath the solar convection zone, placing upper limits of $0.05H_p$, using parameterized models for overshoot. However, several groups have contested this result, including Xiong and Deng (2001), who show that their model (implemented in Xiong, 1986) is consistent with helioseismic observations but penetrates $0.63H_p$, i.e., the convective motions affect the thermal stratification $0.63H_p$ into the radiative region. A few recent works by Baturin and Mironova (2010) and Christensen-Dalsgaard *et al.* (2011) modeled overshoot by fixing the gradients at the convective-radiative interface. Christensen-Dalsgaard *et al.* (2011) showed that the temperature profile of the Sun is consistent with a penetration depth of $0.37H_p$; however, this work assumes the resulting temperature gradients for overshoot without a self-consistent model of overshoot mixing. Taken together, there is a great deal of uncertainty in the nature and extent of convective overshoot (or penetrative convection) beneath the convection zone in the Sun.

The major goals of the doctoral work covered in this thesis are to investigate the impacts of multi-dimensional simulations of minor mixing processes on massive star evolution. Two recent studies on the supernovae and nucleosynthesis of entire massive star populations are summarized to provide background on the field in Chapter 2. Chapter 3 discusses recent models for semi-convection and overshooting convection based on numerical simulations by Wood *et al.* (2013) and Brummell *et al.* (2002), respectively, which were implemented into the stellar evolution code KEPLER. In Chapter 4, we discuss the potential effects of minor mixing processes on the structure of massive stars, developing an analytical understanding of how the internal structure determines the envelope characteristiccs. The next logical step of this work is to apply these new models to a new grid of stellar models followed by another application of the work from Chapter 2. In addition, multi-dimensional simulations with physics more focused on the physical processes relevant to massive stars would be beneficial. Original multi-dimensional work is presented in Chapter 5, which focuses on numerical simulations of fingering convection. Future work would entail implementation of this new model of fingering convection to massive stars. With the additional expertise gained by this study, additional simulations of overshooting convection and semi-convection relevant to astrophysics are possible in the future. Chapter 6 concludes with some ideas of future prospects in these subjects.

Chapter 2

Constraints on Supernova Physics and Stellar Evolution through Integrated Population Synthesis

The contents of this chapter come from Sections 1, 2, 3, and 4 of Brown and Woosley (2013) and Sections 5 and 6 of Sukhbold *et al.* (2016). Example work is presented on the integrated properties of many stellar models for a given set of prescriptions for minor mixing processes (the default KEPLER prescriptions from Weaver *et al.*, 1978, with $\alpha = 0.1$) to provide context for how theoretical massive star evolution calculations can compare to observations and depend on mixing.

2.1 Introduction

Just which massive stars explode as supernovae and which collapse to black holes has been a topic of great interest for a long time. In particular, stellar nucleosynthesis can be used to constrain the maximum mass of the supernova that needs to (or can) explode in order to explain the abundances of the elements seen in the Sun and elsewhere (e.g., Twarog and Wheeler, 1982, 1987; Maeder, 1992). Previous works have generally focused on the production of helium and oxygen and the ratio $\Delta Y/\Delta Z$ where Y is the helium mass fraction and Z is the heavy element fraction. Maeder (1992), for example, concludes that the observed abundances are best fit if stars of all masses contribute their pre-explosive winds, but only stars below 20–25 $M_\odot$ explode as supernovae (though see Prantzos (1994) who contested this conclusion). The remainder presumably end up as black holes which accrete the remaining star, including all its heavy elements.

Added interest in this issue has been generated recently by the increasingly tight observational constraints placed upon the masses of presupernova stars. Smartt (2009, 2015) found no evidence for supernova progenitors with masses over 20 $M_\odot$. Horiuchi *et al.* (2011) also finds an inconsistency with the measured rate of core collapse supernovae and the cosmic star formation rate in the sense that more stars seem to form than are observed to die as supernovae by a factor of about 2. Of course, these arguments are not yet absolute. More massive supernovae may be hidden in dust and the connection between main sequence mass and presupernova mass relies on theory. Star formation rates and supernova rates are not precisely known, but these constraints may become tighter with time and certainly suggest that not all massive stars make luminous supernovae.

On the theoretical front, it has been known for a long time that more massive stars are harder to blow up than lower mass ones (e.g., Fryer, 1999; Fryer and Heger, 2000). With higher mass, the entropies around the collapsing iron cores are greater, and the fall off of density with radius is consequently more gradual. During the collapse this means a greater accretion rate on the proto-neutron star that is harder to reverse. O'Connor and Ott (2011) recently quantified this effect with a "compactness parameter" (see their Fig. 9) and measured the difficulty of blowing up stars in a 1D code as a function of that parameter. Their conclusion, interestingly again was that stars heavier than 20 $M_\odot$ were harder to explode.

These considerations motivate a revisit of the problem of stellar nucleosynthesis as a function of mass. Using a fiducial model set of solar metallicity models from Woosley and Heger (2007), the same model set used by O'Connor and Ott (2011), we study the nucleosynthesis and remnant masses resulting if the supernova explosions are truncated above a certain mass, $M_{\rm BH}$. A Salpeter initial mass function is assumed (Salpeter, 1955) with a slope of $\Gamma = -1.35$. We determine not only the bulk nucleosynthetic properties like oxygen and remnant mass as a function of $M_{\rm BH}$, but also examine the synthesis of the s-process, the individual ratios of important intermediate mass and light elements, and the synthesis of interstellar radioactivities,

^{26}Al and ^{60}Fe.

2.2 Integrating Yields over the Initial Mass Function

The yield tables of Woosley and Heger (2007) give the nucleosynthesis of all species from hydrogen through lead for supernovae resulting from non-rotating massive stars with solar metallicity for the following initial masses: 12–33 (every integer mass), 35–60 (every five masses), 60–80 (every 10 masses), 100, and 120 solar masses. The authors calculated explosions for four sets of models parametrized by the mass cut and explosion energy. Here we use their standard set for which the explosion energy was 1.2×10^{51} erg and the mass cut was located at the entropy jump where $S/N_A k$ = 4.0. These are their "A" models and are the same models for which Zhang *et al.* (2008) calculated compact remnant masses and O'Connor and Ott (2011) studied compactness. Nucleosynthesis ejected by the pre-explosive winds and in the explosions was archived separately, but is not available in Woosley and Heger (2007) (see Figure 4 in their paper for the mass loss prescription used). Using this grid of nucleosynthetic yields, we constructed a stellar population using the Salpeter initial mass function,

$$\Phi(M)\,\mathrm{d}(M) \propto M^{-2.35}\,\mathrm{d}(M). \tag{2.1}$$

The total yields of the stellar population were calculated by integrating the yields over the mass function, as described in Equation 2.2, where m_i is the total production (in solar masses) of isotope i, and $E_i(M)$ and $W_i(M)$ are the total ejecta of isotope i from, respectively, the supernova explosion and the winds in solar masses from a star with initial mass M:

$$m_i = \int_{12M_\odot}^{M_{\rm BH}} \Phi(M)E_i(M)\,\mathrm{d}\,M + \int_{12M_\odot}^{120M_\odot} \Phi(M)W_i(M)\,\mathrm{d}\,M. \tag{2.2}$$

We calculated the yields for the stable isotopes, varying the upper mass limit, $M_{\rm BH}$, to the values listed in Table 2.1. Note that the results presented in Sections 2.6 and 2.7 are calculated using a slightly different method (refer to Section 2.5).

Our results are expressed in terms of a production factor for each isotope defined as

$$P_i = \frac{m_i/\sum_i m_i}{S_i}, \tag{2.3}$$

where S_i is the mass fraction of the isotope in the Sun (Lodders, 2010). The production factor relates to the supernova rate needed to produce that isotope in sufficient abundance. It can be shown that the production factor from high-mass progenitors is inversely proportional to the number of supernovae estimated by the model, assuming that the majority of the isotope's production comes from this stellar population.

For comparison, we plot the production factors as a function of atomic mass for a population with no upper bound in Figure 2.2 (see also Woosley and Heger, 2007).

Table 2.1: Production Factor Statistics: M_{BH} is the heaviest supernova mass, ^{16}O is the production factor of the isotope, Mean (16–59) and σ (16–59) are the arithmetic mean and standard deviation of the production factors for isotopes with atomic mass between 16 and 59, Mean (60–84) and σ (60–84) are the corresponding values for isotopes with atomic mass between 60 and 84, and N_{SN} is the relative number of supernovae needed to produce the same amount of ^{16}O.

M_{BH} ($M_{\odot}$)	^{16}O	Mean (16–59)	σ (16–59)	Mean (60–84)	σ (60–84)	N_{SN}
120	14.7	6.22	5.48	14.6	6.72	1.00
100	14.6	6.19	5.46	14.5	6.69	1.01
80	14.5	6.14	5.43	14.5	6.67	1.03
70	14.2	6.05	5.37	14.3	6.62	1.04
60	14.0	5.96	5.31	14.2	6.55	1.06
50	13.1	5.60	4.79	13.2	6.08	1.13
40	11.5	5.19	4.37	10.6	4.68	1.30
30	7.93	4.24	3.52	5.86	2.37	1.88
28	7.33	3.97	3.26	5.27	2.19	2.03
25	6.52	3.60	2.96	4.35	1.88	2.28
24	6.27	3.50	2.92	3.95	1.78	2.37
23	5.99	3.38	2.88	3.59	1.72	2.49
22	5.69	3.26	2.84	3.29	1.60	2.61
21	5.44	3.14	2.83	3.01	1.52	2.75
20	5.17	3.00	2.81	2.70	1.46	2.88
19	4.93	2.82	2.77	2.51	1.39	3.01
18	4.70	2.67	2.79	2.30	1.32	3.18
17	4.52	2.56	2.82	2.12	1.16	3.30
16	4.36	2.44	2.88	1.94	1.00	3.42
15	4.22	2.31	2.96	1.78	0.834	3.52
14	4.12	2.21	3.07	1.61	0.622	3.61

2.3 The Heaviest Supernova

Observations constrain the mass of the heaviest supernova. These include not only the nucleosynthetic pattern of the intermediate mass elements, but also the light component of the s-process, the frequency of supernovae, and the masses of compact remnants. The existence of a cutoff mass has interesting implications for ΔY/ΔZ and the synthesis of ^{26}Al and ^{60}Fe.

We analyzed the statistics of the isotopes lighter and heavier than the iron group nuclei separately because the nuclei above $A = 60$ are mostly produced by the s-process. Above $A = 90$ the synthesis of the s-process is generally attributed to AGB stars as are ^{12}C and ^{14}N. Most of the iron group is probably made in Type Ia supernovae. The means and standard deviations of the production factors for nuclei below and above $A = 60$ that are attributed to massive stars are tabulated separately in Table 2.1.

2.3.1 The Production of Carbon, Oxygen, and Intermediate Mass Elements

When studying the nucleosynthesis of massive stars, theorists often normalize to the production of ^{16}O, as it is the third most abundant isotope in the universe and comes almost entirely from massive stars (Langer, 1996). For reference, the production factor of ^{16}O is included in Table 2.1.

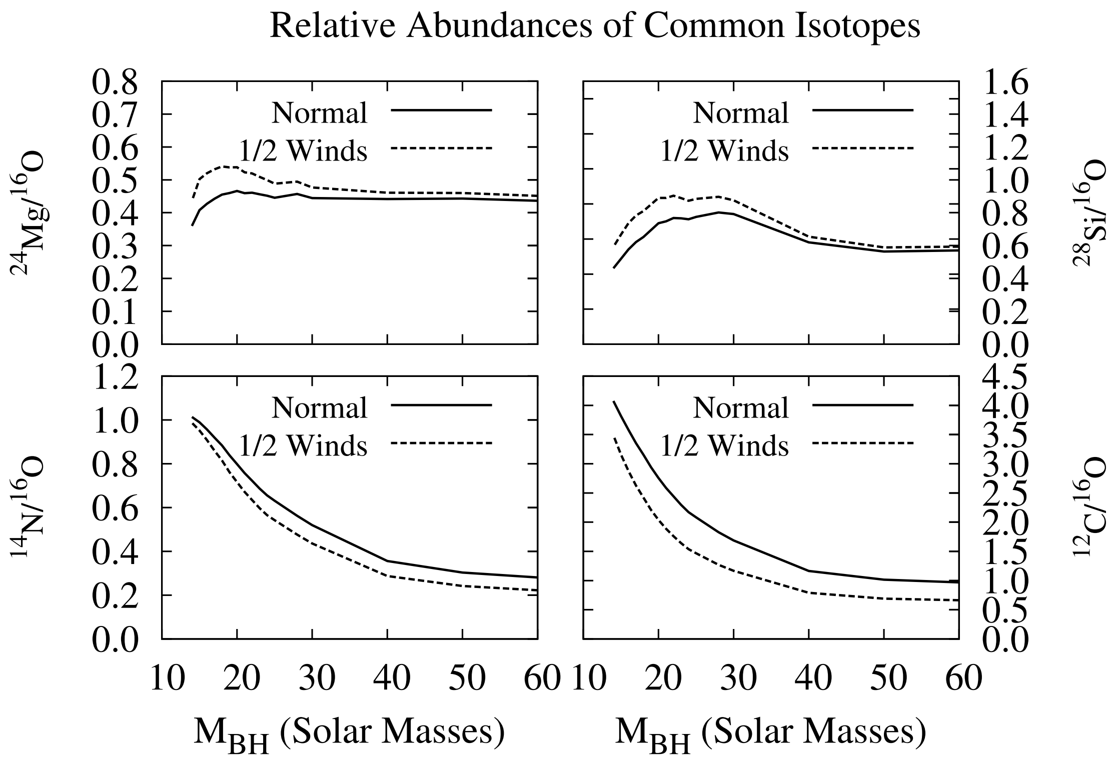


Figure 2.1: The relative abundances of the common isotopes with respect to ^{16}O as a function of $M_{\rm BH}$. The plot has been truncated at $M_{\rm BH} = 60 M_\odot$, beyond which the abundances do not illustrate any significant variations. A relative abundance of 1 indicates solar ratio. The dotted lines represent the ratios produced if the wind contribution is halved to illustrate the dependence on wind. The production factors of these isotopes are consistent with their solar abudances if the upper mass limit is reduced from 120 to 40 $M_\odot$. This supports the preliminary limits set by Heger *et al.* (2003).

In Figure 2.1, we compare the production factor of various elements to that of ^{16}O for variable M_{BH}. Based on the relevant uncertainties in the system, an acceptable nucleosynthetic yield is within a factor of two of the solar abundance. The abundances of the alpha elements probe the strength of oxygen and silicon burning, and that of ^{14}N to measures the strength of the CNO cycle. Figure 2.1 shows that the production factors of these common isotopes are consistent with their solar abudances if the upper mass limit is reduced from 120 to 40 $M_{\odot}$. This supports the preliminary limits set by Heger *et al.* (2003).

However, for reduced upper mass limits, ^{12}C is overproduced in the winds compared to ^{16}O, suggesting either that this bound cannot be much lower than 25 $M_{\odot}$ or that mass loss is significantly less effective than assumed. For the mass range considered, carbon is mostly made in the winds of very massive stars, especially during their Wolf-Rayet stage. In addition, lower mass stars produce substantial carbon, so carbon must be underproduced in these calculations. The final nucleosynthetic yields also depend upon the rate of the $^{12}\mathrm{C}(\alpha,\gamma)^{16}\mathrm{O}$ reaction. A value 1.2 times that of Buchmann (1996) is employed, corresponding to an S-factor at 300 keV of 175 keV, which is well within the uncertainty (Schürmann *et al.*, 2012). Oxygen, which is the major part of the metallicity, is also expelled in winds but more is ejected during the explosion. The ratio of ^{12}C to ^{16}O then is quite sensitive to the mass loss prescription used. To illustrate this, in Figure 2.1, we've also included the ratios for a population for which we've halved the mass loss yields. It should be noted that such a method

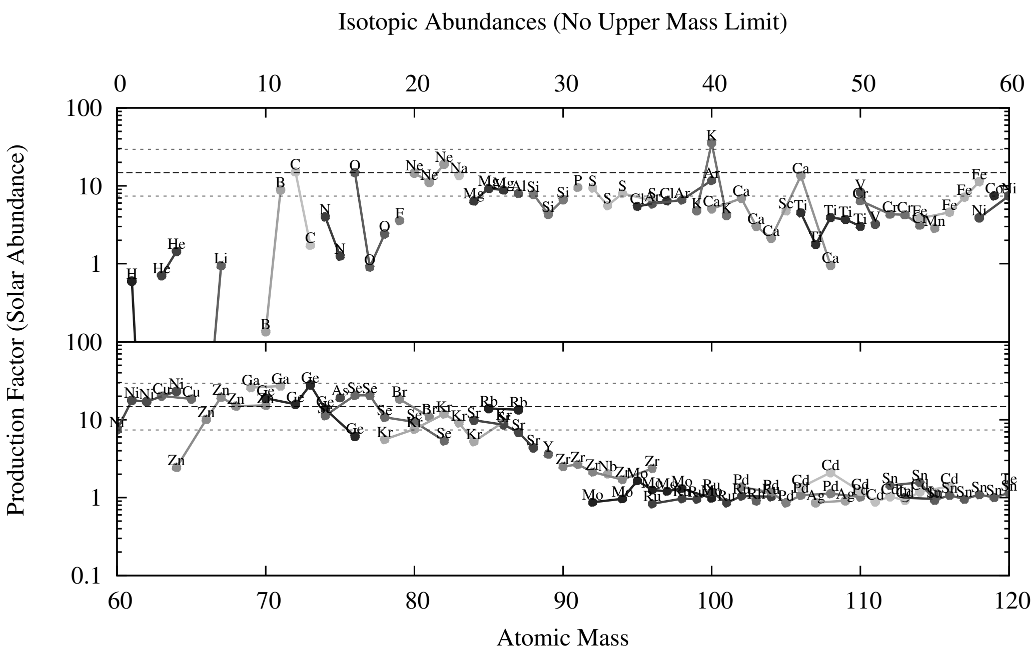


Figure 2.2: The production factor (plotted logarithmically) as a function of atomic mass when no truncation of supernovae is assumed, i.e., when stars as heavy as 120 $M_{\odot}$ continue to explode. The dashed line represents the production factor of ^{16}O. The shorter dashed lines represent a factor of two deviation from ^{16}O. The light s-process isotopes (A from 60 to 90) are overproduced with respect to the intermediate mass elements, which is needed to account for lower mass stars that do not produce much in this range. The apparent large overproduction of ^{40}K in both figures is not problematic as it has a short half life. Any such potassium produced and incorporated into the Sun would have sufficient time to decay.

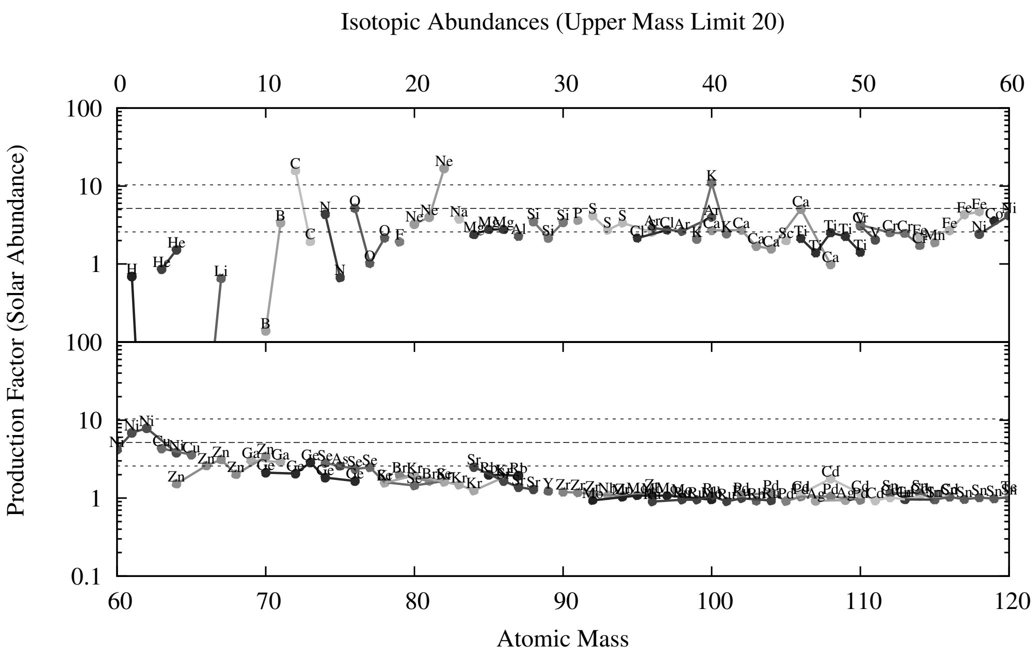


Figure 2.3: The structure of this figure is identical to Figure 2.2. Here it is assumed that all massive stars over 20 $M_\odot$ end up collapsing as black holes and their only nucleosynthetic contribution is from their winds. The large abundance of ^{40}K is not a problem as this species will decay prior to solar system formation, but the large abundances of ^{22}Ne and ^{12}C could indicate unresolved issues with mass loss, nuclear physics, or the assumption of a low threshold for black hole formation. The *s*-process is also weak here (compare with Figure 2.2).

is not internally consistent as the stellar evolution calculations and explosions would need to be recalculated. However, this can provide a first-order estimate as to the results of such calculations. The C/O ratio remains consistent with the solar abundances down to an upper mass limit of about 20 $M_{\odot}$. To illustrate the change in the nucleosynthesis between the upper mass limit extrema, we present Figure 2.3 as a comparison to Figure 2.2. The apparent large overproduction of ^{40}K in both figures is not problematic as it has a short half-life. Any such potassium produced and incorporated into the Sun would have sufficient time to decay. The large yield of ^{22}Ne remains a concern; however, this yield can be reduced if the mass loss is halved.

We checked other intermediate mass isotope ratios to more comprehensively investigate how these stellar populations contribute to solar abundances. These include ^{28}Si/^{40}Ca, ^{40}Ca/^{24}Mg, ^{28}Si/^{56}Fe, and ^{40}Ca/^{56}Fe. As can be seen in Figure 2.4, these all remain acceptably close to solar ratios for the entire range of M_{BH} and, perhaps surprisingly are not tight constraints on $M_{\rm BH}$.

2.3.2 The Supernova Rate

Using the production factor of ^{16}O, we estimated how the supernova rate depends upon $M_{\rm BH}$. An accurate calculation of the rate itself requires a galactic chemical evolution model that includes gas accretion and galactic outflows—as described in Timmes *et al.* (1995)—and is beyond the scope of this work. The supernova rate is estimated by examining N_{SN}, the total number of massive stars required to produce

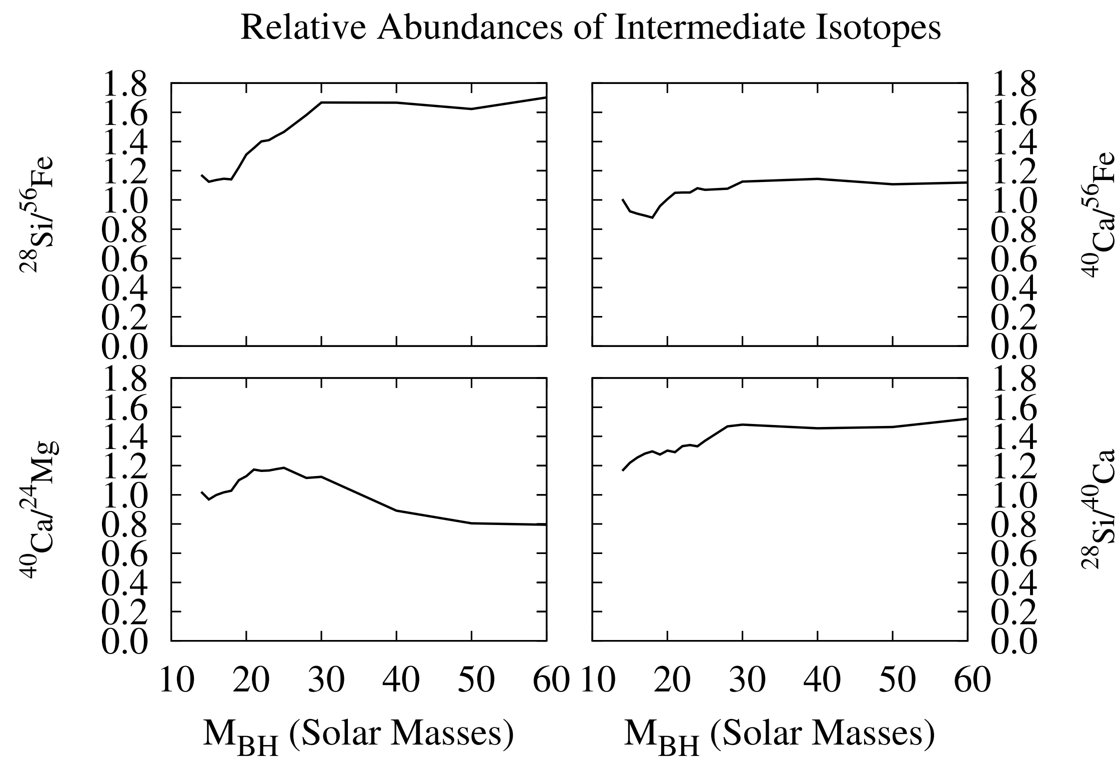


Figure 2.4: The relative abundances of the intermediate mass isotopes as a function of $M_{\rm BH}$. The plot has been truncated at $M_{\rm BH} = 60 M_{\odot}$, beyond which the abundances do not illustrate any significant variations. These all remain acceptably close to solar ratios for the entire range of M_{BH} and, perhaps surprisingly are not tight constraints on $M_{\rm BH}$.

the correct amount of ^{16}O, normalized to the total number required to produce the yields of our control population ($M_{BH} = 120\ M_\odot$). The number of required massive stars increases slowly with the lowering of the upper mass limit until approximately 40 $M_\odot$, below which it increases rapidly up to twice the original number at 28 $M_\odot$ and three times the original number at 19 $M_\odot$ (see Table 2.1).

2.3.3 The Light *s*-Process

The heavier component of the *s*-process, i.e.,those nuclei above $A \approx 90$, is thought to be produced in low mass stars (Pignatari *et al.*, 2010). The light component of the *s*-process, on the other hand, is generally attributed to massive stars and is produced late during helium burning when the temperature rises sufficiently for the ^{22}Ne(α,n)^{25}Mg reaction to occur. Owing to the temperature sensitivity of this rate, production of the *s*-process isotopes where A is in the interval 60–90 is a potentially powerful constraint on M_{BH}. The massive star population must overproduce these isotopes as many of the primary isotopes are also made in lower mass stars that don't produce many *s*-process elements.

Use of this diagnostic is complicated by the possible production of many isotopes in this range by the *r*-process. This analysis instead focused on a few *s*-only isotopes (Figure 2.5)—those isotopes produced exclusively by the *s*-process—^{70}Ge, ^{76}Se, ^{86}Sr, and ^{87}Sr (Kappeler *et al.*, 1989). We excluded ^{80}Kr and ^{82}Kr from our analysis—despite both being *s*-only isotopes—as these isotopes are also significantly produced

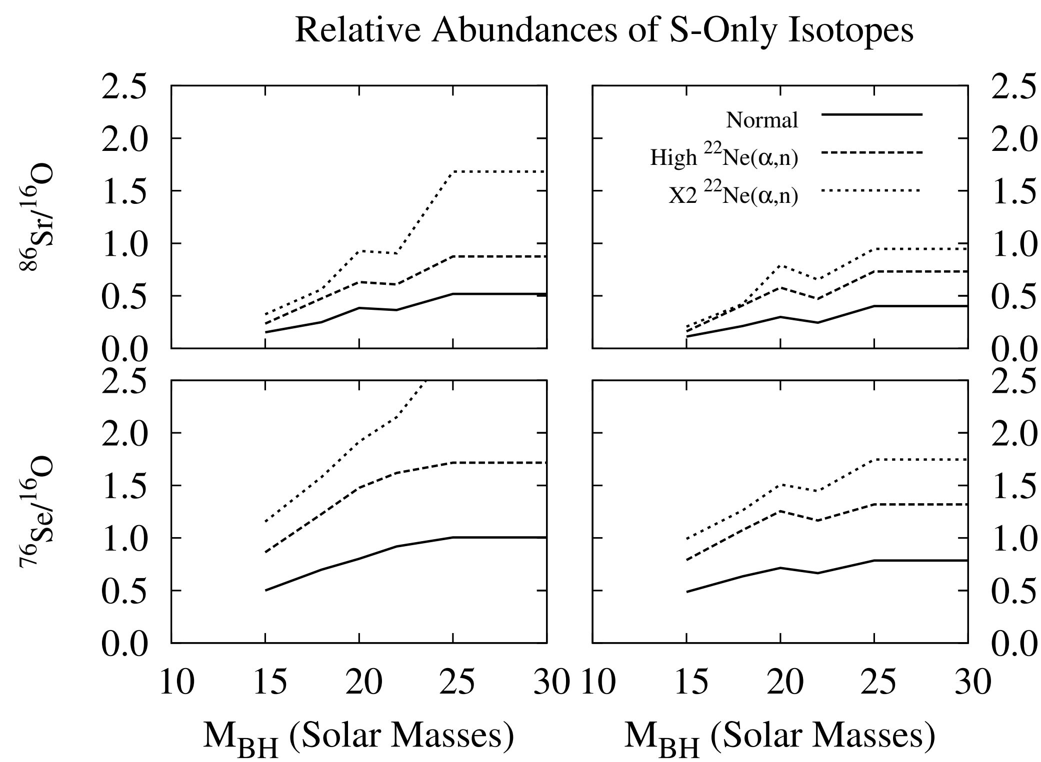


Figure 2.5: The relative abundances of the s-only isotopes with respect to ^{16}O and ^{24}Mg as a function of M_{upper}. This is to indicate the strength of the s-process in this population without worry of contamination of the r-process. The plot has been truncated at $M_{upper} = 60 M_{\odot}$, beyond which the abundances do not illustrate any significant variations. A relative abundance of 1 indicates solar ratio. We plot this for three values of $^{22}Ne(\alpha, n)$, as discussed in the text. Note that these isotopes of Sr are underproduced, which is expected as these isotopes are also made in asymptotic giant branch stars. The production of ^{70}Ge falls below 1/2 the solar abundance at 21 $M_{\odot}$, whereas the production of ^{87}Sr falls below 1/2 solar production at 23 $M_{\odot}$. However, if the rate of ^{22}Ne(α,n) is increased to the maximum experimental value according to Jaeger *et al.* (2001), 18 $M_{\odot}$ becomes a reasonable value for the mass of the heaviest supernovae.

in low-mass asymptotic giant branch stars. The s-process is most predominant in core-collapse supernovae from 30 to 50 $M_{\odot}$. With more complex models of the stellar population, a comparison of the total s-process yields to the solar abundance could prove to provide greater insight as to the upper mass bound. Here, ^{86}Sr and ^{87}Sr are underproduced regardless of the upper bound, which is expected, as these isotopes are also produced in asymptotic giant branch stars (Arlandini *et al.*, 1999). However, the other isotopes, ^{70}Ge and ^{76}Se, are produced primarily in massive stars. The production of ^{70}Ge falls below 1/2 the solar abundance at 21 $M_{\odot}$, whereas the production of ^{87}Sr falls below 1/2 solar production at 23 $M_{\odot}$. However, if the rate of ^{22}Ne(α,n) is increased to the maximum experimental value according to Jaeger *et al.* (2001), 18 $M_{\odot}$ becomes a reasonable value for the mass of the heaviest supernovae.

It is important to note here that this result depends heavily on the $^{22}\mathrm{Ne}(\alpha, n)^{25}\mathrm{Mg}$ rate, as can be seen in Figures 2.5 and 2.6. The values for this rate used in the present study are taken from Jaeger *et al.* (2001). For calibration, the normal value at 3×10^{8} K is 2.58×10^{-11} cm^3 s^{-1} and the high value is 3.14×10^{-11} cm^3 s^{-1}. More recent studies of this reaction rate by Longland *et al.* (2012) suggest that an even higher value may not be unreasonable, extending up to perhaps as much as 4.1×10^{-11} cm^3 s^{-1}. To explore this extreme possibility, we also carried out runs in which the Jaeger *et al.* (2001) normal rate was multiplied by two. These three cases are labeled in the figures as "Normal", "High", and "X2." In all cases, we used the low value from Jaeger *et al.* (2001) for the $^{22}\mathrm{Ne}(\alpha, \gamma)^{26}\mathrm{Mg}$ rate, which corresponds to 0.81×10^{-11}

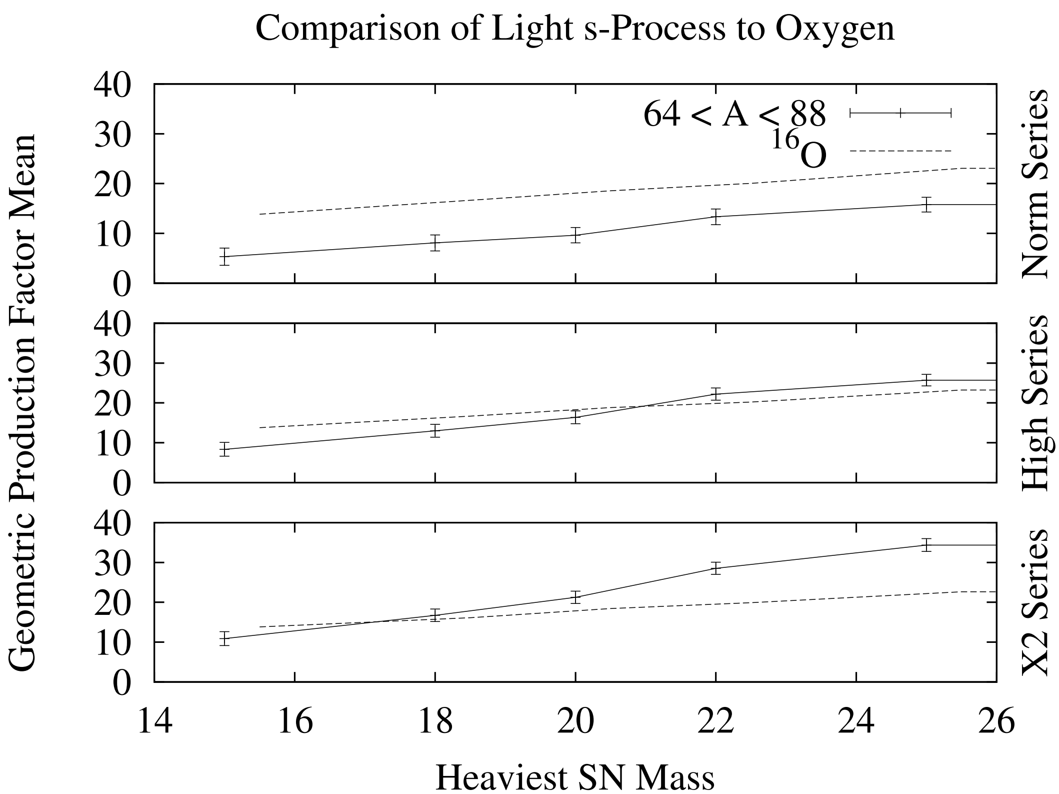


Figure 2.6: A comparison of the overproduction of the light s-process to the production factor of ^{16}O, which is a representation of the lighter isotopes. We plot this for three values of $^{22}\mathrm{Ne}(\alpha, n)$, as discussed in the text. We have plotted the geometric mean of the production factor for atomic masses between 65 and 87. As before, we require that the light s-process is overproduced with respect to the lighter isotopes. This plot is not intended to provide a limit on M_{BH} but rather to show that any such limit based on the average s-process abundance depends heavily on the $^{22}\mathrm{Ne}(\alpha, n)$ rate.

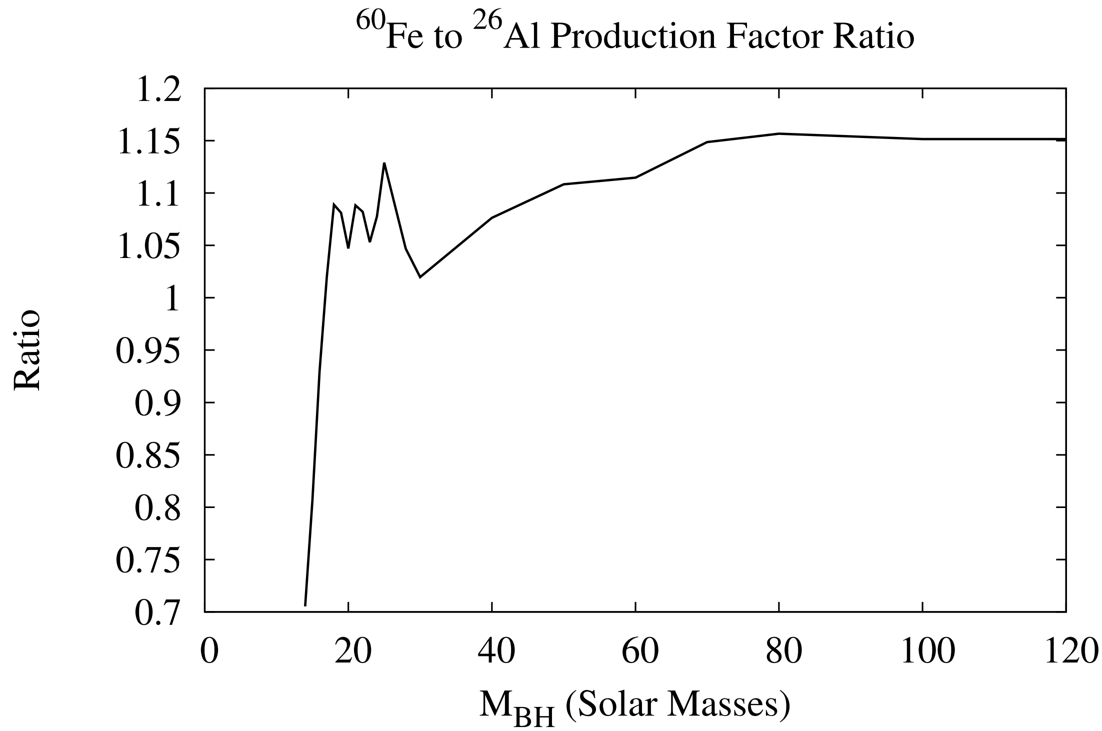


Figure 2.7: The ratio of the total masses of ^{60}Fe to ^{26}Al as a function of upper mass limit. The observational value of this quantity is expected to 0.34, which these calculations cannot reproduce for any upper mass limit. Lowering this limit improves this ratio marginally. The remaining difference between 1.15 and 0.34 chiefly reflects remaining uncertainty in the neutron capture cross sections for ^{59}Fe and ^{26}Al, the treatment of semiconvection, mass loss and rotation in the models, and the choice of explosion energies for low mass supernovae.

cm^3 s^{-1} at 3×10^8 K. Figures 2.5 and 2.6 show that both the production of the *s*-only elements and of the light *s*-process as a whole are improved by increasing this rate.

2.3.4 Iron-60 and Aluminum-26

The nuclei ^{60}Fe ($\tau_{1/2} = 2.62 \times 10^6$ y) and ^{26}Al ($\tau_{1/2} = 7.17 \times 10^5$ y) accumulate in the interstellar medium where the emission generated by their decays can be studied via gamma-ray observations. Observations by Wang *et al.* (2007) give a ratio for the decay rate of ^{60}Fe to that of ^{26}Al of about 0.148, implying a ratio of ^{60}Fe to ^{26}Al mass fractions of $(60/26)(0.148) = 0.34$ (Woosley and Heger, 2007). As pointed out by Prantzos (2004) and discussed in Woosley and Heger (2007), current calculations give a larger value. As Figure 2.7 shows, our mass-averaged estimate of the production is 1.15 for $M_{\rm BH} = 120$. This compares favorably with the value 0.95 given for these models by Woosley and Heger (2007).

The remaining difference between 1.0 and 0.34 chiefly reflects remaining uncertainty in the neutron capture cross sections for ^{59}Fe and ^{26}Al, the treatment of semi-convection, mass loss and rotation in the models, and the choice of explosion energies for low mass supernovae. Kasen and Woosley (2009) suggest that many supernovae, perhaps most, have an explosion energy smaller than the fiducial 1.2×10^{51} erg assumed in many studies. Reducing the explosion energy reduces the yield of ^{60}Fe substantially. Limongi and Chieffi (2006) have also emphasized the dependence of this production ratio on mass loss, convection theory, and the initial mass function. Including the effects of rotation may also increase the ^{26}Al yield (Palacios *et al.*, 2005), especially in very massive stars.

This work does not resolve the debate surrounding ^{60}Fe/^{26}Al, but the ratio depends substantially upon $M_{\rm BH}$. Measurements of gamma-ray line flux ratios may thus ultimately help constrain the masses of stars that explode. Figure 2.7 shows that the ratio of ^{60}Fe/^{26}Al produced by supernovae declines as $M_{\rm BH}$ decreases. Lower kinetic

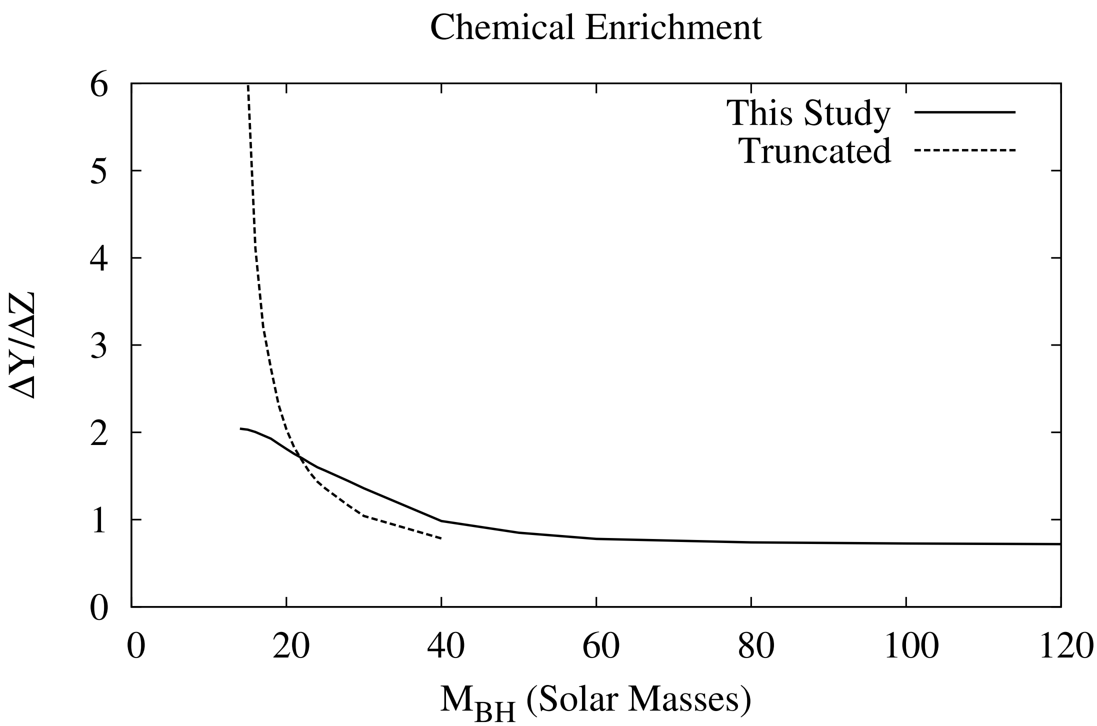


Figure 2.8: The ratio of the total deviations from the initial helium and metal mass fractions of the population as a function of the upper mass limit. Observations suggest a value $\Delta Y/\Delta Z \sim 4$ (Pagel *et al.*, 1992). The truncated curve uses the same assumptions as Timmes *et al.* (1995) in order to recreate the solar metallicity curve of Figure 37 of that same paper. Until $M_{\rm BH}$ is reduced below $\sim$40 $M_{\odot}$, $\Delta Y/\Delta Z$ remains slightly less than unity. Even for $M_{\rm BH} = 18$ $M_{\odot}$, $\Delta Y/\Delta Z$ only reaches 1.9, well below the observed value.

energies of stellar explosions could further reduce the production of ^{60}Fe.

2.3.5 Helium Production

As discussed extensively in Maeder (1992), the measured derivative of the helium mass fraction with respect to metallicity can, in principle, be used to constrain $M_{\rm BH}$. The winds of the most massive stars are rich in helium while the heavy elements are largely confined to the cores. If the cores collapse to black holes, trapping the heavy elements, the synthesis of the winds remains and increases the overall average of $\Delta Y/\Delta Z$. Observations suggest a value $\Delta Y/\Delta Z \sim 4$ (Pagel *et al.*, 1992).

In practice, the application of this metric is fraught with uncertainty. The yields of the massive stars depend on the initial mass function and especially the uncertain mass loss rates employed. After the helium core is uncovered by mass loss, Wolf-Rayet mass loss contributes not only helium, but an increasing amount of carbon and oxygen. The remnant masses additionally depend on an uncertain explosion mechanism. Lower mass stars also contribute appreciably to both helium and metallicity and the yields of stars of all masses are sensitive to metallicity. Even an approximate result requires integration over some uncertain model for galactic chemical evolution.

However, because this metric has been applied extensively in the literature, we give in Figure 2.8 our results for the solar metallicity stars considered in this survey. Until $M_{\rm BH}$ is reduced below $\sim$40 $M_\odot$, $\Delta Y/\Delta Z$ remains slightly less than unity. Even for $M_{\rm BH} = 18\ M_\odot$, $\Delta Y/\Delta Z$ only reaches 1.9, well below the observed value.

These results seem inconsistent with those of Timmes *et al.* (1995) who found $\Delta Y/\Delta Z = 4$ for $M_{\rm BH} = 17\ M_\odot$ for solar metallicity stars and 30 $M_\odot$ for low metallicity stars. However, Timmes et al. used the survey of Woosley and Weaver (1995) which included only stars below 40 $M_\odot$, while our grid extends to 120 $M_\odot$. Furthermore, Timmes *et al.* (1995) assumed that any winds would not contain metals and would not strip the helium core. Using these approximations, we have replicated their results as the dashed curve in Figure 2.8 results. This curve crosses ΔY/ΔZ at $M_{\rm BH}$ = 17 $M_\odot$. However, our results for solar metallicity are more realistic, and a lower value of ΔY/ΔZ from massive stars is appropriate.

2.3.6 Average Mass of Compact Remnants

Zhang *et al.* (2008) calculated fall back and black hole production in the same supernovae employed in the present study, and one can use their results to estimate the properties of the compact remnants left by solar metallicity stars as a function of $M_{\rm BH}$. Their Table 4 gives the baryonic masses of the remnants for our fiducial set for stars, the "sA Models". Results are given for initial masses from 12 to 100 $M_\odot$. To these we add the remnant masses specified in their text: 1.37 $M_\odot$ for the remnant of an 11 $M_\odot$ star and 1.35 $M_\odot$ for a 10 $M_\odot$ star. The gravitational mass for the resulting neutron stars, where one forms, can be calculated using their Eq. (2). It is assumed that the maximum gravitational mass of a neutron star is 2.0 $M_\odot$ so all remnants with a baryonic mass over 2.35 $M_\odot$ form black holes with a gravitational mass equal to its initial baryonic mass. Once again assuming a Salpeter initial mass function, the average properties of the remnants can be determined as a function of, $M_{\rm BH}$, the heaviest supernova mass.

Figure 2.9 shows the resulting average black hole mass, neutron star mass, iron core mass, and the maximum remnant mass as a function of $M_{\rm BH}$. For those stars more massive than the heaviest supernova, we have assumed that the presupernova

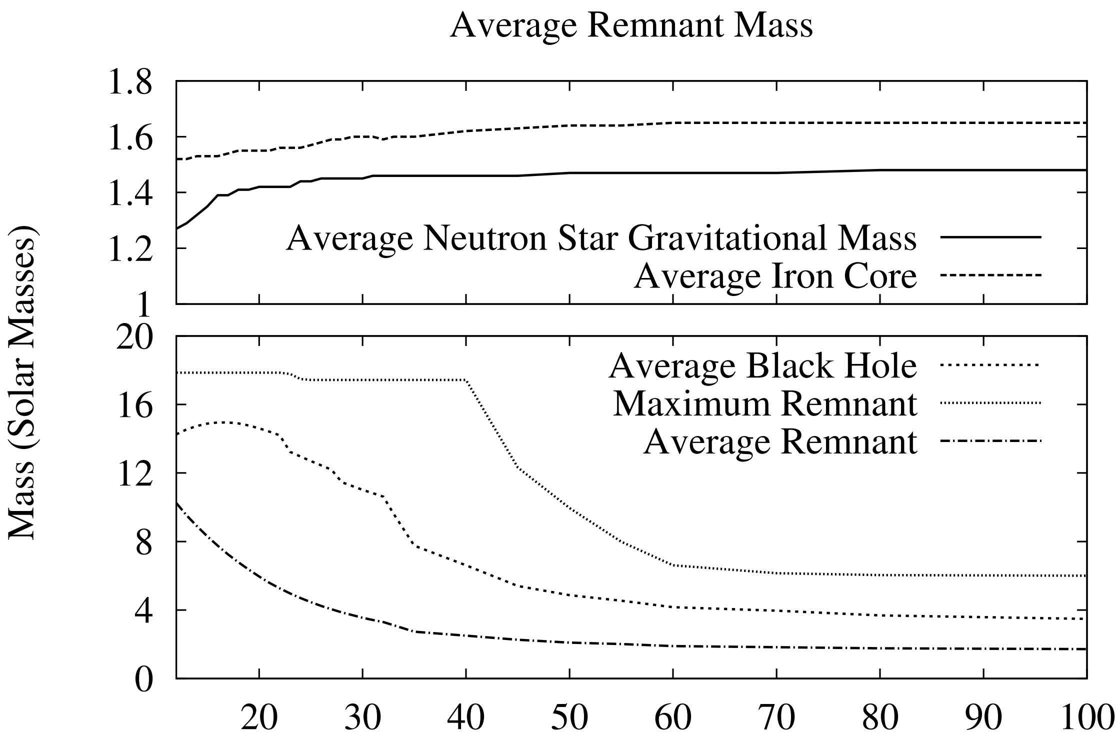


Figure 2.9: The average neutron star remnant mass, iron core mass, average black hole mass, maximum remnant mass, and average remnant mass are plotted as a function of M_{BH}. The average neutron star mass is within the observational uncertainty everywhere below $M_{\mathrm{BH}} = 25M_{\odot}$. The results can be compared with the observed average neutron star mass from Schwab *et al.* (2010). From 14 neutron stars with well-measured masses, they found an average neutron star mass of 1.325 ± 0.056 $M_{\odot}$. The closest average neutron mass to this value is achieved when $M_{\mathrm{BH}} = 14M_{\odot}$, and the most extreme value consistent with the uncertainties from this calculation is $M_{\mathrm{BH}} = 16M_{\odot}$. This suggests that a low value for M_{BH} is more consistent with the observed neutron star masses.

star collapses completely to a black hole, so our black hole masses are larger than those of Zhang *et al.* (2008). The results can be compared with the observed average neutron star mass from Schwab *et al.* (2010). From 14 neutron stars with well-measured masses, they found an average neutron star mass of $1.325 \pm 0.056\ M_\odot$. The closest average neutron mass to this value is achieved when $M_{\rm BH} = 14 M_\odot$, and the most extreme value consistent with the uncertainties from this calculation is $M_{\rm BH} = 16 M_\odot$. This suggests that a low value for $M_{\rm BH}$ is more consistent with the observed neutron star masses. However, there are appreciable errors in both the observational limits and the remnant masses calculated by Zhang *et al.* (2008). Any of the neutron star masses plotted in Figure 2.9 is not badly discrepant with observations.

The maximum and average black hole masses are a more sensitive indicator of $M_{\rm BH}$. Substantially larger values result if $M_{\rm BH}$ is low. Interpretation of these results is complicated though the uncertainties associated with both mass loss and the explosion mechanism. Lower metallicity stars would presumably experience less mass loss, potentially producing more massive black holes.

2.4 Heaviest Supernovae Conclusions

Through a variety of nucleosynthetic diagnostics, the mass of the maximum mass supernova can be estimated, given the assumption that all stars heavier than a critical mass promptly become black holes. At the outset, a number of caveats are worth stating. First, our yield set from Woosley and Heger (2007) reflects a particular choice of many uncertain parameters—the treatment of convection, mass loss, key reaction rates (like $^{12}\mathrm{C}(\alpha, \gamma)^{16}\mathrm{O}$), and explosion physics. We lump together both means of forming a black hole, direct and by fall back, and we ignore mixing of the supernova ejecta. It is probable that some supernovae make black holes and yet still eject some fraction of their presupernova nucleosynthesis. Mixing during the fallback epoch can complicate the estimate of yields. This probably affects the intermediate-mass elements and iron more than helium, carbon, oxygen and the *s*-process. Our models sample a single metallicity—solar. While other models of different metallicity are available, using them would require a more careful treatment of galactic chemical evolution than is presently justified, given all the other uncertainties. Metallicity will not greatly affect the supernova yield of primary elements like oxygen and the intermediate mass elements, but it does affect the mass loss. We attempted to partially compensate for that by multiplying the integrated nucleosynthesis of the wind by a constant factor of 0.5, as shown in Figure 2.1. The use of solar metallicity also overestimates somewhat the production of secondary species such as the elements with odd nuclear charge and the *s*-process (e.g., Timmes *et al.*, 1995), so somewhat larger yields than solar might be required than suggested by Figure 2.3. Finally, all our models ignore rotation. Not only can rotation affect the mechanism and symmetry of the explosion, but it also makes the mass of the helium and carbon-oxygen core larger for a given main sequence. So our derived mass limits may actually need to be smaller.

Given these limitations, the best we can say at the present time is what supernova mass limits might be consistent with observations. The idea of a limiting mass is itself an approximation, since the compactness of the core is not a monotonic function of main sequence mass (O'Connor and Ott, 2011), especially in the interesting range 20 – 35 $M_\odot$. For still heavier stars, mass loss may shrink the helium core so much that the presupernova helium core mass of say a 100 $M_\odot$ star differs little from that of a 20 $M_\odot$ star. Such massive stars are rare however, and their nucleosynthesis is mostly due to their presupernova winds.

We have looked at several processes that limit $M_{\rm BH}$. As $M_{\rm BH}$ is decreased, the necessary rate of successful supernovae rises. For $M_{\rm BH}$ = 20 the rate is 2.88 times greater than for $M_{\rm BH}$ = 120. Surprisingly, in reducing $M_{\rm BH}$ to 20 $M_\odot$, the overall nucleosynthesis of most isotopes from Ne to Ca with respect to ^{16}O is altered little and the production of some isotopes in the iron group is actually improved. The greatest apparent problem is ^{22}Ne; however, this can be mitigated somewhat by reducing the amount of mass loss or by increasing the ^{22}Ne(α,n) rate. Of the *s*-process only isotopes produced mainly in massive stars, ^{70}Ge becomes the limiting isotope for the mass of the heaviest supernova, reaching half of solar value at 23 $M_\odot$. This can be extended to 18 $M_\odot$ by increasing the ^{22}Ne(α,n) rate.

In total, we find that we can easily reduce the mass of the heaviest supernovae to 40 $M_\odot$ without any significant changes. For KEPLER's standard values of nuclear rates and mass loss, there are only moderate changes in these processes down to 25 $M_\odot$. The limits become increasingly severe for smaller masses. For $M_{\rm BH} = 25\ M_\odot$, the stellar winds overproduce ^{12}C with respect to ^{16}O by a factor of two, unless we reduce the mass loss in these stars by two. At $M_{\rm BH} = 21\ M_\odot$, the lighter elements are overproduced in these more massive stars, and the *s*-process produces only half the needed ^{70}Ge unless the ^{22}Ne(α,n)^{25}Mg rate is increased. At 20 $M_\odot$, even with halved mass loss, the winds overproduce ^{12}C by a factor of two. At 19 $M_\odot$ the supernova rate is three times what it would have been had all stars contributed their full nucleosynthesis.

2.5 Integrating Yields, Calibrated to SN 1987A

In what follows, a new series of yields from from Sukhbold *et al.* (2016) were integrated in a similar method as that from Section 2.2. The properties of the piston used to explode the star in Sukhbold *et al.* (2016) were given by a consistent engine, calibrated to SN 1987A for stars of zero age main-sequence mass greater than $13 M_\odot$ and the Crab nebula progenitor for those less than $13 M_\odot$. Multiple engines were evaluated, based on assumptions of the progenitor structure for SN 1987A: the two investigated in the most detail in that study were an 18 $M_\odot$ model from Woosley and Heger (2007) (W18) and a 20 $M_\odot$ model from Nomoto and Hashimoto (1988) (N20). Two other calibrations used a 15 $M_\odot$ model (W15) and a 20 $M_\odot$ model (W20). For these calculations, rather than evaluating the results for a changing upper mass limit, we take the explosive yields from those stars that explode based either on the W18 or N20 engines from Sukhbold *et al.* (2016). Again, we include the wind yields from all stellar models.

2.6 Compact Remnant Distributions, Calibrated to SN 1987A

2.6.1 Neutron Stars

The baryonic masses from the remnants are calculated in Sukhbold and Woosley (2014). Lattimer and Prakash (2001) provide a formula to calculate the neutron star's gravitational mass, M_g, from its calculated baryonic mass, M_b and an assumed neutron star radius of 12 kilometers:

$$\beta = \frac{GM_g}{c^2 12\mathrm{km}}\,, \tag{2.4}$$

$$\frac{M_b - M_g}{M_g} = 0.6\frac{\beta}{1-0.5\beta}\,, \tag{2.5}$$

where G is the gravitational constant and c the speed of light.

A distribution of gravitational masses can be constructed by weighting the occurrence of each of the models that successfully explode according to a Salpeter IMF. The resulting frequencies are plotted as a function of neutron star mass in Figure 2.10 and are seen to be in reasonably good agreement with the observed values of

Lattimer (2012)[1] in terms of overall spread and mean value. Our maximum gravitational neutron star mass for the W18 series is 1.68 $M_\odot$ and the minimum is 1.23 $M_\odot$. Use of 10 km for the neutron star radius reduces these numbers to 1.64 $M_\odot$ and 1.21 $M_\odot$, and reduces the average mass by about 0.02 $M_\odot$.

Different mass neutron stars, for the most part, come from different ranges of main sequence mass with lower-mass progenitors producing low-mass neutron stars. There is some overlap however. Remnants from stars between 15 and 18 solar masses have slightly lower masses than some below 15 solar masses. Stars above 18 $M_\odot$ contribute little to the distribution, but do so over a broad range of neutron star masses. This reflects both the variable mass of the very massive stars at death, due to extensive mass loss, and inherent variability in their compactness. While we have not carried out a statistical analysis, Figure 2.10 does suggest some bimodality in the distribution, with a separate peak around 1.25 $M_\odot$ (see also Schwab *et al.*, 2010). It is interesting to note that these low mass neutron stars are produced without invoking electron-capture supernovae (e.g., Schwab *et al.*, 2010), none of which are in our sample. Indeed, heavier stars that experience iron-core collapse supernovae are capable, in principle, of producing lighter neutron stars than electron-capture supernovae since the values of Y_e in the cores of the latter are close to 0.5 at the onset of collapse, while they are substantially less after oxygen and silicon burning in the former. In quite heavy stars, this destabilizing effect is more than compensated for

[1] http://stellarcollapse.org/nsmasses; retrieved 6 Sep. 2015

by the high entropy and mild degeneracy, but in lighter 9 - 11 $M_\odot$ stars, Y_e becomes a major determinant for the onset of core collapse.

An average neutron star mass can also be generated from a IMF-weighted sampling of the successful explosions. The results for four different calibrations of the central engine for SN 1987A, each including the normalization to the Crab for stars below 13 $M_\odot$, are given for a Salpeter IMF in Table 2.2. These values are not very sensitive to the central engine used for SN 1987A, but are determined more by presupernova properties. All lie close to the (error-weighted) observational means given by Lattimer (2012)[1] of 1.368 $M_\odot$ for x-ray and optical binaries, 1.402 $M_\odot$ for neutron star binaries, and 1.369 $M_\odot$ for neutron star–white dwarf binaries. They are not far from the average neutron star mass found by Schwab *et al.* (2010), $1.325 \pm 0.056 M_\odot$. These new theoretical values are closer to the observations than the results from Brown and Woosley (2013).

Table 2.2 also gives the average baryonic mass of the remnants of successful explosions, $\bar{M}_b$, the average supernova kinetic energy at infinity, $\bar{E}$, and the range of average ^{56}Ni masses, all as calculated in P-HOTB. The lower bound ignores the contribution from the neutrino-powered wind while the upper bound assumes all of the wind not in the form of α-particles is ^{56}Ni. Based on post-processing with KEPLER this may be an overestimate of the actual ^{56}Ni production by 10 to 20%. The column labelled SN% is the IMF-weighted fraction of all stars studied that exploded and left neutron star remnants. The last columns give the numerical fraction of all successful

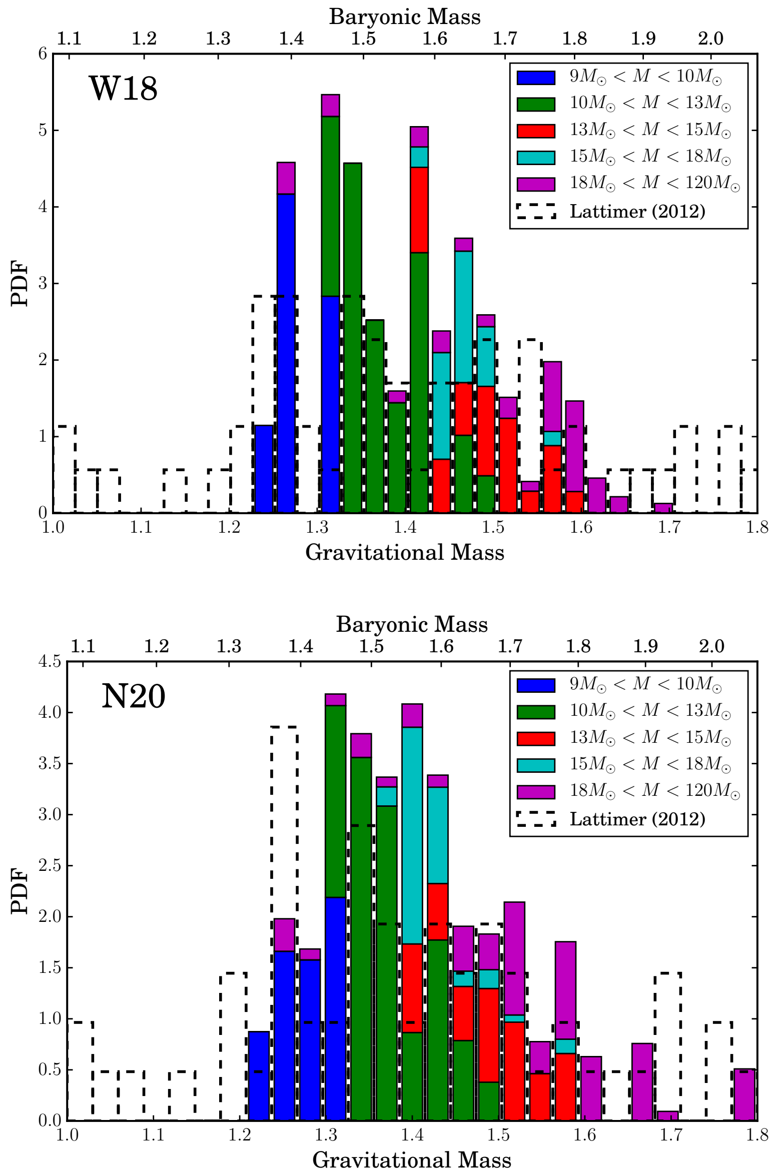


Figure 2.10: Distributions of neutron star masses for the explosions calculated using P-HOTB compared with the observational data from Lattimer (2012). Several neutron stars with masses greater than the maximum mass plotted here have been observed and the lightest observational masses have large error bars. In the top figure, the W18 calibration is used, and in the bottom, the N20. Both distributions show some weak evidence for bimodality around 1.25 and 1.4 $M_\odot$. Results have been color coded to show the main sequence masses contributing to each neutron star mass bin.

Cal.	$\overline{E}$ (erg)	$\overline{M_b}$	$\overline{M_{\mathrm{BH},l}}$	$\overline{M_{\mathrm{BH},u}}$	$\overline{M_{\mathrm{Ni,l}}}$	$\overline{M_{\mathrm{Ni,u}}}$	SN%	(> 12)	(> 30)
W15	0.68×10^{51}	1.55	8.40	13.3	0.040	0.049	66	47	2
W18	0.72×10^{51}	1.56	9.05	13.6	0.043	0.053	67	48	2
W20	0.65×10^{51}	1.54	7.69	13.2	0.036	0.044	55	37	0
N20	0.81×10^{51}	1.56	9.23	13.8	0.047	0.062	74	52	5

Table 2.2: Integrated Statistics (see Section 2.7.4 for descriptions; all masses in $M_\odot$)

explosions that have main sequence masses above 12 $M_\odot$ and 30 $M_\odot$. The successful explosions above 30 $M_\odot$ are all Type Ib or Ic.

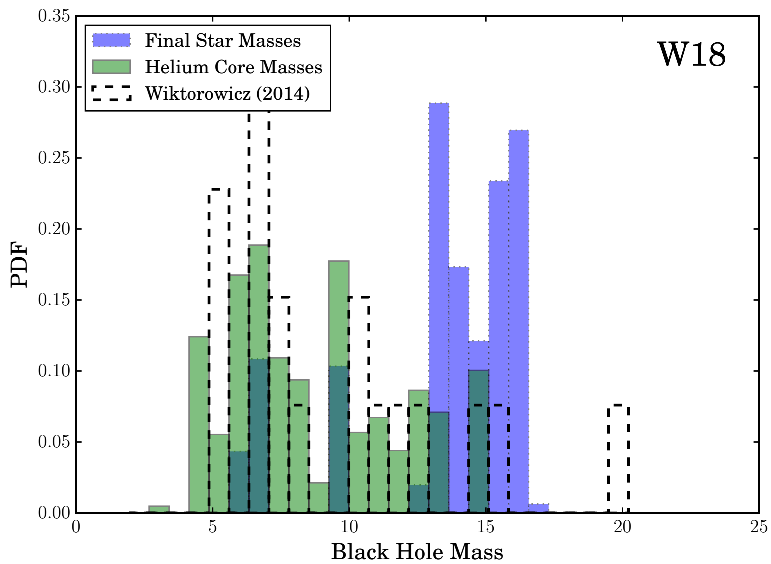
2.6.2 Black Holes

Figure 2.11 shows the distribution of black hole masses under each of two different assumptions for the accretion of the pre-supernova star: a) that only the helium core accretes onto the black hole, and b) that the entire presupernova star falls into the black hole. A distribution of IMF-weighted black-hole frequency, calculated just as it was for neutron stars, is given in Table 2.2. No subtraction has been made for the mass lost to neutrinos, i.e., the gravitational mass has been taken equal to the baryonic mass. It is expected that a proto-neutron star will form in all cases and radiate neutrinos until collapsing inside its event horizon. The amount of emission before trapped surface formation is uncertain, but unlikely to exceed the binding energy of the maximum mass neutron star, about 0.3 $M_\odot$ (O'Connor and Ott, 2011; Steiner *et al.*, 2013).

In addition to their production by stars that fail to launch a successful outgoing shock, black holes can also be made in successful explosions that experience a large amount of fallback. Only a few cases of this were found in the present survey, and the resulting black hole masses were always significantly less than the helium core mass. They were made in some of the most massive stars that exploded. Series W18 produced black holes by fallback at 27.2 and 27.3 $M_{\odot}$ with masses of 3.2 and 6.2 $M_{\odot}$. The stronger engine, N20, gave a few such cases at a slightly higher mass, resulting into black holes in the range 4.1 to 7.3 $M_{\odot}$. In all cases, the black hole mass was substantially less than the helium core mass, which ranged from 9.2 to 10.2 $M_{\odot}$.

2.7 Final Nucleosynthetic Yields, Calibrated to SN 1987A

Rather than discuss the yields of individual stars, this section gives and discusses a summary of the nucleosynthesis for the two main explosion series, averaged over an initial mass function. To give an indication of the variation with mass, separate results are given for three mass ranges: 9–12 $M_{\odot}$; 12–30 $M_{\odot}$; and 30–120 $M_{\odot}$. These ranges are chosen to represent the qualitatively different nature of supernova progenitors with very compact core structures (first group); supernovae that have made most of the elements (second group); and events where mass loss or an uncertain initial mass function plays a key role (third group). At the end, we also summarize the total nucleosynthesis (9–120 $M_{\odot}$) and consider the effect of adding a SN Ia component to pick up deficiencies in iron-group and intermediate-mass element production (Section



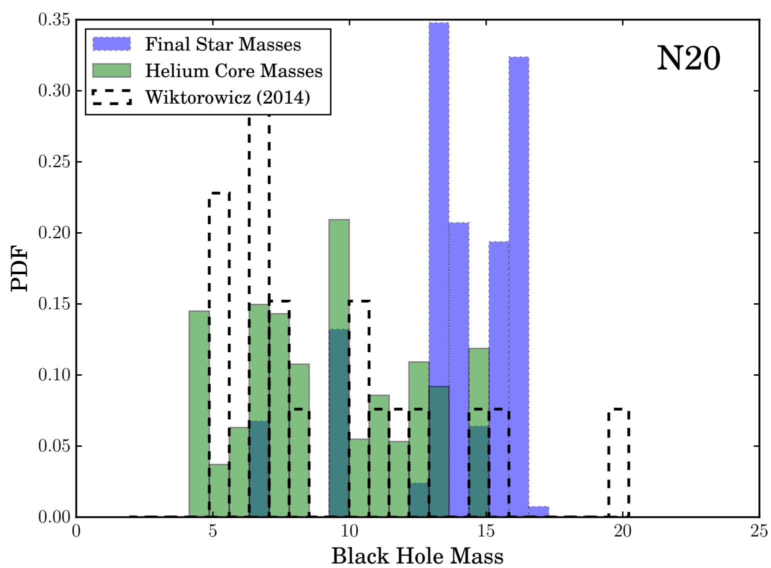


Figure 2.11: Distributions of black hole masses for the explosions calculated using P-HOTB compared with the observational data from Wiktorowicz *et al.* (2014) in gray. Theoretical results are shown based on two assumptions: 1) that only the helium core implodes (green); or 2) that the whole presupernova star implodes (blue). Observations are more consistent with just the helium core imploding.

2.7.5).

To proceed, the yields of each star, M_i, are separated into two categories: those species produced in the winds of the star during its presupernova evolution, and those ejected during the supernova itself. For each star, these yields, in solar masses, are multiplied by the fractional area under a Salpeter initial mass function (IMF) that describes the number of stars in each bin between M_i and M_{i+1}. The area has been normalized so that the total area under the IMF curve from 9 to 120 $M_\odot$ is 1. Summing the results across a mass range provides the average ejecta, in solar masses per massive star within that range.

A characteristic mass fraction is then computed which is the ratio of the IMF-weighted ejected mass of a given element (or isotope) produced in a mass range divided by the total mass of all species ejected. The computation may or may not include the wind; both cases are considered. The production factor is then the ratio of this mass fraction to the corresponding mass fraction in the Sun (Lodders, 2010).

Defined in this way, a production factor of 1 implies no change in the mass fraction of an isotope from the star or mass range considered compared to the presupernova progenitor. What came in is what goes out. The net production of a given species is proportional to 1 - P. There is also a change in all species since some mass stays behind in the bound remnant and is lost. This is not accounted for in the definition used here for P. If half of the mass of the star is lost to the remnant, but the mass ejected from the star is all of exactly solar composition, the production factor by our definition is still 1, but the net production would be this factor times $(M_{\rm ZAMS} - M_{\rm rem})/M_{\rm rem} = 1/2$. Values of remnant masses are given in Sukhbold *et al.* (2016). As presently designed, the production factor can only be less than 1 if the species is destroyed by nuclear reactions (e.g., deuterium, beryllium).

For purposes of presentation, all production factors are divided by that of ^{16}O so that the production factor of ^{16}O—an abundant element known to be produced in massive stars—is one. To retain information about the unnormalized production factor, a line is given in each figure at the reciprocal of the unnormalized production factor of ^{16}O. Points lying on this line have no net production, i.e, have an unnormalized production factor of 1.

In order to demonstrate the dependence on uncertain mass loss rates, especially in the heavier stars, some yields are presented with and without mass loss. The "without mass loss results" are an artificial construct because we have not carried out a self-consistent evolution and explosion of constant mass stars. The explosion of such stars would be quite different. For example, the more massive stars would all make black holes. The less massive ones might explode and eject the same matter the wind would have removed. Here, the mass lost in the wind is simply discarded. The effect of removing it shows its importance.

Because of the lower bound on the mass, 9 $M_\odot$, no contribution from lighter stars is included here, so it is to be expected that species like the heavy *s*-process, ^{13}C, ^{14}N etc. which are known to be efficiently produced in asymptotic giant brach stars,

Mass Range	Engine(s)	$\overline{\mathrm{M_{SN}}}$	$\mathrm{M_{frac}}$ (no wind)
≤12	Z9.6+W18	10.25	0.19 (0.40)
12-30	-	14.5	0.37 (0.54)
30-120	-	60.0	0.44 (0.010)
9-120	-	12.0	1
≤12	Z9.6+N20	10.25	0.18 (0.40)
12-30	-	15.2	0.39 (0.58)
30-120	-	80.0	0.43 (0.026)
9-120	-	12.25	1

Table 2.3: Selected Isotopes Production in $M_{\odot}$

will be under-represented. Any r-process from merging neutron stars or the neutrino wind is also missing, as are the products of cosmic ray spallation.

Some interesting statistics relevant to nucleosynthesis are given in Table 2.3. For the W18 and N20 series, the typical supernova mass, based only upon those stars that actually exploded, is given along with the fraction of the total ejecta that come from each mass range.

2.7.1 9 to 12 Solar Masses

While roughly half of all observed supernovae lie within this range, (Table 2.2), their contribution to the overall stellar nucleosynthesis is relatively small. Figure 2.12 shows reasonably good agreement with solar system elemental abundances for elements heavier than beryllium (Z = 4) all the way through the iron group, but this can be misleading since a lot of the yield was in solar proportions initially. The actual production factor for oxygen, before renormalization, is 2.37, so a value of

$1./2.37 = 0.422$ in the figure indicates no net change in the abundance due to stellar nucleosynthesis. This is apparent for the heaviest elements plotted, around Z = 40. To the extent that other elements have the same production factor as oxygen, their abundance is also about 2.37 times greater than what the star had, external to its final gravitationally bound remnant, at birth.

Figure 2.12 reveals some interesting systematics for low mass supernovae. First, the abundances are roughly solar, even for the iron group. Some elements, Co, Ni, and Cu are even overproduced. This is surprising given that only a minor fraction of iron-group elements is expected to come from massive stars (e.g., Timmes *et al.*, 1995; Cayrel *et al.*, 2004). More should come from Type Ia supernovae (SN Ia). Their apparent large production here reflects the fact that these stars make very little oxygen. As we shall see later, massive stars altogether make only about one-quarter of solar iron. The ratio of iron to oxygen in the remnants of low mass supernova remnants may be roughly solar.

2.7.2 12 to 30 Solar Masses

Stars from 12 to 30 $M_{\odot}$ are responsible for most of the nucleosynthesis that happens in supernovae. Figure 2.13 shows the production factors for this range. The intermediate mass elements are made in roughly solar relative proportions compared with oxygen, though the low yield of calcium is a concern. That the iron-peak elements are deficient is not surprising since most iron comes from Type Ia supernovae.

Some species in the iron group, especially Co, Ni, and Cu, are produced here in roughly solar proportions.

Of particular note in this mass range is a significant underproduction of *s*-process elements, especially Sc and the elements just above the iron group. This problem will persist in the final integrated nucleosynthesis from all masses and is a challenge for the present models and approach. Part of this deficiency may be due to an uncertain reaction rate for ^{22}Ne(α,n)^{25}Mg.

2.7.3 30 to 120 Solar Masses

Most stars above 30 $M_\odot$ in the present study become black holes, and their remnants absorb most of the core-processed elements. Their winds will escape prior to collapse and contribute to nucleosynthesis, especially of the lighter elements like CNO. Some stars in this range also explode, but only because winds have stripped the core of the star down to a manageable size. Such stars, having carbon–oxygen cores comparable to stars in the 12–39 $M_\odot$ range will contribute similar nucleosynthesis, but more enhanced in the products of helium burning.

Figure 2.14 shows that the winds contain excesses of ^{4}He, both from the envelopes of these massive stars that have been enriched by hydrogen burning and convective dredge up, and from the winds of Wolf-Rayet stars that have lost their envelopes altogether. This production significantly augments the primordial helium in the original star from the Big Bang and previous generations of stars. The nuclei ^{12,13}C, ^{14}N, and ^{16,18}O are also significantly enriched, but ^{15}N and ^{17}O are not. ^{22}Ne is greatly enhanced owing to Wolf-Rayet winds. This nucleus comes from two alpha captures on ^{14}N and is abundant in the helium shell from convective helium burning. Intermediate mass elements are underproduced compared with the large oxygen synthesis in the winds.

2.7.4 Integrated Yields from Massive Stars 9 to 120 Solar Masses

The integrated production factors from massive stars of all masses above 9 $M_\odot$ are shown in Figure 2.15 summed over isotopes. The individual isotopic abundances are given in Figure 2.16. The integrated production factor is calculated from a combined weighting of the three mass ranges just discussed. If $P_{\rm low}$ is the production factor for stars 9–12 $M_\odot$; $P_{\rm mid}$, for stars 12–30 $M_\odot$; and $P_{\rm high}$, for stars 30–120 $M_\odot$, the total production factor, $P_{\rm tot}$ is

$$P_{\rm tot} = f_1 P_{\rm low} + f_2 P_{\rm mid} + f_3 P_{\rm high} \qquad (2.6)$$

where f_i is given in Table 2.3. Because $P_{\rm low}$ is generally low and $P_{\rm high}$ is low except for species produced in the winds like CNO, the total production for most elements is dominated by $P_{\rm mid}$.

The probable nucleosynthetic sites for producing the various isotopes heavier than lithium have been summarized in Table III of Woosley *et al.* (2002), and our discussion follows that general delineation and omits references given there. We also include the full integrated yield of the N20 series in Figure 2.17. The primary difference is that the N20 model produces a bit more ^{16}O, which can be seen by comparing the baseline solar abundance of the heavy isotopes.

Additionally, one could vary the power of the initial mass function. Lowering the power from that of a Salpeter IMF of -2.35 to a value as low as -2.7 affects the results in a similar way to a reduction in the contributions of mass loss rates. The value of -2.7 was chosen as being a low value within the known uncertainties (e.g., Scalo, 1986; Chabrier, 2003). The results in Figure 2.18 show that a steeper IMF acts in the same way as reducing the contributions from the winds of very heavy stars, but with a smaller effect.

2.7.5 A Type Ia Supernova Contribution

Nucleosynthesis

The correct models for Type Ia supernovae are currently under much debate. In terms of nucleosynthesis, however, the models segregate into two general categories: explosions near the Chandrasekhar mass by a combination of deflagration and detonation, and prompt detonations in sub-Chandrasekhar mass white dwarfs. The requirement that any successful model makes $\sim$0.6 $M_\odot$ of ^{56}Ni in order to explain the light curve and $\sim$0.2 $M_\odot$ of intermediate mass elements determines the acceptable bulk nucleosynthesis, so that the models differ only in detail. Here, as a representative example of the Chandrasekhar mass model we use the nucleosynthesis from Nomoto's highly successful "W7" model (Iwamoto *et al.*, 1999). For the sub-Chandrasekhar mass model we use Model 10HC of Woosley and Kasen (2011). This detonation of a 1.0 $M_\odot$ carbon-oxygen white dwarf capped by 0.0445 $M_\odot$ of accreted helium produced 0.636 $M_\odot$ of ^{56}Ni in an explosion with final kinetic energy 1.2×10^{51} erg. The contribution from each SN Ia model was normalized so as to produce a solar proportion of ^{56}Fe/^{16}O when combined with the integrated yield of massive stars from 9 to 120 $M_\odot$.

The combined nucleosynthesis is given in Figure 2.19 for a sub-Chandrasekhar white dwarf and in Figure 2.20 for the Chandrasekhar mass model. Most of the intermediate mass elements are unchanged, though some of these isotopes do change. Not surprisingly, the introduction of Type Ia supernovae improves the overall fit especially for the iron group. The sub-Chandrasekhar mass model greatly improves the production of ^{44}Ca (made by helium detonation as ^{44}Ti) and seems to be required if only because of that unique contribution.

Implications for Supernova Rates

As noted in Sukhbold *et al.* (2016), massive stars are responsible for producing about 28% of the iron in the sun. The average SN Ia makes very nearly 0.6 $M_\odot$. That being the case, and assuming that the typical successful core collapse supernova (mostly SN IIp in the present study) makes 0.04 to 0.05 $M_\odot$ of new iron (Table

2.2), the implied ratio of SN Ia event rates to SN IIp is 1/6 to 1/5. If core-collapse supernovae made a little less iron, as seems likely, then the SN Ia rate would need to be a bit bigger. This ratio is consistent with observations for spiral galaxies similar to the Milky Way (Li *et al.*, 2011).

We can also use these abundances, particularly the fact that each massive star above 9 $M_\odot$ produces an average of 0.57 $M_\odot$ of ^{16}O, to infer a rate of Type II supernova (assuming stars below 9 $M_\odot$ produce little oxygen in their evolution). From the solar abundance of ^{16}O of 6.6×10^{-3} (Lodders, 2010), we find that there must have been 0.012 massive stars per solar mass of Population I material. Assuming that 66% of these stars become supernovae, as we found in our models, we find that integrated over galactic time, there must be 0.0076 core collapse supernovae (visible or invisible) per solar mass of Pop I material in order to explain the present solar abundance. This seems a reasonable, albeit very approximate number, especially if the supernova rate was different in the past.

All of the calculations here of yields and supernova properties have been calculated using the default prescriptions of convective overshoot and semi-convection for KEPLER (Weaver *et al.*, 1978). With a cohesive understanding of how the properties of these models can be compared to observations, we now move on to discuss how changing the prescriptions for these processes can affect the internal properties of these stars. These changes will affect both the supernovae and the nucleosynthesis of these populations of stars.

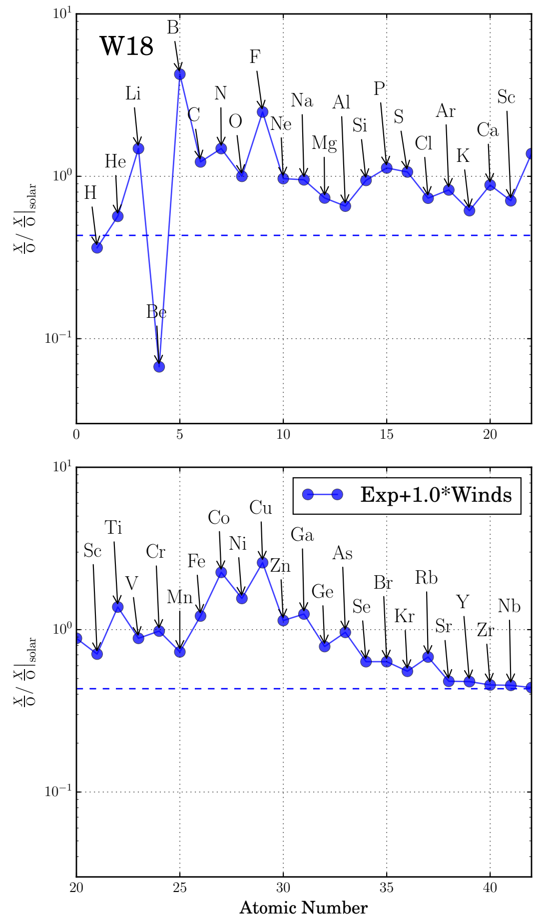


Figure 2.12: Nucleosynthesis in low mass supernovae. The IMF-averaged yields of just the explosions from 9 –12 $M_\odot$ are given for the elements from hydrogen through niobium. The calculations actually included elements up to bismuth (Z = 83), but no appreciable production occurred for the heavier elements. Production factors (defined in the text) have been normalized to that of ^{16}O, which is 2.37, by dividing them all by this factor, making the oxygen production unity. The dashed line at $(2.37)^{-1} =$ 0.43 is thus a line of no net change. Elements below the dashed line are destroyed in the star; those above it experience net production. The relative large yields of iron group elements, which are nearly solar compared to oxygen, is a consequence of the low oxygen yield in these light supernovae, and does not characterize heavier stars where more nucleosynthesis happens. The s-process just above the iron group is significantly underproduced. Note, however, the large yields of Li and B produced by the neutrino process. Fluorine is also overproduced by a combination of neutrinos and helium burning. Stellar winds of all models have been included.

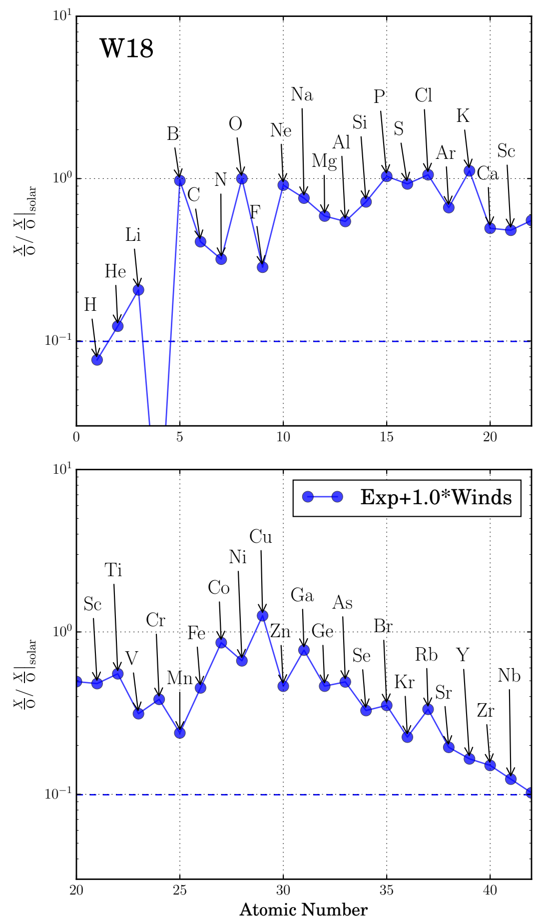


Figure 2.13: Similar to Figure 2.12, the IMF-averaged production of elements in supernovae from 12 to 30 $M_\odot$ is shown using the W18 calibration for the central engine. This is the range of masses where most supernova nucleosynthesis occurs. The production factor of oxygen before normalization was 10. Production factors above the dashed line at 0.10 thus indicate net nucleosynthesis in the stars considered. The production of iron-group elements is substantially lower than in Figure 2.12, but will be supplemented by SN Ia. The production of light s-process elements is substantially greater, but still inadequate to explain their solar abundances. Production of Li and B by the neutrino process is diminished, but still significant, especially for B. Contributions to nucleosynthesis by the winds of all stars, including those that became black holes is included.

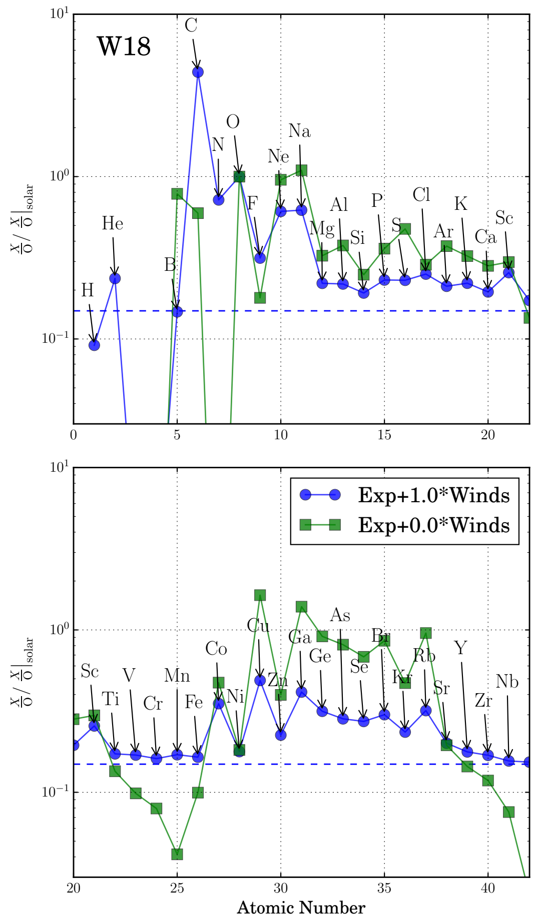


Figure 2.14: Similar to Figure 2.12, the IMF-averaged production of the elements is shown for stars from 30 to 120 $M_{\odot}$, using the W18 calibration for the central engine. Most, though not all stars in this very massive range become black holes, so their nucleosynthesis is dominated by the winds that they eject before dying. Production factors are normalized to oxygen. If winds are included, net production of an element occurs in this mass range if its average production factor is above 0.15 (blue dashed line). Also shown are the production factors if the wind is discarded. Since oxygen was being appreciably synthesized in the winds, neglecting them results in an overall renormalization of all production factors. The cut-off for the net production without winds falls to 0.010, off the scale of the plot. Including winds, these stars make so much carbon, nitrogen and oxygen that they can be responsible for little else.

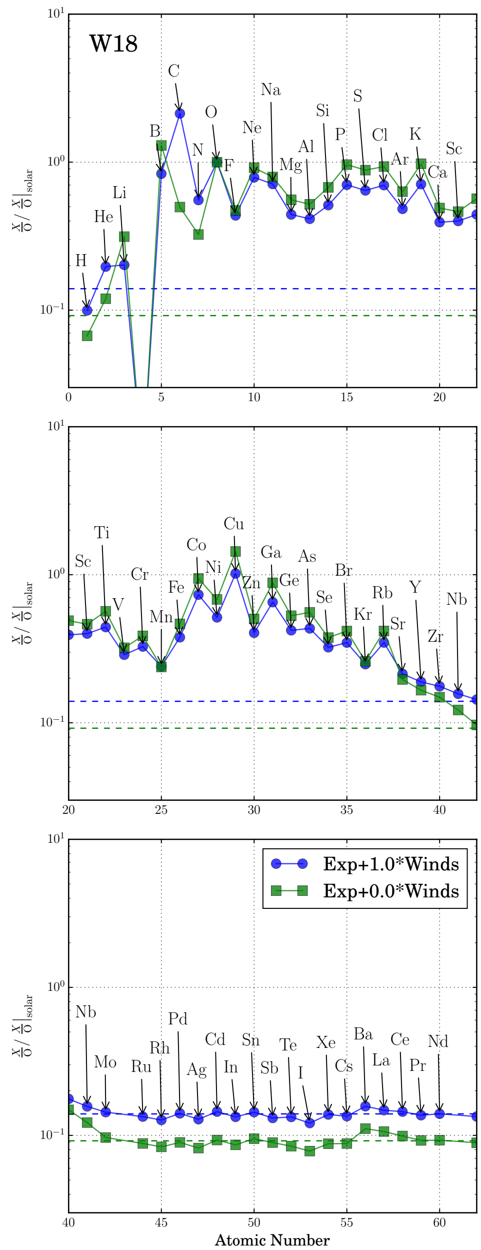


Figure 2.15: Integrated nucleosynthesis of the elements for all stars considered (9 – 120 $M_\odot$) using the W18 central engine. The results have been normalized to oxygen and the dashed line show net production occurring for production factors bigger than 0.14, if winds are included, 0.092 if they are not. The displacement of the overall normalization is due to the production of oxygen itself in winds. While individual isotopes may be affected (Figure 2.16), no net nucleosynthesis occurs for elements above Z = 40.

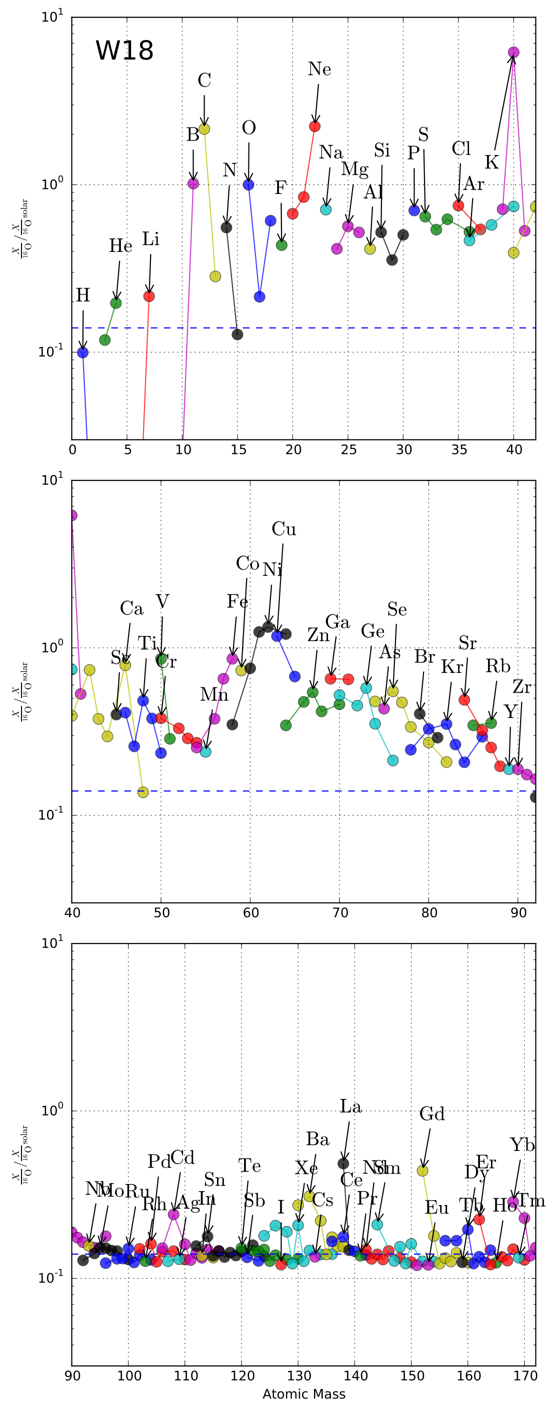


Figure 2.16: The same yields as in Figure 2.15 for the W18 central engine including the wind, but now given for the individual isotopes of each element.

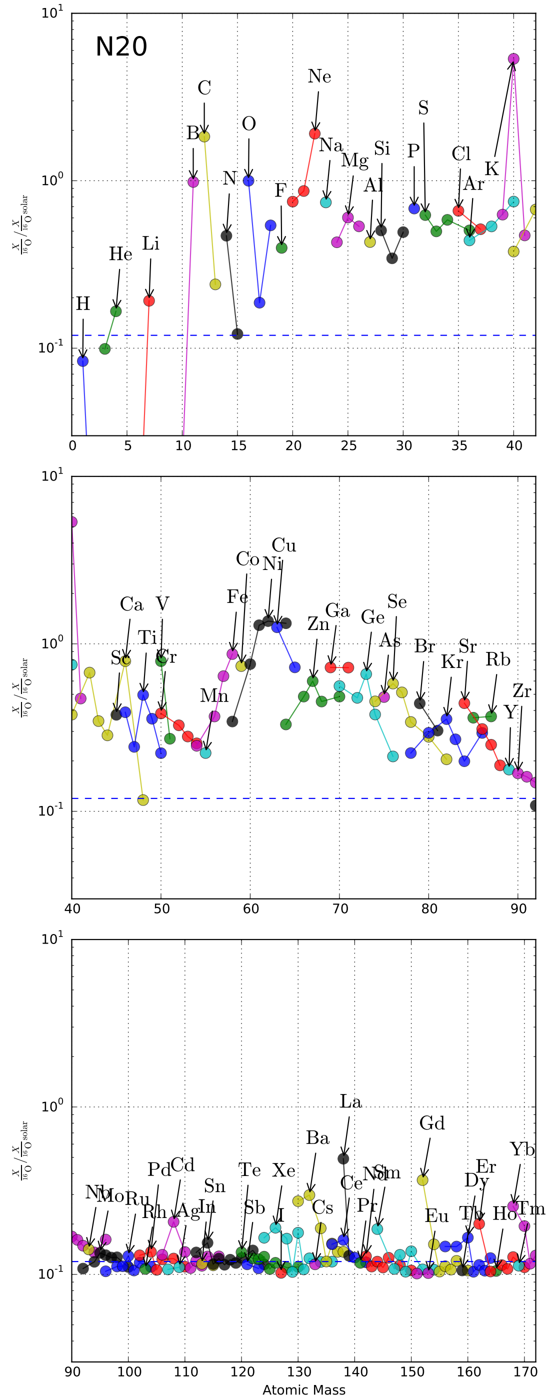


Figure 2.17: Similar to Figure 2.16, this figure gives results using the N20 central engine instead of W18. There are no significant differences.

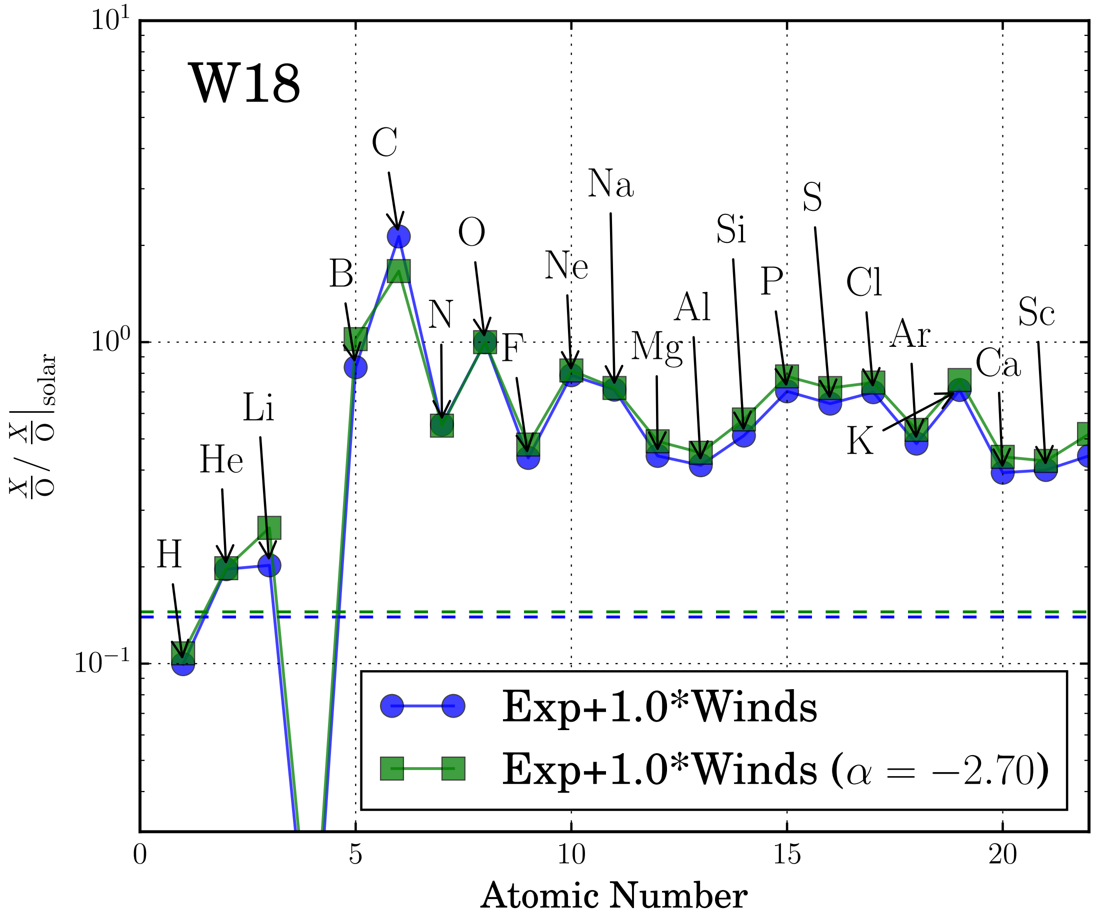


Figure 2.18: Similar to Figure 2.15, but using a steeper mass dependence for the IMF. Here the results of using a Salpeter IMF with a power-law dependence of -2.35 are compared to those obtained using a slope of -2.7. The dashed lines here indicate net production for both ensembles at 0.14.

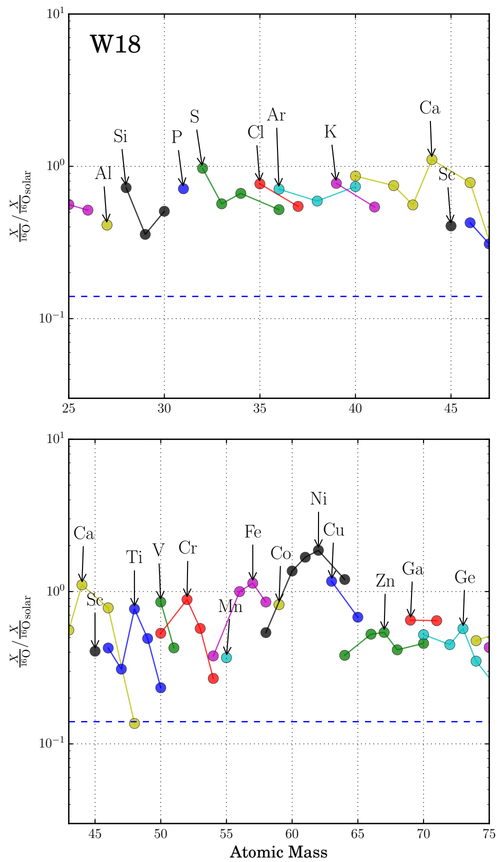


Figure 2.19: Similar to Figure 2.16, this figure includes contributions from all 9–120 $M_\odot$ supernovae and their winds, but also an additional component from SN Ia. Added to the average yield in Figure 2.16 is that of a typical sub-Chandrasekhar mass SN Ia (Woosley and Kasen, 2011) with a sufficient quantity of iron to make its solar abundance relative to oxygen. Not only does this variety of SN Ia make iron, but also raises the production of ^{44}Ca to its full solar value and increases the abundances of Si, S, Ar, and Ca so that they are closer to solar. Note, however, the underproduction of Mn/Fe. The dashed line for net production here is at 0.14.

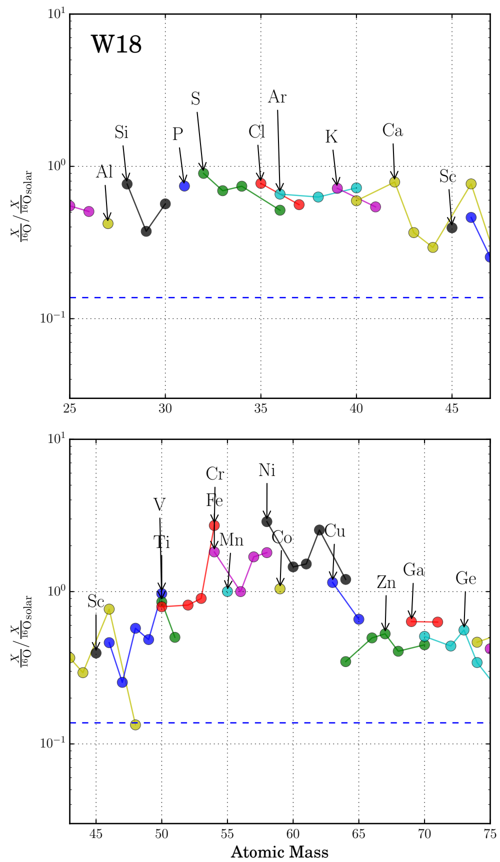


Figure 2.20: Similar to Figure 2.19, but including the yields of a Chandrasekhar mass model for SN Ia (Nomoto's W7) rather than the sub-Chandrasekhar model. The dashed line for net production here is at 0.14. Manganese is much improved, but ^{44}Ca is lost

Chapter 3

Minor Mixing Processes in Massive Stars

Chapter 2 provided some pertinent material on the nucleosynthesis and supernova properties expected from a theory where certain assumptions were made regarding convective mixing and energy transport. In this chapter, the central work of this thesis—namely the introduction of improved models for minor mixing processes—is now developed.

This work focuses on mixing processes in massive stars, which remain a large source of uncertainty in stellar evolution. These processes affect the appearance of stars as they evolve in the HR diagram, their structure when they die as supernovae, and the nucleosynthesis they produce. Two such processes are particularly uncertain: semi-convection and overshooting convection.

Novel prescriptions for semi-convection and convective overshoot are developed here based upon multi-dimensional hydrodynamical simulations of Wood *et al.* (2013) and Brummell *et al.* (2002), respectively. These models are implemented into the stellar evolution code KEPLER and used to investigate the role of semi-convection and overshooting convection in altering the evolution of a typical massive star, one with an initial mass of 15 $M_\odot$.

3.1 Traditional Treatment of Semi-Convection

As discussed in Chapter 1, semi-convection is a form of "double-diffusive instability" that occurs when a fluid is stable in density but is unstable in temperature, and when two of its components (temperature and composition in the stellar case) diffuse at different rates. The condition for density stability is the Ledoux criterion, which can be expressed as

$$\nabla_{\mathrm{rad}} < \nabla_{\mathrm{ad}} + \frac{\phi}{\delta}\nabla_{\mu}, \tag{3.1}$$

where $\nabla_{\mathrm{rad}} \equiv \frac{\partial \ln T}{\partial \ln p}$ for a fluid whose energy transport is solely due to photon diffusion, $\nabla_{\mathrm{ad}} \equiv \left.\frac{\partial \ln T}{\partial \ln p}\right|_{\mathrm{ad}}$, $\nabla_{\mu} \equiv \frac{\partial \ln \mu}{\partial \ln p}$ (where μ is the mean molecular weight), $\delta \equiv -\frac{\partial \ln \rho}{\partial \ln T}$, and $\phi \equiv -\frac{\partial \ln \rho}{\partial \ln \mu}$, as are standard in the field of stellar evolution. In the absence of compositional gradients, the Ledoux condition simplifies to the Schwarzschild criterion, which is a temperature instability criterion that can be expressed instead as

$$\nabla_{\mathrm{rad}} < \nabla_{\mathrm{ad}}. \tag{3.2}$$

Hence semi-convection is traditionally invoked when

$$\nabla_{\mathrm{ad}} < \nabla_{\mathrm{rad}} < \nabla_{\mathrm{ad}} + \frac{\phi}{\delta}\nabla_{\mu}. \tag{3.3}$$

The dynamics of semi-convection depend on several physical quantities. As this is a double-diffusive instability, diffusion of momentum, temperature, and composition are all critical to its manifestation. Mathematically, they can be accounted for in terms of two non-dimensional numbers: the Prandtl number, Pr, which is the ratio of the kinematic viscosity to the thermal diffusion coefficient (ν/κ_T); and τ, the ratio of the concentration diffusion coefficient to the thermal diffusion coefficient (κ_μ/κ_T). In addition, the relative stratification of temperature and mean molecular weight also plays a fundamental role, appearing non-dimensionally as the density ratio:

$$R_0 \equiv \frac{\alpha\left(\frac{dT}{dr} - \frac{dT}{dr}\big|_{\mathrm{ad}}\right)}{\beta\frac{d\mu}{dr}} = \frac{\nabla - \nabla_{\mathrm{ad}}}{\frac{\phi}{\delta}\nabla_{\mu}}, \tag{3.4}$$

where $\frac{dT}{dr} - \frac{dT}{dr}\big|_{\mathrm{ad}}$ is the superadiabatic temperature gradient, $\frac{d\mu}{dr}$ is the gradient of the mean molecular weight, α is the thermal expansion coefficient $(-\rho^{-1}\left(\partial\rho/\partial T\right)|_{\mu,p})$, and β is an analogous coefficient for the mean molecular weight $(\rho^{-1}\left(\partial\rho/\partial\mu\right)|_{T,p})$; the quantities $\frac{dT}{dr}\big|_{\mathrm{ad}}$, α, and β come directly from the equation of state.

Walin (1964) and Kato (1966) found through a linear stability analysis of the governing equations that a doubly-stratified system is unstable to overstable oscillations only when

$$\frac{\mathrm{Pr}+\tau}{\mathrm{Pr}+1} \leq R_0 \leq 1, \text{ or equivalently,} \tag{3.5}$$

$$\frac{\mathrm{Pr}+\tau}{\mathrm{Pr}+1}\left(\nabla_{\mathrm{ad}} + \frac{\phi}{\delta}\nabla_{\mu}\right) \leq \nabla_{\mathrm{rad}} \leq \nabla_{\mathrm{ad}} + \frac{\phi}{\delta}\nabla_{\mu}. \tag{3.6}$$

Note that $R_0 = 1$ is equivalent to the Ledoux criterion and so the system is unstable to convection if $R_0 > 1$. Conversely, the Schwarzschild condition occurs at $R_0 = 0$, so these oscillations do not span the entire semi-convective regime, defined in Equation 3.3. In practice, however, Pr and τ are both quite small in stellar interiors ($\sim 10^{-6}$), so the parameter space that is unstable to the semi-convective instability, which spans $\sim 10^{-6} \leq R_0 \leq 1$, is essentially equivalent to it. Stellar evolution of massive stars reveals that, in semi-convective zones, the typical value of R_0 is roughly one half, implying that the contributions to buoyancy from thermal and compositional stratifications are comparable.

Traditionally, models of semi-convection merely provide prescriptions of the mixing of species concentration. Semi-convection may also provide substantive heat transport, but this has received very little attention and so is not discussed here.

Semi-convection has been discussed in the context of stellar evolution since Schwarzschild and Härm (1958). Many of the models from that era treated semi-convection as an instantaneous mixing process. Schwarzschild and Härm (1958) and Sakashita *et al.* (1959) phenomenologically argued that—in main sequence stars with mass greater than 10 $M_\odot$—semi-convective regions will initially extend a core convection zone; however, any entrained helium from the core will lower the opacity, shutting down convection and bringing the zone into marginal convective–radiative stability. Numerically, this is done by altering the composition profile to make the semi-convective region marginally stable to convection. Several groups (e.g., Paczyński, 1970; Castellani *et al.*, 1971) later applied a similar phenomenology to helium core burning stars.

Both these models make the assumption that the mixing of a semi-convective region will approach convective neutrality without paying too much attention to the physical nature of the process. It is not guaranteed that semi-convection will bound a convection zone, and so semi-convection may not always behave according to this phenomenological picture. Instantaneous mixing models like these offer only a limited view into the physics of the problem and are certainly not applicable to times in a stellar life where the mixing and evolution time scales are comparable. This limitation was removed by the pioneering work of Eggleton (1972) among others, who argued for a diffusive implementation of mixing length theory and semi-convection. This has been the common treatment of semi-convection since.

Diffusive models of semi-convection assume that semi-convective mixing of species concentration can be described diffusively. A diffusion coefficient that depends on local quantities is associated to each zone of the stellar model The evolution of the concentration profile for each species takes the form of a diffusion equation with a spatially varying diffusion coefficient. At present, the most commonly used model for semi-convection is from Langer *et al.* (1983).

Langer *et al.* (1983) constructed a diffusion coefficient to represent semi-convection from the work of Kato (1966). Assuming that the dynamics of semi-convection are dominated by the overstable oscillations described in Section 3.1, Langer *et al.* (1983)

proposed a model in which the diffusion coefficient, D_{semi}, takes the form (dimensionally)

$$D_{\mathrm{semi}} \sim l^2 a, \tag{3.7}$$

where l is the length scale of the waves and a is their growth rate, both obtained from linear theory (i.e., by ignoring non-linear terms and assuming plane wave solutions that could grow and oscillate in amplitude). They then showed that, approximately,

$$D_{\mathrm{semi}} = \alpha_{\mathrm{semi}} \frac{\kappa_T}{6} \frac{\nabla - \nabla_{\mathrm{ad}}}{\nabla_{\mathrm{ad}} + \frac{\phi}{\delta}\nabla_\mu - \nabla}, \tag{3.8}$$

where α_{semi} is an order-one free parameter, κ_T is the thermal diffusion coefficient from photon diffusion, and ∇, ∇_{ad}, ∇_μ, ϕ, and δ take their usual meanings in stellar evolution.

As mentioned in Chapter 1, it has long been known from geophysical studies that semi-convection can also take the form of stacks of large-scale convective layers. Layer formation in double-diffusive convection has been a focus of substantial experimental (e.g., Turner, 1965; Shirtcliffe, 1973) and more recent numerical work (Noguchi and Niino, 2010; Carpenter *et al.*, 2012).

Spruit (1992) (see also Spruit, 2013) originally applied this layered picture of semi-convection to propose an alternative to the model of Langer *et al.* (1983). He assumed that the interfaces between layers were primarily diffusive and solved for the layer structure that would transport a given luminosity. For this equilibrated system, he calculated an effective concentration diffusion coefficient of

$$D_{\rm semi} = (\kappa_\mu \kappa_T)^{\frac{1}{2}}\,\frac{\nabla_{\rm rad} - \nabla_{\rm ad}}{\nabla_\mu}, \tag{3.9}$$

for an ideal gas where the definitions are the same as in Equation 3.8; $\nabla_{\rm rad}$ takes its usual astrophysical meaning. A more general form is provided in Spruit (1992) to include contributions of radiation pressure. We now move away from these existing prescriptions to propose a new model based on recent numerical simulations (see Section 3.3).

3.2 Traditional Treatment of Convective Overshoot

As discussed in Chapter 1, "convective overshoot" refers to the dynamical processes by which fluid motions originating in a convective zone travel past the theoretical edge of the zone. This process could behave in at least two ways: strong convective overshoot may mix both composition and entropy and weak convective overshoot may mix only composition. Convective overshoot that is strong enough to mix some fraction of the radiative region adiabatically is known as "penetrative convection." Whether convective overshoot is penetrative or not in stellar contexts remains an open question and has received much recent numerical attention (e.g., Hurlburt *et al.*, 1994; Brummell *et al.*, 2002; Rogers and Glatzmaier, 2005). When necessary to distinguish between outward and inward convective fluid motion, "undershooting convection" specifies the inward motion and "overshooting convection," the outward motion. The processes are generally treated the same in stellar evolution, but as has been shown in compressible simulations (e.g., Hurlburt *et al.*, 1994), they can behave quite differently due to effects of compression and spherical geometries.

In early stellar models, overshoot was viewed as an extension of the convection zone and simply served to advance the boundary beyond nominal stability (e.g., Castellani *et al.*, 1971; Maeder, 1975; Bertelli *et al.*, 1981; Doom *et al.*, 1986). In this sense, there was no distinction between overshoot and penetration, and mixing in the overshoot region was instantaneous.

Deng *et al.* (1996), Freytag *et al.* (1996), and Herwig (2000) were some of the first to treat overshoot as a diffusive process instead of an instantaneous one over a region some fraction of a pressure scale height; this marked a significant departure in thinking of the nature of the zone. The model of Freytag *et al.* (1996) remains one of the most commonly used models today for convective overshoot—it is used for example in the MESA star code (Paxton *et al.*, 2013). This model is inspired from the general behavior of hydrodynamical simulations. The two-dimensional hydrodynamical simulations of Freytag *et al.* (1996) suggest that the root-mean-square velocity fluctuations decay exponentially away from the edge of the convection zone. They recommended the following diffusion coefficient, D_{osht}:

$$D_{\mathrm{osht}} = D_{\mathrm{conv},0} e^{-\frac{2|r-r_0|}{f H_p}}, \tag{3.10}$$

where D_{osht} is the composition diffusion coefficient provided by the model, H_p is the pressure scale height, $D_{\mathrm{conv},0}$ is the diffusion coefficient predicted by mixing length theory just above the radiative–convective boundary, r_0 is the radius of the boundary, and f is a free parameter, usually taken to be a few percent or less. The model of Freytag *et al.* (1996) ultimately relies on comparisons with observations to determine f.

More recently, Canuto (2011) has been advocating a more physically motivated turbulence scheme to replace mixing length theory and indeed all turbulent diffusive processes in stars with a single all-encompassing theory, often referred to as a turbulent convection model. These models are claimed to be more physical than the current prescriptions of radiative and convective transport than are currently used in stellar evolution codes; however, they would be difficult to implement numerically and particularly so in any code with implicit hydrodynamics. These models also depend on a closure scheme, which is parameterized and on which the model is quite dependent. Nevertheless, a turbulent convection model has been implemented into a stellar evolution code by Zhang (2012), and they find they are able to reproduce stellar observations.

Much work has been done in recent years on simulations of convective overshoot with direct numerical simulations (i.e., using codes that explicitly dissipate momentum). Some of the major results are summarized here. Sofia and Chan (1984) and Hurlburt *et al.* (1986) were among the first to simulate compressible convective overshoot in two dimensions with full hydrodynamics. In these simulations, three parameters dominate the system dynamics.

The most important parameter is the total energy of the convective motions, or equivalently their turbulent Reynolds number, which is the ratio of inertial forces to viscous forces. This number can be determined from the Rayleigh number in the convection zone, defined as

$$\mathrm{Ra} \equiv \frac{\mathrm{g}\alpha \left| \frac{\partial \mathrm{T}}{\partial \mathrm{r}} - \frac{\partial \mathrm{T}}{\partial \mathrm{r}}\Big|_{\mathrm{ad}} \right| \mathrm{H}^4}{\nu \kappa_{\mathrm{T}}}, \tag{3.11}$$

where g is the acceleration due to gravity, α is the coefficient of thermal expansion, $\frac{\partial T}{\partial r} - \frac{\partial T}{\partial r}\big|_{\mathrm{ad}}$ is the superadiabatic temperature gradient, H is the height of the convection zone, ν is the kinematic viscosity, and κ_T is the thermal diffusion coefficient. This parameter measures the stratification of the convection zone and is instrumental in determining the final strength of the convection.

Hurlburt *et al.* (1994) realized the need to introduce another important parameter, the stiffness, S, which is a comparative measure of the buoyancy between the convective and radiative regions: how stably stratified is the radiative zone compared with how strongly driven is the convection. The formal definition of the stiffness is presented in Section 3.4 and can vary in stars from values near unity for convection in a supergiant envelope to several hundred for the undershooting convection beneath the helium burning shell in later stages of evolution. In particular, for very low stiffness ($\lessapprox 1$), convective overshoot becomes penetrative.

The final parameter is the Prandtl number, the ratio of kinematic viscosity to the thermal diffusion coefficient (ν/κ_T). This parameter has a less direct effect on the system dynamics but is much lower than one in the astrophysical regime. Because the viscous scale is generally much smaller than the relevant physical scales in astrophysics, approximations must be made in order to simulate such a fluid. Thus, for direct numerical simulations of astrophysical fluids, the viscosity must be enhanced by many orders of magnitude in order for the viscous scale to be resolved and the Prandtl number may be artificially high.

The pressure scale height of the system, $H_p \equiv -p\,(dp/dr)^{-1}$ also plays a prominent role in many implementations of convective overshoot as an effective length unit. The pressure scale height typically enters into the convective mixing length from mixing length theory as the typical distance a buoyant plume can travel. The mixing length, and therefore the pressure scale height, serves as the typical expected travel distance of convective overshoot. This may or may not be the typical physical extent of a buoyant plume in turbulent convection with rotation, magnetic fields, degenerate gases, etc., and much information could be gained from more detailed numerical and experimental studies of convection.

Hurlburt *et al.* (1994) later parameterized their two-dimensional simulations of convective overshoot in terms of the stiffness parameter. Informed by their simulations, they analytically derived expected power laws of the behavior of penetrative and non-penetrative convective overshoot, which will be discussed in Section 3.4.

Additionally, Meakin and Arnett (2007) modeled stellar convection with an implicit compressible large eddy simulation, which uses numerical dissipation to model viscosity and species diffusion. This is in contrast to any other numerical schemes (e.g., Hurlburt *et al.*, 1994; Brummell *et al.*, 2002), which explicitly model viscosity. It is not the goal of this paper to argue as to whether convection in the limit of vanishing viscosity is better modeled with or without explicit viscosity. The latter case has been discussed substantially in astrophysics, and the goal here is to complement the literature with some discussion of implications of the former. For an excellent review of the effects of viscosity in large Reynolds number flows, the authors recommend the reader to Batchelor (2000).

Meakin and Arnett (2007) simulated turbulent convection with stellar parameters in 3D. They modeled the shell and core convection of a 23 $M_\odot$ star burning oxygen in its core. They found that convective overshoot is best described by a mass entrainment model, where mass is added to the convection zone according to

$$\dot{M}_E = \left(4\pi r^2 \rho\right) \sigma_H 10^{A - n \log \mathrm{Ri_B}}, \tag{3.12}$$

where $\dot{M}_E$ is the mass entrainment rate, σ_H is a characteristic turbulent velocity, A is a free parameter, $\mathrm{Ri_B}$ is the bulk Richardson number, and n is the power law dependence of the mass entrainment on $\mathrm{Ri_B}$. This prescription will be discussed in more detail in Section 3.6.1.

In this paper, we shall seek a new prescription for convective overshoot, calibrated against modern fully compressible direct numerical simulations. This is discussed in Section 3.4.

This paper is organized as follows. Sections 3.3 and 3.4 present new numerically-inspired models for semi-convection and convective overshoot, respectively. Section 3.5 contains some results acquired from implementing these models into the stellar evolution code KEPLER. In Section 3.6, the implications of these results on stellar evolution and potential observable properties of massive stars are discussed. Some concluding remarks are included in Section 3.7.

3.3 Semi-Convection Model

Recently, much numerical progress has been made in studies of semi-convection. Rosenblum *et al.* (2011) investigated semi-convection under the Boussinesq approximation while approaching astrophysical parameter space. They found that their simulations could develop into homogeneous turbulence or stacks of convective layers. They explained the existence of layers by reapplying the γ-instability of Radko (2003) to the semi-convection. Later, Mirouh *et al.* (2012) used numerical simulations to determine the physical parameters under which layers were likely to form. By extrapolating their results to the astrophysical regime, they found that the critical R_0 for layer formation at low Pr was roughly $\tau^{-\frac{1}{2}}$. Thus, for stellar interiors, the layered regime extends from $R_0 \sim 10^{-3}$ to $R_0 = 1$. The typical empirical range of R_0^{-1} in semi-convective regions in KEPLER simulations is 1.4 ± 0.1 and so the majority of parameter space is likely in the layered regime. This paper then focuses mainly on semi-convection in the layered regime.

Following the work of Rosenblum *et al.* (2011) and Mirouh *et al.* (2012), Wood *et al.* (2013) simulated layered semi-convection in three-dimensions under the Boussinesq approximation. They derived expressions for the flux of temperature and species concentration as a function of Pr, τ, R_0, and of the layer height observed in their simulations.

Wood *et al.* (2013) simulated semi-convection in a three-dimensional domain with periodic boundary conditions under the Boussinesq approximation. For values of R_0 close to unity, their simulations take the form of convective layers, i.e., a vertical stack of convection regions separated by thin turbulent interfaces. While the density gradient is unstable within a layer, the gradient is stable when averaged over the entire domain. Using this setup, Wood *et al.* (2013) measured the vertical fluxes of composition and heat through these staircases.

Wood *et al.* (2013) found that these layers initially form with small height (an order of magnitude larger than the size of the initial instability). Over time, the layers merge through a mechanism that is as of yet unknown, increasing in height. As the layers become larger, the flux through the staircase increases dramatically. This process stops only when one layer fills the entire simulated domain.

The results of Wood *et al.* (2013) can be implemented into a one-dimensional flux model. They provide analytical expressions for the composition fluxes from their simulations in terms of a Nusselt number (the ratio of the total flux of the mean molecular weight or concentration to the background diffusive flux). The Nusselt number measures the total flux of a quantity (in this case, species concentration) in units of the diffusive flux (e.g., a Nusselt number of unity indicates no turbulent motion, whereas a value of 100 indicates that semi-convection mixes concentration at 100 times the rate of molecular diffusion). Wood *et al.* (2013) empirically found that their simulations could be represented by to their Equation 43:

$$\mathrm{Nu}_{\mu} - 1 = \mathrm{B}\tau^{-1}\mathrm{Pr}^{\frac{1}{4}}\mathrm{Ra}^{0.37}, \tag{3.13}$$

where B has a typical value of 0.03 and Ra is the Rayleigh number for a single semi-convective layer. The Rayleigh number is defined as (their Equation 38)

$$\mathrm{Ra} = \frac{\mathrm{g}\alpha \left| \frac{\mathrm{dT}}{\mathrm{dz}} - \frac{\mathrm{dT}}{\mathrm{dz}}\Big|_{\mathrm{ad}} \right| (\mathrm{l}_{\mathrm{semi}})^4}{\nu\kappa_{\mathrm{T}}}, \tag{3.14}$$

where g is the acceleration due to gravity, ν is the kinematic viscosity, and κ_T is the thermal diffusion coefficient. The quantity l_{semi} is the physical height of each layer. While l_{semi} can be directly observed in a numerical simulations, it remains unknown what the actual layer height in a star may be. This, unfortunately, remains a free parameter.

The form of Equation 3.13 can be understood physically as follows. The dependence on τ will be made clear when the equation is recast into an effective diffusion coefficient in Equation 3.15. The dependence on Ra is not surprising; a larger convective layer leads to a larger value of Ra and therefore an increased flux. The exact power is empirically fit to their simulations but is not far from the 1/3 power usually associated with Rayleigh Bènard convection. It is also expected from dimensional analysis. However, to determine the effective value of the Rayleigh number for the semi-convective layers in stellar evolution, l_{semi} must be known.

Equation 3.13 can be expressed in terms of a dimensional turbulent diffusion coefficient if the total flux is assumed to take the form $\mathbf{F}_\mu = -\left(\kappa_\mu + \kappa_{\mu,\text{turb}}\right)\nabla\mu$. Such an assumption has only a marginal basis in fluid dynamics, but it makes for a somewhat physically justified prescription that is simple to implement. Given this assumption, $\text{Nu}_\mu - 1$ can be expressed as $\kappa_{\mu,\text{turb}}/\kappa_\mu$, and so Equation 3.13 becomes:

$$\kappa_{\mu,\text{turb}} = B\text{Pr}^{\frac{1}{4}}\text{Ra}^{0.37}\kappa_{\text{T}}. \tag{3.15}$$

All of these quantities, save the layer height, can be calculated from the local properties within the semi-convection zone. Given this prescription, a local diffusion coefficient that characterizes mixing by semi-convection can easily be constructed within a stellar calculation; the only unknown quantity is the layer height.

3.3.1 Implementation of Semi-Convection in Stellar Evolution Codes

In practice, a model for l_{semi} is needed to implement the above calculation into a stellar evolution code. Moore and Garaud (2016) let l_{semi} be a constant multiplier of a pressure scale height and investigated the effect of varying l_{semi} on the evolution of stars with initial masses from 1.2 to $1.7M_\odot$. Moore and Garaud (2016) found that semi-convection using a model with a layer height larger than a critical value of $10^{-8}H_p$ was nearly equivalent to running simulations that ignored semi-convection altogether and used the Scwarszchild criterion for convection. Because the transition occurs at scales smaller than the minimum expected layer height from simulations, they concluded that for long-lived stages of stellar evolution (i.e., core hydrogen and helium burning), using the Schwarschild criterion for convection adequately approximates the results. However, because semi-convection is relevant also on shorter time scales during the late stages of massive star evolution, the choice of l_{semi} could be important.

The layer height may depend on any of the fundamental parameters of the system in a way that is thus far unknown. Spruit (2013) constructed an analytic formalism in which these layers were forced to transport a fixed quantity flux. He assumed that the fluxes through the interfaces between the layers were primarily diffusive. Given this assumption, he determined the equilibrium layered state for which the fluxes through the layers matched those through the interfaces and found that the layer height depended on the relative gradients of composition and temperature, specifically, the density ratio. A diffusive picture of interfacial transport disagrees with the simulations of Wood *et al.* (2013); however, there the interfaces are primarily turbulent in nature. Because the quantitative picture of the work of Spruit (2013) contradicts the simulations from which the prescription of this paper is constructed, one should be cautious using his scaling relation as a function of R_0 with this prescription.

Given the current lack of any simulations to elucidate the answer, it is possible that the qualitative picture of layer equilibration for a given flux may be accurate for turbulent interfaces as well. Physically, if l_{semi} depends on R_0, this implies that the strength of convecting layers is intrinsically limited by the compositional and thermal gradients of the system. Such a limit is physically motivated since the available potential energy that can be converted to convective motion is stored in the mean entropy gradient. The exact dependence on R_0 in Spruit (2013) is tied to the assumption of interfacial properties, so here, l_{semi} is instead assumed to have a power law dependence on R_0 with an unknown exponent, m. A relevant length scale is still needed to construct l_{semi}, and for ease of comparison with Moore and Garaud (2016), the pressure scale height is used here, resulting in a final expression of

$$l_{\mathrm{semi}} = \alpha_{\mathrm{semi}} R_0^m H_p, \tag{3.16}$$

where α_{semi} and m are free parameters. This ad-hoc prescription is used to explore whether a dependence on this quantity has discernable effects on the mixing within semi-convection zones.

This prescription must be interpreted in the context of physical limits however. The value of $\alpha_{\mathrm{semi}} H_p$ should never be greater than the mixing length from mixing length theory, as this would suggest that semi-convection can be more efficient than convection. In KEPLER, the mixing length defaults to one pressure scale height, and as such, α_{semi} should certainly not be larger than unity for the work discussed here. If a value must be chosen for α_{semi}, it is not unreasonable to assume a value of unity, as this establishes that a semi-convection layer with a height comparable to that of an entire convection zone should mix nearly as efficiently as pure convection.

3.4 Convective Overshoot Model

In this work, we choose to create a model for convective overshoot based on the hydrodynamical simulations of Brummell *et al.* (2002). The specifics of the simulations of Brummell *et al.* (2002) are discussed here, followed by a detailed description of the model and its implementation into a stellar evolution code.

Brummell *et al.* (2002) used a setup similar to the one of Hurlburt *et al.* (1994) and simulated convective overshoot with compressible hydrodynamics, varying the stiffness of the barrier between the convective and radiative regions. The convection in these simulations was driven by enforcing a background state with a superadiabatic polytrope; that is $p = K\rho^{\frac{n+1}{n}}$, where n, the polytropic index, is smaller than $n_{\rm ad}$, the adiabatic polytropic index. The adiabatic poltropic index ($n_{\rm ad} = 1/(\gamma - 1)$, where γ is the ratio of the specific heats, c_p/c_v) describes a fluid marginally stable to convection. On the other hand, the radiative region of these simulations has a background state with a subadiabatic polytrope. The relative strengths of these are measured via a quantity called the "stiffness" (Hurlburt *et al.*, 1994):

$$S = \frac{n_{\rm rad} - n_{\rm ad}}{n_{\rm ad} - n_{\rm conv}}, \tag{3.17}$$

where $n_{\rm conv}$ and $n_{\rm rad}$ are the polytropic indices of the convective and radiative regions, respectively.

For ease of interpretation, the stiffness can be recast in terms of derivatives of pressure with respect to density by recognizing that $n = \left(\frac{\partial \ln p}{\partial \ln \rho} - 1\right)^{-1}$, which comes from manipulating the polytropic equation of state. Through simple algebra, the stiffness can be expressed as

$$S = \frac{\gamma - \left.\frac{\partial \ln p}{\partial \ln \rho}\right|_{\mathrm{rad}}}{\left.\frac{\partial \ln p}{\partial \ln \rho}\right|_{\mathrm{conv}} - \gamma} \frac{\left.\frac{\partial \ln p}{\partial \ln \rho}\right|_{\mathrm{conv}} - 1}{\left.\frac{\partial \ln p}{\partial \ln \rho}\right|_{\mathrm{rad}} - 1}. \tag{3.18}$$

In practice, the deviation from adiabaticity is usually small, but the deviation of $\partial \ln p/\partial \ln \rho$ from 1 (the value for an isothermal polytrope) is usually large and does not change substantially across the interface. Thus, the stiffness can be approximated as

$$S \approx \frac{\left.\frac{\partial \ln p}{\partial \ln \rho}\right|_{\mathrm{ad}} - \left.\frac{\partial \ln p}{\partial \ln \rho}\right|_{\mathrm{rad}}}{\left.\frac{\partial \ln p}{\partial \ln \rho}\right|_{\mathrm{conv}} - \left.\frac{\partial \ln p}{\partial \ln \rho}\right|_{\mathrm{ad}}}. \tag{3.19}$$

This shows clearly that the stiffness is effectively a ratio of the superadiabaticity of the convective region to the subadiabaticity of the radiative region. For fixed convection zone properties, a small value of the stiffness (less than unity) indicates only weak resistance to convective overshoot. A larger stiffness (of order 100) indicates strong resistance to convective overshoot. The simulations of Brummell *et al.* (2002) span a stiffness of 0.5 to 30.

The primary output of the simulations of Brummell *et al.* (2002) is the penetration depth, l_{osht}, here called the overshoot length to avoid confusion with penetrative convection. Brummell *et al.* (2002) measured the overshoot length as the vertical distance between the nominal convective–radiative boundary and the position where the horizontally averaged flux of kinetic energy drops to one percent of its maximum. This identifies a region within the radiative zone where fluid motion is substantial (i.e., comparable to convection). The overshoot length depends on many of the input parameters to these simulations.

Brummell *et al.* (2002) explored the effects of changing the strengths of viscosity and thermal diffusion. The fundamental parameter for the strength of viscosity is the Prandtl number, the ratio of the kinematic viscosity to the thermal diffusion coefficient. They found a weak scaling with the Prandtl number of $l_{\rm osht} \propto {\rm Pr}^{0.05}$. As was the case with semi-convection, this implies that as the relevant length scales become much larger than the viscous scale, the increased turbulent motions provide additional mixing. On the other hand, they described the dependence on the thermal diffusion coefficient through two non-dimensional numbers: the Péclet number, the ratio of the thermal timescale to the advective timescale; and a parameter C_k, the ratio of the sound crossing time to the thermal timescale. For substantially subsonic convection—as is usually the case in stellar interiors—the Péclet number is the more reasonable choice for parametric dependence. The Péclet number dependence found by Brummell *et al.* (2002) is $l_{\rm osht} \propto {\rm Pe}^{-\frac{1}{2}}$ for large Pe (> 100); for lower Pe, $l_{\rm osht}$ is roughly constant. This implies that the overshoot length depends directly on the thermal coefficient, which makes physical sense. In a case with no thermal diffusion, plumes entering the radiative zone experience a strong buoyancy force returning them to the convection zone quickly. Conversely, strong thermal diffusion results in a plume equilibrated with its surroundings, resulting in no buoyancy force and therefore in the plumes traveling further.

The overshoot length also depends inversely on the stiffness to a weak power, which is consistent with non-penetrative pictures of convective overshoot (Zahn, 2002). Brummell *et al.* (2002) found in their simulations that the overshoot length scales as $S^{-\frac{1}{4}}$. The stiffness parameter, which was introduced by Hurlburt *et al.* (1994), can affect l_{osht} in two analytically derivable scalings, depending on the assumptions made in the derivation. The scaling of l_{osht} is roughly S^{-1} if the kinetic energy flux is assumed to be substantial and $\mathrm{Pe} \gg 1$ in the overshoot region; physically, this describes penetrative convection, where the overshooting region is mixed to be nearly adiabatically stratified. On the other hand, if Pe is of order unity in the overshoot region (i.e., the advective time is comparable to the thermal time) within the overshooting region, this scaling is instead $l_{\mathrm{osht}} \propto S^{-\frac{1}{4}}$, which is the scaling seen in the simulations of Brummell *et al.* (2002).

Combining these results into a single expression for the overshoot length as a function of the fundamental parameters of the system, we obtain

$$l_{\mathrm{osht}} = \lambda \alpha_{\mathrm{osht}} \mathrm{Pe}^{-\frac{1}{2}} \mathrm{S}^{-\frac{1}{4}} \mathrm{L}, \tag{3.20}$$

where L is the vertical extent of the convection zone, α_{osht} is a constant calibrated to the simulations of Brummell *et al.* (2002), and λ is a free parameter. The Péclet number, Pe, is defined to be UL/κ_T, where U is a characteristic velocity, taken to be the average root mean square velocity of convective motion. The parameter λ is included to provide the user with direct control over the model. A value of $\lambda = 1$ reproduces the exact results from Brummell *et al.* (2002) for a given stiffness and

Péclet number. Together, this provides a length scale of convective overshoot, but this is not enough to construct a complete mixing prescription.

Most modern implementations of convective overshoot create a diffusion coefficient dimensionally from a length scale and a velocity:

$$\kappa_{\mu,\mathrm{eff}}\left(r\right) = \frac{1}{3} l_{\mathrm{osht}} v_{\mathrm{osht}}\left(r\right), \tag{3.21}$$

where the relevant length scale is the overshoot length, and the velocity of the convective overshoot is assumed to decay exponentially with radius, as is seen in simulations (Freytag *et al.*, 1996; Brummell *et al.*, 2002).

The scaling of the overshoot velocity can be extracted from the simulations of Brummell *et al.* (2002). Recall that l_{osht} is defined as the distance from the convective–radiative boundary where the horizontally-averaged kinetic flux falls to 1% of its maximum. Assuming that the horizontal and vertical kinetic energies are roughly in equipartition, the kinetic energy flux is w^3, the cube of the vertical velocity. Thus, at l_{osht}, the vertical velocity has fallen to a $\sigma \equiv 0.01^{\frac{1}{3}}$ fraction of the maximum characteristic velocity of the convection zone. Assuming an exponential decay from the convective velocity given by mixing length theory (e.g., $v_{\mathrm{conv}}^2 = g\rho\left(\nabla - \nabla_{\mathrm{ad}}\right)\alpha_{\mathrm{ml}}^2 H_p/8$, where $\alpha_{\mathrm{ml}} H_p$ is the mixing length), the velocity has the form

$$v_{\mathrm{osht}}\left(r\right) = v_{\mathrm{conv}} \sigma^{-\frac{\left|r-r_b\right|}{l_{\mathrm{osht}}}}, \tag{3.22}$$

where r_b is the radius of the radiative–convective boundary.

A point of concern might be regarding the seeming inverse dependence of $\kappa_{\mu,\mathrm{eff}}$

on the strength of convective motion. The concern is that since Pe is proportional to the convective velocity, the overshoot length decreases as the convective velocity increases. Note, however, that Equation 3.21 also depends of S. In order to increase the convective velocity and maintain the same radiative region, the stiffness parameter must also decrease, which increases $l_{\rm osht}$, and preliminary analytic analysis shows that this is the dominant effect.

There are many caveats with this choice of model for convective overshoot though. This work assumes that overshooting convection above and undershooting convection below a convection zone behave comparably, which is certainly incorrect (e.g., Hurlburt *et al.*, 1986). Any major effects due to compositional changes, as can occur between shells of material in the cores of stars, are not modeled. These models also likely do not apply for transonic convection or for any kind of convection where geometric effects are of substantial importance. The role of degeneracy has also not been explored.

3.4.1 Implementation into Stellar Evolution Codes

A choice must be made to determine the appropriate length scale, L, from Equation 3.20 when applying this model to stellar evolution codes. One possible option would be to use the actual width of the convection zone. Such an implementation would be difficult to compare with existing work but could have interesting consequences and would be worth exploring. It is difficult to implement numerically, however, because most stellar codes are purely local. Another option is to express the extent of the convection zone in the simulations of Brummell *et al.* (2002) in terms of pressure scale heights, which is what is done here. This allows the calculation of convective overshoot primarily in terms of parameters relatively local to the convective overshoot region. It is not clear from the simulations of Brummell *et al.* (2002) whether the appropriate overshoot length unit is convective zone extent or pressure scale height, so for simplicity of comparison with previous models, the latter is used. In Brummell *et al.* (2002), the vertical extent of the convection zone is approximately two pressure scale heights; thus, $L = 2H_p$ for the purposes of this work.

In practice, several complications arise for some extreme values of the input parameters. Rarely, for convective–radiative boundaries with very low stiffness as can occur in massive star envelopes, the overshoot length can become comparable to a pressure scale height. This usually does not pose any problems numerically; however, if the pressure scale height changes substantially across the overshooting region (e.g., when crossing from the convective envelope into the core), this must be accounted for. Physically, in such a circumstance, any ballistic plumes attempting to pass such a barrier would experience substantial resistance. In the code, this is handled by separating the extent L from the expression of the overshoot length, $\hat{l_{\mathrm{osht}}} \equiv \lambda \alpha_{\mathrm{osht}} \mathrm{Pe}^{-\frac{1}{2}} \mathrm{S}^{-\frac{1}{4}}$. Because $L = 2H_p$ is no longer assumed to be evaluated at the boundary, it is written as

$L(r)$. Equation 3.22 then becomes

$$\Gamma(r) = \left| \int_{r_b}^{r} \frac{dr}{L(r)} \right| \tag{3.23}$$

$$v_{\text{osht}}(r) = v_{\text{conv}} \sigma^{-\frac{\Gamma(r)}{\hat{l}_{\text{osht}}}}. \tag{3.24}$$

Effectively, this restricts the distance that can be traveled to be in units of cumulative pressure scale heights rather than in physical distance.

In addition, weak convection zones can have relatively small values of Pe. The simulations of Brummell *et al.* (2002) show that l_{osht} for low Pe simulations do not obey the same $-1/2$ power. This can be remedied by replacing $\text{Pe}^{-\frac{1}{2}}$ with $(\text{Pe} + \text{Pe}_0)^{-\frac{1}{2}}$, where Pe_0 is fit to be roughly 150 from the low Pe simulations of Brummell *et al.* (2002).[1] The exact expression used is then

$$l_{\text{osht}} = \lambda \alpha_{\text{osht}} (\text{Pe} + \text{Pe}_0)^{-\frac{1}{2}} S^{-\frac{1}{4}} 2H_p, \tag{3.25}$$

where $\text{Pe}_0 = 150$ and $\alpha_{\text{osht}} = (196 + 150)^{\frac{1}{2}} 7^{\frac{1}{4}}/0.4 \approx 76$ for $S = 7$, $\text{Pe} = 196$, $l_{\text{osht}} = 0.40L$, the parameters of Case 5.

3.5 Simulations

Using the new implementations for semi-convection and overshoot, described in Sections 3.3 and 3.4, a detailed grid of stellar models was constructed that spans the

[1]This is slightly different from the prescription used in the results of this study. The actual prescription relevant to the results here is included in Appendix .1 and is somewhat nonphysical. Preliminary simulations using the new model do not show substantial variations from the old simulations. The old model has been superseded by the model described here.

relevant parameter space. From these models, several diagnostics were measured at various stages of the evolution. The conditions used to obtain the core masses are defined as the mass coordinate (the mass enclosed at a certain radius) at which the condition given in Table 3.1 is met. The time stages of the star mentioned are defined as the first time when the condition given in Table 3.2 is met at the specified location. The simulations spanned overshooting multipliers (λ) from 0.1 up to 1.0 in bins of 0.1 and semi-convective exponents (m) ranging from 1/8 to 32 in powers of two.

3.5.1 Limits

The parameterization of convective overshoot is a multiplier on the depth of penetration, calibrated to be 1.0 at the depth predicted by the simulations of Brummell *et al.* (2002) ignoring any potential effects of viscosity scaling. In their simulations, the convective region was held at a given superadiabaticity; whereas in stars, the convective region will soon become nearly adiabatic, likely weakening the strength of convection. Taking into account the noted dependence of the overshoot depth with Pr, which is very weak ($l_{\mathrm{osht}} \sim \mathrm{Pr}^{0.05}$), the recommended value would be closer to 0.6 for stellar Pr $\sim 10^{-6}$. The clear lower limit is 0.0; however, in practice the helium core bifurcates unphysically around a value of 0.1, which is discussed in greater detail in Section 3.6.2.

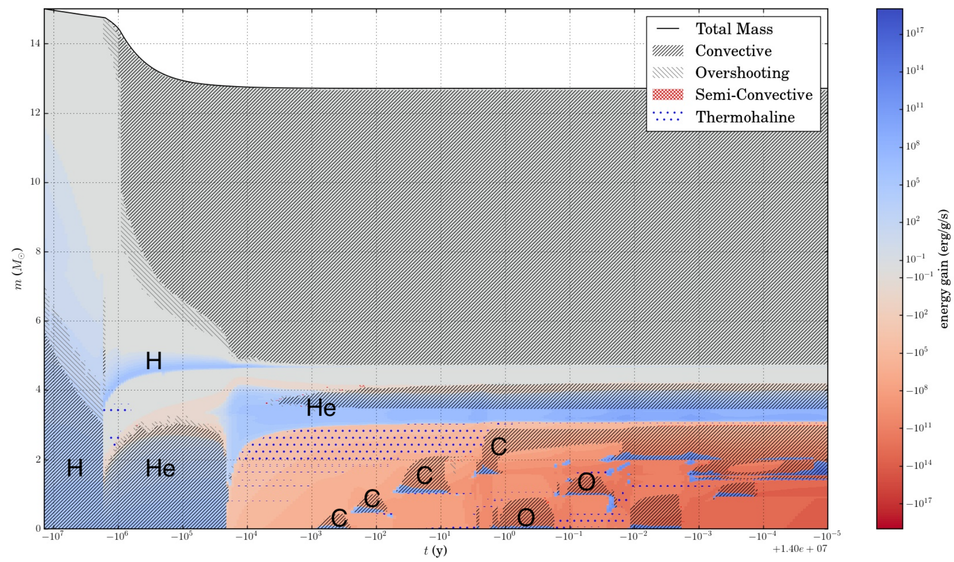


Figure 3.1: A Kippenhahn diagram of the recommended model of zero age main sequence mass 15 $M_{\odot}$. The abscissa is time until death and is plotted logarithmically. The ordinate is the mass coordinate of the star. Energy generation is plotted in blue, and energy losses are in red. Convection is given by dense, black hashing; overshoot by sparse, gray hashing; semi-convection by red crosshatching; and thermohaline convection by blue dots. The overshoot multiplier in this simulation is 0.6, and the semi-convective exponent is 0.5, i.e. strong; however, the strength of semi-convection does not strongly affect the outcome of this model. For reference, the primary fuels of most major burning regions up until silicon ignition are labeled. This star immediately becomes a red supergiant upon exhausting hydrogen in its core and remains that way for the remainder of its life.

3.5.2 Convergence Tests

The models were checked for substantial variations with resolution to ensure that the simulations are converged. In this study, doubling the resolution changed the helium, carbon–oxygen, and neon core masses by less than 2%.

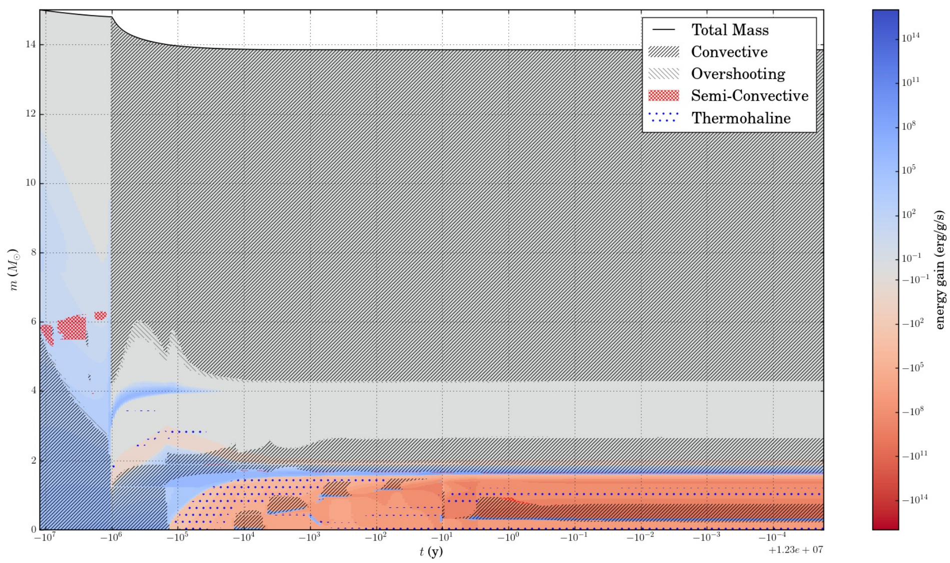


Figure 3.2: The evolution of a 15 $M_\odot$ star with weak overshoot (multiplier of 0.1) and semi-convection (exponent of 32) plotted up to oxygen ignition. The plot description is given in Figure 3.1. In this simulation, the helium convective core has bifurcated; note the small separation in the hashing. This has effectively ignited a convective helium shell immediately above the convective helium core. This setup is unlikely to be physical as these interfaces are likely to be quite turbulent, and so this is likely a numerical artifact. The end result of this bifurcation is that the convective core effectively runs out of fuel earlier than if the core hadn't bifurcated, leading to a much smaller carbon–oxygen core than would normally be predicted for a given helium core mass.

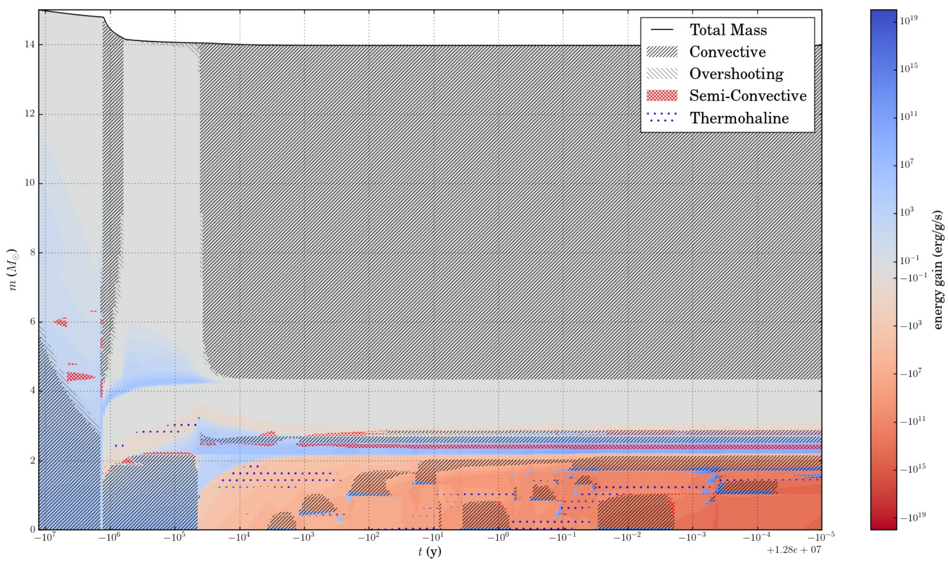


Figure 3.3: The evolution of a 15 $M_\odot$ star with weak overshoot (multiplier of 0.2) and semi-convection (exponent of 16); however, both processes are stronger than in Figure 3.2. The plot description is given in Figure 3.1. Similar to Figure 3.2, this star has undergone a core bifurcation; however, in this case the helium burning shell is not convective. It seems unlikely that there would be no entrainment whatsoever of this helium burning shell, particularly as its opacity increases as the helium fuses to carbon and oxygen (see, e.g. Castellani *et al.*, 1985) and the helium burning shell becomes more convectively unstable. However, this case is not as physically unlikely as that of Figure 3.2. This simulation is an example of one that undergoes a blue loop, as can be seen by the convective envelope receding during helium burning.

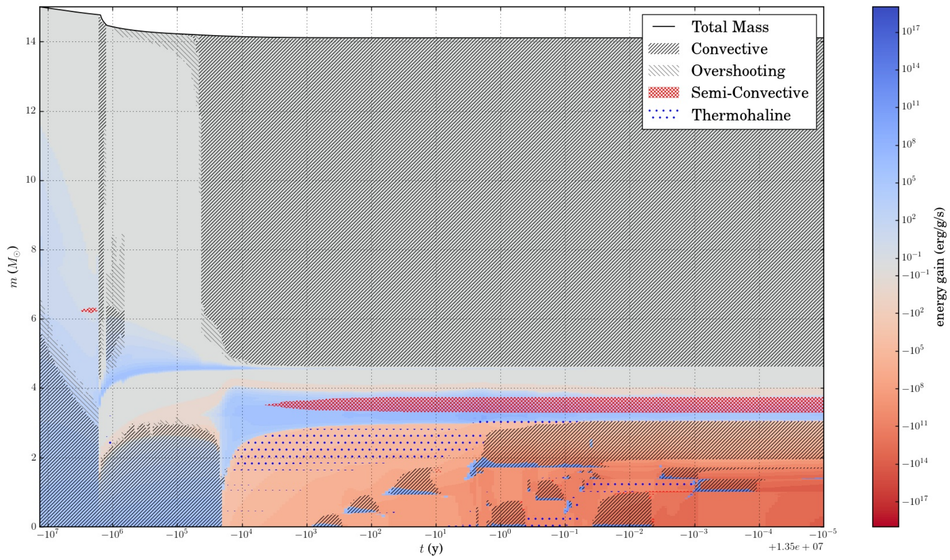


Figure 3.4: The evolution of a 15 $M_\odot$ star with moderate overshoot (multiplier of 0.4) and weak semi-convection (exponent of 16). The plot description is given in Figure 3.1. This star also undergoes a blue loop, but doesn't exhibit the same bifurcating helium core as in Figure 3.3. Also note the existence of a convective hydrogen burning shell when the star becomes a blue supergiant.

3.5.3 Overall Evolution

Figures 3.1 through 3.4 illustrate the four main possibilities observed in this survey of parameter space. The recommended model is given in Figure 3.1, with a convective overshoot multiplier of 0.6. This star becomes a red supergiant upon exhausting the hydrogen in its core and remains so until core-collapse. Most of the simulations in this sample appear qualitatively like this one, though the total core masses of various elements can depend substantially on the parameters chosen for convective overshoot and semi-convection.

Some of the simulations exhibited a unphysical bifurcation of their helium core. Figure 3.2 and Figure 3.3 illustrate the two realizations of this bifurcation. In Figure

Table 3.1: Core Mass Definitions

Core	Definition
Helium	$X_{\mathrm{H}} > 0.1$
Carbon–Oxygen	$X_{\mathrm{He}} > 0.2$
Neon	$X_{\mathrm{C}} > 0.1$

3.2 the outer component of the convective core bifurcates into a convective helium-burning shell. However, in Figure 3.3, the outer component of the core becomes a radiative helium-burning shell. A sharp transition develops between the inner convective core and the burning shell (usually no more than a few zones). Multi-dimensional simulations of convection, such as those of Meakin and Arnett (2007) indicate strong turbulence at convective boundaries, and even in cases with minimal convective overshoot, convection zones launch substantial gravity waves that blur these interfaces. It is thus unlikely that such strong interfaces would persist in the turbulent interiors of stars.

The last possibility, which requires moderate overshooting and weak semi-convection, is shown in Figure 3.4. These simulations exhibit a blue loop, and the fraction of time spent on this loop depends on the helium core mass of the star, as will be shown.

Table 3.2: Stellar Stage Definitions

Stage	Condition	Location
Hydrogen Ignition	$X_{\rm H} - X_{\rm H,0} > 0.01$	Center
Hydrogen Depletion	$X_{\rm H} < 0.01$	Center
Helium Ignition	$X_{\rm C} > 0.01$	Center
Helium Depletion	$X_{\rm He} < 0.01$	Center
Carbon Ignition	$T > 5. \times 10^8$K	Center
Carbon Depletion	$T > 1.2 \times 10^9$K	Center
Oxygen Depletion	$X_{\rm O} < 0.05$	Center
Presupernova	$v > 9. \times 10^7$cm/s	Anywhere

Table 3.3: Typical Convection Values of Pe

	H Ign.	H Dep.	He Ign.	He Dep.	C Ign.	C Dep.	O Dep.
H Burn	2.5E7	1.8E7	13.1	58.5	63.4	55.7	55.7
He Burn			5.0E8	8.3E8	–	3.9E8	4.5E8
C Burn					–	–	5.0E8
O Burn							2.6E10

Table 3.4: Typical Values of S Over (Ov) and Under (Un) Convection Zone

	H Ign.	H Dep.	He Ign.	He Dep.	C Ign.	C Dep.	O Dep.
Ov H Burn	35.2	38.0	–	–	–	–	–
Un H Burn			9.2	1.8	1.7	2.5	2.4
Ov He Burn			12.4	99.1	–	8.1	4.7
Un He Burn					–	26.9	17.3
Ov C Burn					–	–	97.5
Un C Burn							1.1
Ov O Burn							63.6

Table 3.5: Typical Convection Values of $\Delta r/r$

	H Ign.	H Dep.	He Ign.	He Dep.	C Ign.	C Dep.	O Dep.
H Burn	1.0	1.0	0.76	0.96	0.99	0.996	0.996
He Burn			1.0	1.0	–	0.61	0.62
C Burn					–	–	0.43
O Burn							1.0

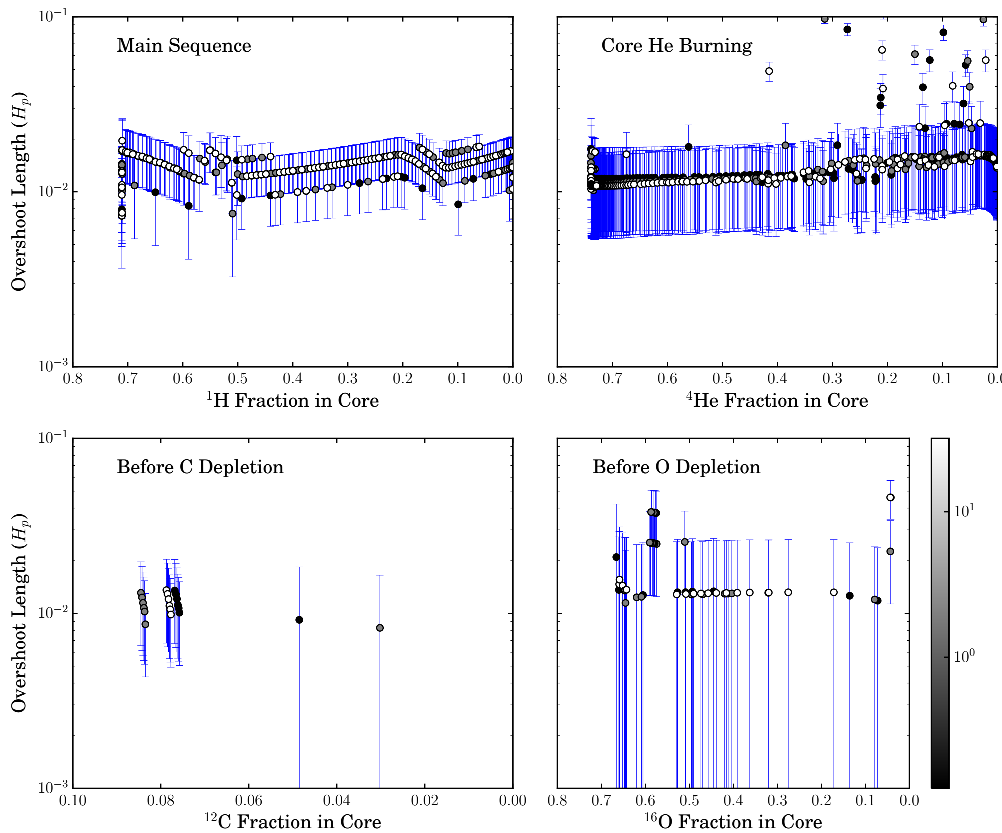


Figure 3.5: The overshooting extent from the convective core at various stages for some characteristic models. The overshoot factor here is exactly 0.6, which is the predicted value based on viscous extrapolation from Brummell *et al.* (2002). The semi-convective power is given by shading—black points have strong semi-convection with a semi-convective power of 1/8, gray points have moderate semi-convection with a power of 2, and white points have weak semi-convection with a power of 32. These are plotted as functions of an effective time coordinate for different states of the stellar lifetime, using the depletion of hydrogen on the main sequence (top left), the depletion of helium on the helium core burning phase (top right), the depletion of carbon prior to carbon core depletion (bottom left), and the depletion of oxygen prior to oxygen core depletion (bottom right). The errorbars indicate the grid resolution by indicating the predicted change if the edge of the overshooting region were one zone further or closer from the convection zone. The variation in the overshoot extent from the core changes only slightly from phase to phase in units of the pressure scale height. The core overshoot for both the main sequence and helium stage is well approximated as $10^{(-1.84\pm0.22)}H_p$ (excepting when the convective extent is minimal or when the star is undergoing a transition, as near the end of the helium burning life).

Table 3.6: Typical Convection Values of η

	H Ign.	H Dep.	He Ign.	He Dep.	C Ign.	C Dep.	O Dep.
H Burn	-5.9	-6.4	-14.9	-14.0	-12.6	-25.6	-12.4
He Burn			-3.7	-4.0	–	-5.7	-4.7
C Burn					–	–	-2.4
O Burn							1.7

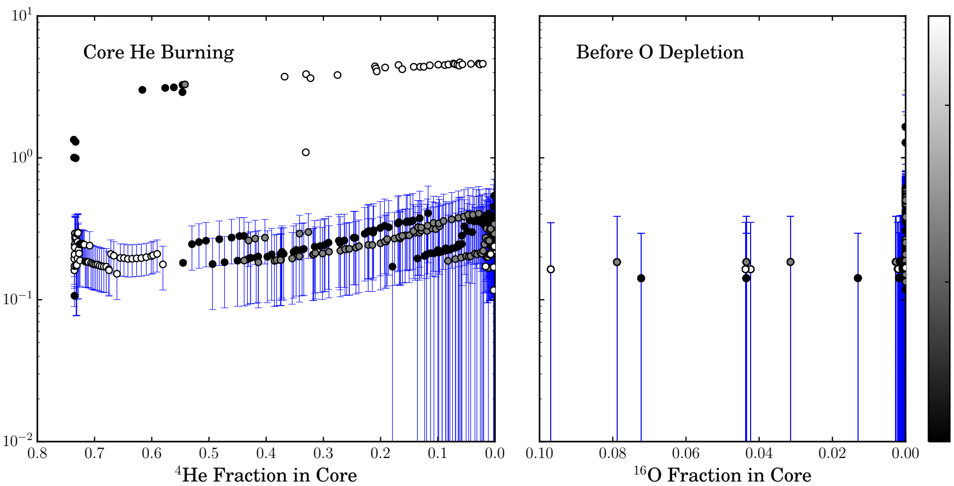


Figure 3.6: The undershooting extent from the convective envelope at various stages for some characteristic models in the same style and for the same simulations as in Figure 3.5. For this plot, only the core helium burning phase (left) and post helium depletion phase prior to oxygen depletion (right) are plotted. The undershooting extent depends strongly on whether the convection zone boundary is near the envelope-core transition, resulting in a larger stiffness and therefore a smaller undershooting extent, or whether the boundary is far out in the envelope, resulting in a smaller stiffness and a larger undershooting extent. This describes the difference between the lower sequence of points in the core helium burning phase and the higher sequence, respectively. The runs with large and moderate semi-convection start as blue supergiants with no outer convection zone. As the outer convection zone appears and works down to the envelope-core transition, the undershooting extent is substantial ($\sim 3H_p$) The variation in the overshoot extent from the core changes only slightly from phase to phase in units of the pressure scale height. The envelope undershoot extent is well approximated as $10^{(-0.66\pm0.36)}H_p$.

3.6 Discussion

3.6.1 Convective Overshoot

The evolution of the overshoot depth for the main sequence through core oxygen depletion is shown in Figure 3.5. The overshoot depth is measured at the location where the diffusion coefficient drops by a factor of e; the dependence of this length scale on the various properties of this system is not strongly affected by the choice of this factor. The main sequence core overshooting length increases from effectively no overshoot in previous KEPLER models to around 1.4% of a pressure scale height. The undershooting length below the outer envelope in the later phases extends to 22% of a pressure scale height on average, which has substantial consequences for surface abundances. For those wanting to take advantage of this work—using, for example, the model from Freytag *et al.* (1996)—main sequence core overshoot should be taken to be $10^{(-1.84\pm0.22)}H_p$ and envelope overshoot, $10^{(-0.66\pm0.36)}H_p$.

These results should be taken with several caveats in mind. Tables 3.3 through 3.6 list typical values of various non-dimensional numbers to help determine which approximations from Brummell *et al.* (2002) are most appropriate to massive star interiors. The organization of these tables are as follows: the columns each correspond to a stellar stage of evolution as defined in Table 3.2. Each stage of the stars discussed here has at most one major convection zone for each burning region of the star (including both the core and any shells), and the rows correspond to the burning region associated with the convection zone. A blank cell indicates that no such burning region exists at that stage of stellar evolution (e.g., no oxygen burning at hydrogen ignition). A cell with a dash (–) indicates that though the burning region exists, no associated convection zone was present at that stage.

From Table 3.3, the values of Pe vary substantially throughout the star, but typically take values between 10^7 and 10^{10} for the core and shells and are much reduced (~ 10) in a convective envelope. The simulations of Brummell *et al.* (2002) have Pe ~ 100 and so are appropriate for envelope but unlikely to represent convection well in the interiors of these stars. It is likely that overshooting convection will become increasingly penetrative as Pe continues to increase, but this has not yet been observed in direct numerical simulations.

Table 3.4 lists typical values of the stiffness, which vary from 1 to 100. In the table, for each convection zone, a stiffness is given for each possible overshooting region (i.e., above and below). Typically, the stiffness below shell convection can be relatively large due to the change in density across the burning shell. Such does not appear to be the case for the carbon–oxygen shell, whose lower convective bound is well above the carbon-burning region. These numbers are typical of the simulations of Brummell *et al.* (2002).

Table 3.5 attempts to address the importance of spherical effects of convection by measuring the difference in radius between the outer and inner extent of the convection zone (Δr) and dividing by the outer extent. For convection that spans an entire sphere (i.e., core convection), $\Delta r/r = 1$, and for a thin shell, $\Delta r/r \ll 1$. The Sun, for example, has $\Delta r/r = 0.3$. The simulations of Brummell *et al.* (2002) are only appropriate for small values of this parameter (< 0.3 or possibly less), and so it is unlikely that spherical effects can be ignored in the contexts of massive stars.

Table 3.6 also shows the degeneracy parameter η, which is a renormalization of the electron chemical potential (see, e.g., Timmes and Arnett, 1999). Degenerate effects become important when this number exceeds 0. This does happen in the deep interiors and late stages of massive star evolution, and so could be a relevant process for some convection zones in stars.

Comparison with Meakin and Arnett

In order to compare this quantitatively with the expression from Meakin and Arnett (2007), their mass entrainment must be converted into a diffusion coefficient:

$$\frac{\partial M_{\mathrm{conv}}}{\partial t} \propto 4\pi r^2 \rho v_{\mathrm{conv}} \mathrm{Ri}_{\mathrm{b}}^{-n}, \tag{3.26}$$

where n is the exponential dependence, probably between 1 and 1.75. This mass entrainment effectively depends on one non-dimensional number, the bulk Richardson number, Ri_{b}, defined here as

$$\mathrm{Ri}_{\mathrm{b}} = \frac{\Delta b L_t}{U_{\mathrm{rms}}^2}, \tag{3.27}$$

Where Δb is the “buoyancy jump across the interface,” L_t is a length scale characteristic of the turbulence, and U_{rms} is the root-mean-square velocity of the turbulence.

For ease of comparison, this last quantity is taken to be of order the root-mean-square convective velocity.

To construct a diffusion coefficient, a velocity (in this case, the entrainment velocity, $\frac{\partial M_{\mathrm{conv}}}{\partial t}\frac{1}{4\pi r^2 \rho}$) and a length scale are required. The characteristic length scale is unclear, but is probably related either to L_t or to H_p. Assuming that L_t is the relevant length scale, comparison to Equation 3.21 is straightforward:

$$D_{\mathrm{osht}} \propto v_{\mathrm{conv}} L_t \mathrm{Ri}_{\mathrm{b}}^{-n}. \tag{3.28}$$

For this, the comparison between the two models is clear in that the differences are solely due to the length scale of the two. In the Meakin and Arnett (2007) scaling, the characteristic length is $L_t \mathrm{Ri}_{\mathrm{b}}^{-n}$; whereas in the Brummell *et al.* (2002), the characteristic length is $H_p \mathrm{Pe}^{-1/2} \mathrm{S}^{-1/4}$.

Though it is nontrivial to evaluate Ri_{b} in terms of S, it can be shown that $\mathrm{Ri}_{\mathrm{b}} = 2\,(1+S)$. Thus, in the case of $n = 1$, the length scale from Meakin and Arnett (2007) is roughly S^{-1}. Such a scaling is not surprising, as this is the expected scaling of penetrative convection (Hurlburt *et al.*, 1994). The scaling from Brummell *et al.* (2002) is consistent instead with that of non-penetrative convective overshoot. Whether convective overshoot in stars is penetrative or not or some combination thereof continues to be a subject of much debate (e.g., Viallet *et al.*, 2015). It remains quite possible, as is discussed in Viallet *et al.* (2015), that convective boundaries could behave as convective overshoot (or “diffusion-dominated” in their paper) for small Pe and as penetrative convection for large Pe. No three-dimensional compressible direct numerical simulations of convective boundaries have yet seen penetration. Granted, direct numerical simulations cannot explore as extreme values of Pe as a large eddy simulation, and these calculations may present the most physical view of such processes. Such a possibility of penetrative convection is left to a future publication.

3.6.2 Core Masses

Helium Core and Compactness

Strong convective overshoot increases the size of the helium core, which is illustrated in the first plot of Figure 3.7. For extremely efficient convective overshoot, the convectively burning hydrogen shell begins to penetrate into the helium core, reducing its size. The recommended convective overshoot from Brummell *et al.* (2002) (i.e., after taking into account the scaling due to viscous effects) yields helium cores up to 12% larger than comparable runs with the previous overshooting prescription in KEPLER.

The effect of semi-convection on the helium core mass is less extreme. A semi-convection exponent greater than four increases the mass of the helium core just following helium ignition, on average. This increase is due to reduced mixing of the convective hydrogen shell: inefficient semi-convection mixes less fresh material into the hydrogen shell, so the star burns further in mass coordinate to use the same amount of fuel, leading to a larger initial core mass. Moore and Garaud (2016) found

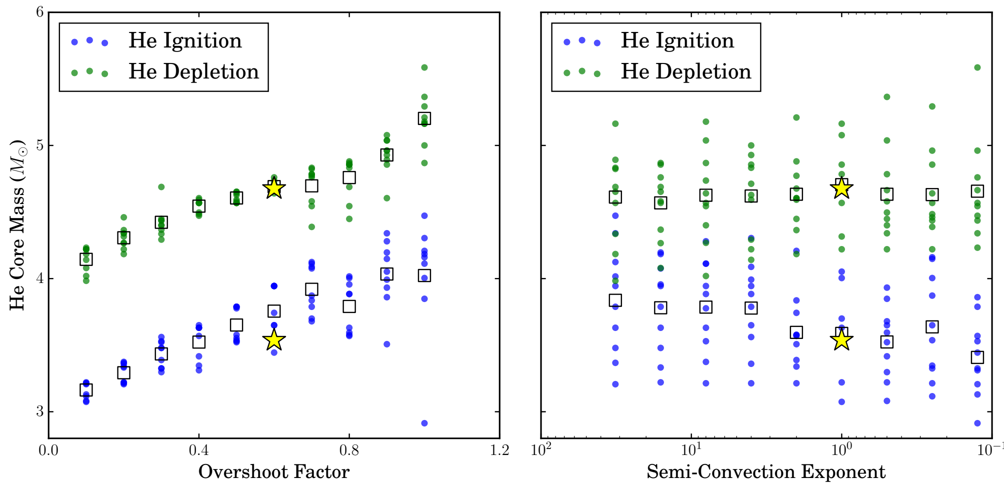


Figure 3.7: Each circle indicates the helium core mass for a choice of semi-convection and convective overshoot measured at helium core ignition (blue) and depletion (green). The black boxes mark the averages, binned according to the quantities of each abscissa. These boxes should be interpreted as illustrating the trend of varying the quantity on the abscissa while holding the other process fixed. The circles are present to indicate the range of possible results for a given choice of one parameterization while varying the other and should not be thought of as an uncertainty in the results. It is particularly clear from this that convective overshoot has a much stronger potential impact than semi-convection. The size of the helium core depends strongly and mostly linearly on the factor of the overshoot length. The deviation from this linear trend for large overshoot occurs due to the convection zone outside of the helium core forcing hydrogen into the helium core. The dependence on semi-convection is more subtle but mostly has to do with the efficiency of mixing as the star enters the core helium burning phase. Efficient semi-convection more uniformly mixes the material that will become the helium core, which leads to slower burning through that region and therefore a smaller initial core. However, all of that material is eventually burned, so there is little long effect on the helium core mass.

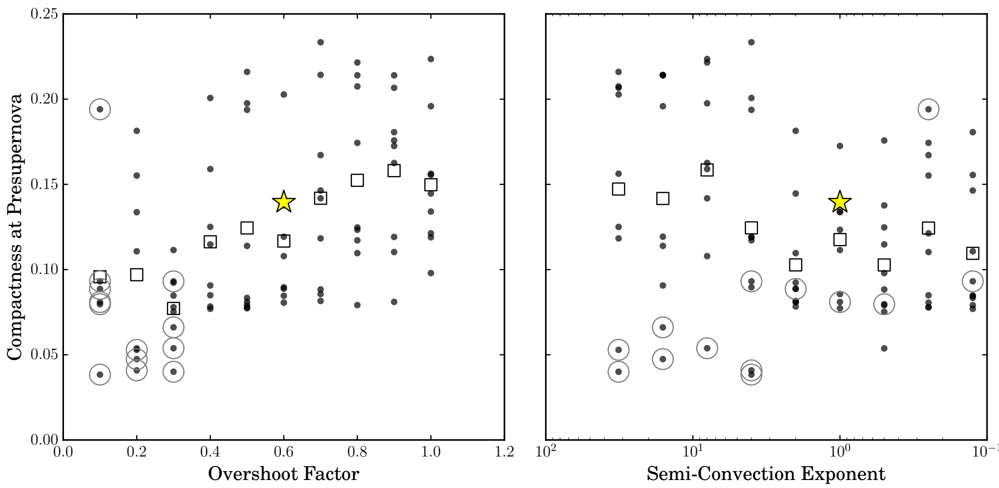


Figure 3.8: The effects of convective overshoot and semi-convection on the final compactness of the star in the same style as in Figure 3.7. The simulations are categorized according to whether the helium core bifurcates or not, as described in Figures 3.2 and 3.3, by encircling bifurcated cores in gray. This quantity traces relatively well with the initial helium core mass. There is clearly strong variation of compactness both with overshooting convection and semi-convection; however, it should be noted that all of these models would likely explode as they have relatively low compactness (see e.g. Sukhbold and Woosley, 2014), so whether this variation is meaningful is not clear.

a similar duality of semi-convection and concluded that this dual behavior is indistinguishable from the choice of the Ledoux and Schwarzschild criteria for convection. At later times, the shell burns past the initial mixed zone, leading to a final core mass that is insensitive to the strength of semi-convection.

The compactness also depends on convective overshoot and semi-convection in an analogous way as the initial helium core mass, as seen in Figure 3.8. This plot shows substantially more scatter across both semi-convection and convective overshoot and at least part of this is due to the inherently chaotic nature of the compactness, as is discussed in Sukhbold and Woosley (2014). Nevertheless, trends do appear in the compactness when binned by overshooting and semi-convective parameters. The compactness grows linearly with overshoot, analogous to the helium core mass. The abrupt change at the semi-convective break is also present, likely due to the initial core mass setting the initial size of the convective core. As the core evolves, increasing amounts of carbon and oxygen increase the free-free opacity of the core, inhibiting the convection zone growth. Thus, the initial core mass proves to be of substantial consequence for the later stages.

Carbon/Oxygen Core

As mentioned in Section 3.6.2, the helium core mass generally determines the late-stage behavior of these stars, including the final carbon–oxygen core mass. This dependence is plotted for various codes, initial stellar mass, and convective overshoot strength in Figure 3.9. An increase in the strength of convective overshoot is degenerate with an increase in the initial stellar mass (given identical physics) for both KEPLER and MESA. The carbon–oxygen core relates almost linearly with the helium core for changing either quantity. This relation, however, is not the same between different stellar evolution codes, and it is unlikely that convective overshoot is a major contributor to these differences, based on these effects.

This relation can also be seen in Figure 3.10, which also illustrates the effect of semi-convection. Some of the simulations do not fall along this trend, and many of these have extremely weak overshooting (factor of less than 0.3) and semi-convection

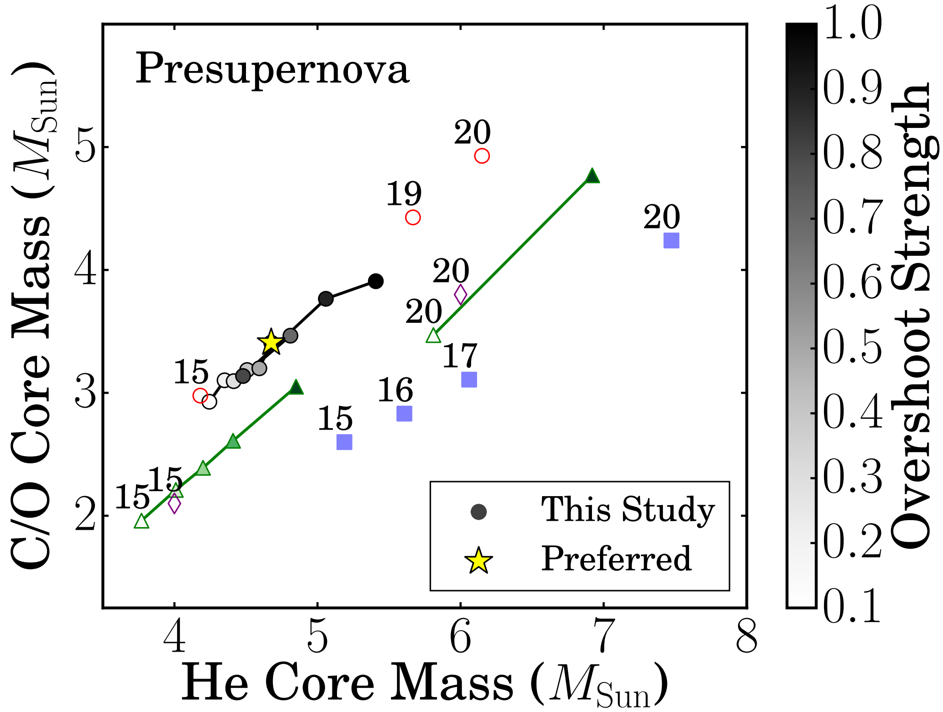


Figure 3.9: The final carbon–oxygen core masses of non-rotating stars plotted against the final helium core mass for various codes. Each code is given by a different symbol, and runs of the same mass with the same code are connected by lines and labeled. A moderate value for semi-convection of $m = 0.5$ is presented; however, the general results hold for other values of semi-convection. Strength of convective overshoot is indicated by shading, with shaded symbols having stronger convective overshoot. The star indicates the suggested overshooting factor of 0.6. Three models from Woosley and Heger (2007) are included as red circles to illustrate that increasing the strength of convective overshoot is analogous to increasing the main sequence mass. Similar simulations using MESA from Sukhbold and Woosley (2014) are included in green triangles, taken from their Table 2. The shaded green triangles indicate the inclusion of overshoot up to $f = 0.025$; note that the magnitude and direction is comparable to these results. Simulations from Limongi and Chieffi (2010) are presented as blue squares from their NL00 series, taken from their Figure 7. That each code demonstrates substantially different relations is concerning but not new. Drastic changes in semi-convection and convective overshoot do not appear to move the results off the C/O to He relation unique to each code. It is likely that a similar implementation of these processes in these other codes would have analogous results, though the extent to which this would occur is difficult to predict.

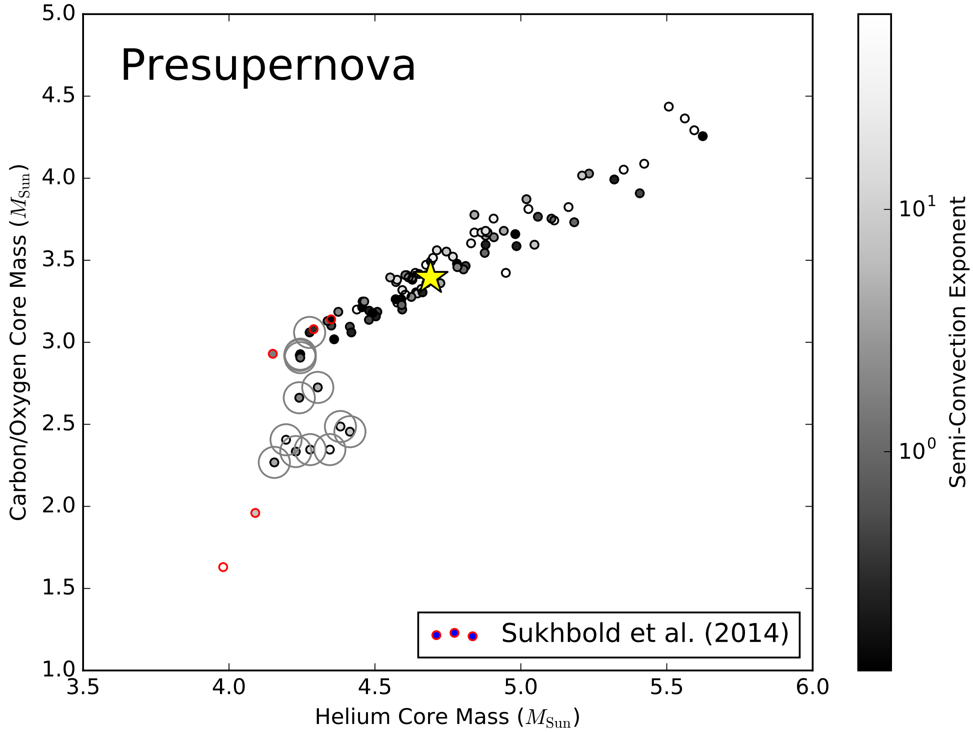


Figure 3.10: The circles on this plot are the same as in Figure 3.9 but with the shading of the points corresponding to the strength of semi-convection. Here, all possible combinations of semi-convection and overshoot have been included. These are categorized according to whether the helium core bifurcates or not, as described in Figures 3.2 and 3.3, by encircling bifurcated cores in gray. The red circles correspond to the 15 solar mass star KEPLER runs from Sukhbold and Woosley (2014) with changing strength of semi-convection (here the shading is intended to be a suggestive rather than quantitative measure of semi-convective strength). The strength of convective overshoot is not illustrated here, but it correlates very strongly with the mass of the He core. As in Figure 3.9, the stars indicate the recommended parameter regime. It is clear that two distinct populations exist in this diagram: one that falls along a clear core mass relation that is consistent with previous KEPLER models and one that falls off this relation. These simulations have the peculiar property of having bifurcated helium cores; that is, the helium core is comprised of a convective helium burning core and a hydrogen burning shell that are separated by a few numerical cells. Because stellar convection is a strongly turbulent process, and the boundary between these two zones contains a fraction of stellar material, it is unlikely that this is a physical effect. The only way to produce a small carbon–oxygen core with a large helium core is for the helium core to bifurcate, which calls into question many supernova 1987A models.

(exponent greater than 4). In these limits, approximately halfway through helium core burning, the helium core has bifurcated, as is discussed in Section 3.5. In such simulations, a helium shell has ignited immediately above (within several zones) a convective helium burning core. Because this behavior is unlikely in the turbulent interior of a star, this range of parameter space is likely prohibited. The only way to produce a small carbon–oxygen core with a large helium core is for the helium core to bifurcate, which calls into question many supernova 1987A models.

3.6.3 Envelope Structure

The envelope structure of these stars depends on the mixing effects of these minor mixing processes. In particular, convective overshoot and semi-convection substantially effect the time spent as red supergiants and on blue loops.

Red Supergiants

Minor mixing processes have substantial effects on the structure of the envelope in the advanced stages of these stars. The stars labeled as having weak semi-convection and convective overshoot behave quite differently from the others, and will be addressed in Section 3.6.3. The other stars clearly behave by spending more time as red supergiants as the helium core mass increases. This is not surprising as a denser core leads to a more extended envelope. This can be extended to indicate that the denser the core, the deeper the convective envelope extends, which is consistent with the re-

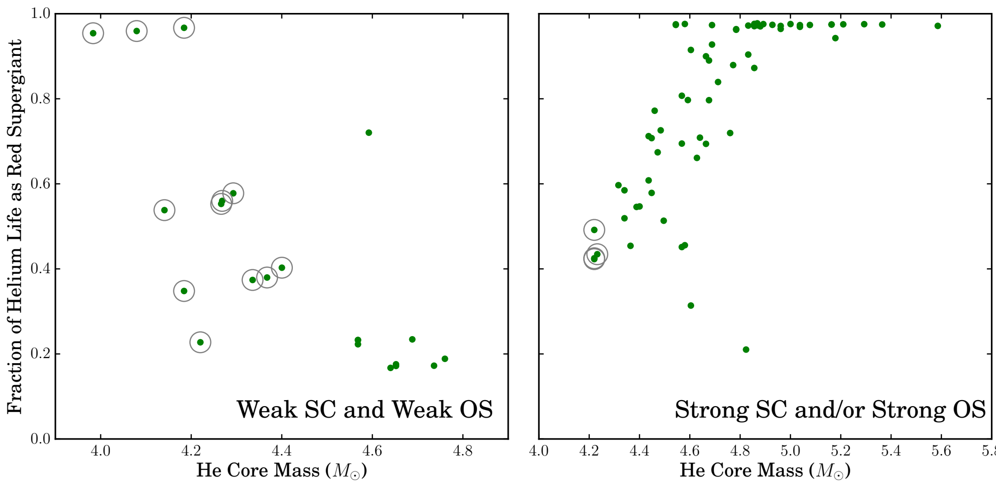


Figure 3.11: The time spent as a red supergiant star as a function of helium core mass. Each circle indicates a particular choice of semi-convection and overshooting convection. These are categorized according to the total mixing by semi-convection and overshoot according to Equation 3.29, where the left plot includes only runs with little mixing from semi-convection and overshoot together and the right plot to those with large amounts of overshoot, semi-convection, or both. As in Figure 3.10, runs with bifurcated helium cores are marked with gray circles. The duration of time spent as a red or blue supergiant is an observable diagnostic of stellar evolution as it can be quite sensitive to interior physics and can be measured simply by calculating the red to blue supergiant ratio (see e.g., Langer and Maeder, 1995). The fraction of time spent as a red supergiant can depend quite differently on the internal structure of the star depending on the strength of minor mixing. Weak mixing leads to the fraction being a decreasing function of helium core mass; whereas, strong mixing leads to an increasing function of core mass. The second major point is that most runs with weak minor mixing have bifurcated helium cores. Those stars with initially large helium cores become red supergiants more quickly, which is not surprising based on the relation in Lauterborn *et al.* (1971) that states that a stronger gravitational potential leads to a red supergiant solution. Those with large cores but small mixing tend to initially become red supergiants and mix a substantial amount of hydrogen deep into the star. The other half of the relation in Lauterborn *et al.* (1971) states that a lower mean molecular weight just outside the star core leads to a blue supergiant solution. This causes the star to enter a blue loop, decreasing the amount of time spent as a red supergiant.

sults of Lauterborn *et al.* (1971), who showed that the depth that the red supergiant extends depends on the gravitational potential of the core.

Thus, overshoot and semi-convection have consequences on the time spent as a red supergiant (and therefore on the red to blue supergiant ratio). This time is illustrated in Figure 3.11, which shows that stars with initially large helium cores become red supergiants more quickly. Strong overshoot leads to less time as a blue supergiant before transitioning to a red supergiant. Recall from Section 3.6.2 that low semi-convection leads to a strong initial helium core, which leads to strong red supergiant convection. The properties of the red to blue supergiant ratio on the whole remain elusive as only the effects on a 15 $M_\odot$ star have been investigated here.

Blue Loops

Almost all of the aforementioned stars with weak minor mixing undergo a blue loop on the helium core burning phase, i.e., they become blue supergiants after an initial red supergiant phase. This term is defined here to mean any simulations falling within both

$$\begin{aligned} \alpha &< \log m - 0.2, \text{ and} \\ \alpha &< 0.6. \end{aligned} \tag{3.29}$$

It is well known (Lauterborn *et al.*, 1971) that an increase in the metallicity near the hydrogen burning shell tends to make blue supergiants, so an initial red supergiant with deep convective penetration tends to become blue as low-mean-molecular weight material enters near the hydrogen burning shell. Because, as mentioned in Section 3.6.3, models with low semi-convection have initially vigorous convective envelopes, they easily mix down enough hydrogen to transition to being a blue supergiant. This explains the strange behavior of the weakly mixed cases in Figure 3.11, where stars with initially strong helium cores spend little time as red supergiants.

The early helium core burning lifetime has a short but important stage of convectve hydrogen shell burning, whose dynamics impact the evolution of the star. Before becoming either a red or blue supergiant, each star ignites its hydrogen shell. This shell burns convectively, and the strength of the mixing depends strongly on minor mixing processes. Strong semi-convective mixing tends to produce a uniform, high-metallicity region just outside the hydrogen shell; whereas, weak semi-convective mixing produces a sharp gradient in metallicity. If the hydrogen burning shell has high metallicity, as in the cases with strong mixing, the structure tends to have shallower convection, as mentioned above. Thus, only the weakly mixing stars form blue loops for this particular mass range.

Much work must be done regarding the later phases of these stars, but this discussion describes the effects of semi-convection and overshooting convection on the later stages of a fifteen solar mass star in detail. The next intended steps are to understand the effects of these processes on the envelopes in more depth and to investigate the effects across the entire mass range of massive stars.

3.7 Conclusions

New prescriptions for overshooting convection and semi-convection were implemented that are based on the results of numerical hydrodynamical simulations. Using the numerical results of Wood *et al.* (2013), a prescription for the compositional mixing of semi-convection has been implemented into KEPLER. Using the numerical results of Brummell *et al.* (2002), an expression is used for the overshoot depth as a function of the local variables. This is the first time that predictions for overshooting length in massive stars have been attempted based on the work of Brummell *et al.* (2002).

In this large suite of simulations of a 15 solar mass star, the strengths of semi-convection and overshooting convection were varied throughout their reasonable ranges. The overshooting factor is limited between 0.1 and 1.0 and very likely must be greater than 0.3 to prevent the unnatural bifurcation of the helium core and likely less than 0.6, according to the predicted scalings from Brummell *et al.* (2002) with respect to viscosity. According to the behavior of convection from Brummell *et al.* (2002), the core hydrogen and helium burning convection zones extend by overshooting by approximately 2% of a pressure scale height, and the envelope convection zone by approximately a quarter of a pressure scale height. Semi-convection mixes nearly as efficiently as convection for these stars, which is consistent with the findings of Moore and Garaud (2016) for lower mass stars.

The cores of these stars illustrate the strong dependence of the final star properties on these minor mixing processes. Strong overshooting convection, in particular, can lead to core masses of up to 20% larger than previous models of KEPLER have predicted. Given the predicted dependence on the overshoot depth on the viscosity of the system, helium cores are likely only about 5% larger than previous estimates.

Massive star calculations in the past generally result in stellar cores that lie along a particular helium core-carbon/oxygen core mass relation. This precise relation varies from code to code, with KEPLER producing the most massive carbon–oxygen cores for a given helium core mass of the codes surveyed. Altering the depth of overshooting convection in KEPLER and in MESA seems primarily only to shift along this relation, strengthening the conclusion that the helium core mass is one of the primary concerns regarding the final state of such a calculation. Deviations from this relation are possible but are only seen them when the helium core bifurcates in a way that appears unphysical.

Regarding the envelope structures of these stars, there are two general trends of behavior regarding the time spent as a red supergiant, which will play in importantly for determining the observable red to blue supergiant ratio. As the amount of overshooting convection increases, the denser cores lead to increasing time spent as red supergiants, which is due to the deeper gravitational potential of the core. Conversely, for those simulations with weak mixing, the opposite trend appears as these stars tend to spend more time on blue loops. This behavior can be explained by looking at the hydrogen abundance outside of the core, as detailed in Section 3.6.3.

Next reasonable steps for this work involve simulating a grid of stellar models across the massive star spectrum at least at the new recommended values for semi-convection and convective overshoot. These models can be exploded and the integrated properties can be examined, as in Sukhbold *et al.* (2016).

Chapter 4

An Analytic Understanding of Supergiant Envelopes

As was shown in Chapter 3, the strength of minor mixing processes can substantially affect the core masses and abundance profiles of massive stars. Lauterborn *et al.* (1971) found that these two properties seem to uniquely determine whether an evolved massive star is a blue or red supergiant through numerical simulations. An analytical understanding of this dependence is derived here, based on a series of simulations of supergiant envelopes with excised cores using KEPLER. This work is to be submitted to the Astrophysical Journal Letters.

4.1 Indroduction

The advanced stages of the lives of massive stars are quite complex, and a lack of understanding of the behaviors of supergiant atmospheres impedes gathering more information from observations. Supergiants are massive stars in the range of roughly $8M_\odot$ to $40M_\odot$ that have exhausted core hydrogen but that still have a substantial hydrogen envelope. Like intermediate-mass (Sun-like) stars, the onset of hydrogen shell burning leads to an expansion of the envelope. In the case of massive stars, however, the envelope of these advanced stages usually takes the form of either a red or blue supergiant, which differ by an order of magnitude in radius and a factor of 4 in effective temperature (with the red supergiant larger and cooler). As the properties of such envelopes inform all primary observable properties of these stars, a detailed understanding of how the envelope depends on the inner properties is needed.

An important observable in galactic and extragalactic studies of massive star evolution is the blue-to-red supergiant number ratio. This ratio has been the object of many observations of extragalactic stellar populations (e.g., van den Bergh, 1968; Meylan and Maeder, 1982; Humphreys and McElroy, 1984). The supergiant ratio is a strong function of metallicity, with a value near 4 for the low metallicity environment of the Small Magellanic Cloud and a value near 48 for the high metallicity environment of the inner Milky Way (Humphreys and McElroy, 1984). Such variations of the ratio with metallicity have been difficult to explain in simulations of massive stars because they depend on many uncertain choices regarding the physics of these stars. Langer and Maeder (1995) found, for example, that no calibration of convection physics could explain the ratio at all metallicities.

Many different factors contribute to whether a star evolves as a red or blue supergiant in simulations, including mass loss (Salasnich *et al.*, 1999; Meynet *et al.*, 2015), convection physics (Stothers and Chin, 1976; Langer and Maeder, 1995), and rotation (Ekström and others, 2012). Depending on the nature of these uncertain physical processes, a massive star can begin core helium burning as a red or blue supergiant and may transition from one to the other. The blue-to-red supergiant ratio can be used to test the properties of stellar interiors, and observations of this ratio are often used to calibrate the models of the aforementioned uncertain physical processes (e.g., the investigation of core helium burning supergiants in Dohm-Palmer and Skillman, 2002). One such process is convection (e.g., the use of the Ledoux or Schwarszchild criterion for convection in Stothers and Chin, 1976) which can substantially change the evolution on the H–R diagram, with the Ledoux criterion representing the observations more accurately. In addition, the addition of overshooting convection was shown to more accurately represent the colors and luminosities of blue loops (i.e., initially red supergiants that transition to blue before again becoming red) (Stothers and Chin, 1991). However, more recent observations of these stars still show substantial differences between the state-of-the-art isochrones in both magnitude and shape (e.g., McQuinn *et al.*, 2011)

One of the most blatant challenges to the understanding of massive star evolution remains the event of supernova 1987A, which raised serious questions as to the surface structures of these stars immediately prior to explosion. This supernova occurred in the Large Magellanic Cloud, originating from the blue supergiant Sanduleak -69 202 of class B3 I (Herald *et al.*, 1987). The accepted theory of massive star evolution at the time suggested that bright supernovae should come primarily from red supergiants (see e.g., Arnett, 1980). For a more thorough description of the supernova, the authors refer the reader to the review by Arnett (1989).

Despite these difficulties in theory, progress has been made. In particular, Lauterborn *et al.* (1971) quantified the presence of blue loops by the gravitational potential of the stellar core and the hydrogen fraction of the hydrogen burning shell. But so far, a physical explanation of these dependences has eluded stellar astrophysicists.

In this work, a series of simulations of stellar envelopes were run using KEPLER to explain the precise nature of the blue to red supergiant transition. In Section 4.2, the nature of these KEPLER simulations is outlined. In Section 4.3, the simulations are summarized and the main results are presented. Section 4.4 discusses these results in the context of previous work and presents a novel analytic understanding of the interior behavior of these envelopes.

4.2 Setup

To systematically investigate the behavior of the envelopes of evolved massive stars, a series of models was simulated with the one-dimensional stellar evolution code KEPLER (Weaver *et al.*, 1978). The goal is to determine how the equilibrium properties of the envelopes of these stars depend on the internal structure. Thus, the internal properties are held fixed while the envelope structure is allowed to adjust.

One of the main quantities that affects the envelope is the gravitational potential of the core (Lauterborn *et al.*, 1971). Thus, the mass and radius of the core were two of the primary input parameters of these simulations. To fix this potential, the computational domain began at a non-zero inner radius, r_i, within which a given mass was contained, M_i. The computational domain extended in mass coordinate from M_i to the total mass of the star M_o. Calculations such as this have been done since supernova 1987A to understand the nature of red and blue supergiant envelopes and their transition (see e.g., Saio *et al.*, 1988). A hypothetical core within the inner radius of the simulation would produce some luminosity through core fusion, and so the inner boundary was given a fixed incident luminosity, L_i. All of these properties were assumed constant in time.

The composition of the envelope is also known to affect whether a massive star becomes a red or blue supergiant (Lauterborn *et al.*, 1971; Walmswell *et al.*, 2015). In these simulations, the composition of the envelope was not allowed to evolve with time again to isolate the effects of the internal properties on the envelope structure. From $(M_o - M_i)/3+M_i$ to M_o, the composition of the envelope was assumed to be of mostly solar composition (Lodders, 2010). The only deviation from solar composition was that the ^{3}He was removed from the calculation as this became the primary source of energy generation otherwise. From M_i to $(M_o - M_i)/3 + M_i$, the composition differed from solar composition only in the hydrogen and helium mass fractions, X_i and Y_i, which were varied in the simulations, and in ^{3}He, which was removed from this region as well. The inner and outer envelopes are treated separately to focus solely on effects of composition changes near the core rather than the effects of the composition of the outer envelope, which was not explored here.

Each simulation was run until it reached a steady state, i.e., until the inner luminosity plus the energy generation on the grid equaled the luminosity coming out of the star. Each gridpoint could transport energy either radiatively using photon diffusion or convectively using mixing length theory based on the Ledoux criterion for convection (see e.g., Clayton, 1969). Within a Kelvin–Helmholtz time, the envelope adjusted to a steady state, either a red supergiant state where convection transported energy throughout most of the envelope or a blue supergiant state where radiation did so instead. In some of these simulations, hydrogen fusion ignited near the base of the simulated domain and produced energy comparable to L_i. Though the composition did not evolve in time, the energy generation could still be calculated for the hydrogen burning shell based on the composition and thermodynamic properties, and the star could evolve into a steady state appropriate to all of the fixed internal properties in these simulations.

Table 4.1: Parameters used in the envelope calculations

Parameter	Lower	Upper	Interval
Core Radius (10^x cm)	10.3	10.4	0.1
Core Luminosity (10^x erg/s)	37.7	38.2	0.1
Core Mass (10^x g)	33.75	34.0	0.05
Total Mass ($M_\odot$)	14.2	15.8	0.8
H Shell Abundance	0.411	0.711	0.1

4.3 Simulations

The parameters used in the series of 864 envelope calculations are given in Table 4.1. The choices for the ranges of these parameters come from typical values of a previous study of $15M_\odot$ stars (Chapter 3). After the simulation has come into steady state, the properties of that state were extracted, including the final structure and the total luminosity of the star, L_o.

4.3.1 Luminosity

The total luminosity of these models tracks almost exactly as

$$L_o = (1 - \beta_i) L_{\mathrm{edd},i}, \tag{4.1}$$

where L_o is the luminosity of the entire star, and $L_{\mathrm{edd},i}$ and β_i are defined as

$$L_{\mathrm{edd},i} = \frac{4\pi c G M_i}{\kappa_i}, \tag{4.2}$$

$$\beta_i = \frac{p_{\mathrm{ideal},i}}{p_i}, \tag{4.3}$$

where $L_{\mathrm{edd},i}$ is the Eddington luminosity at the internal boundary of the domain, κ_i is the opacity at that boundary, and β_i is the ratio of ideal gas pressure to the total

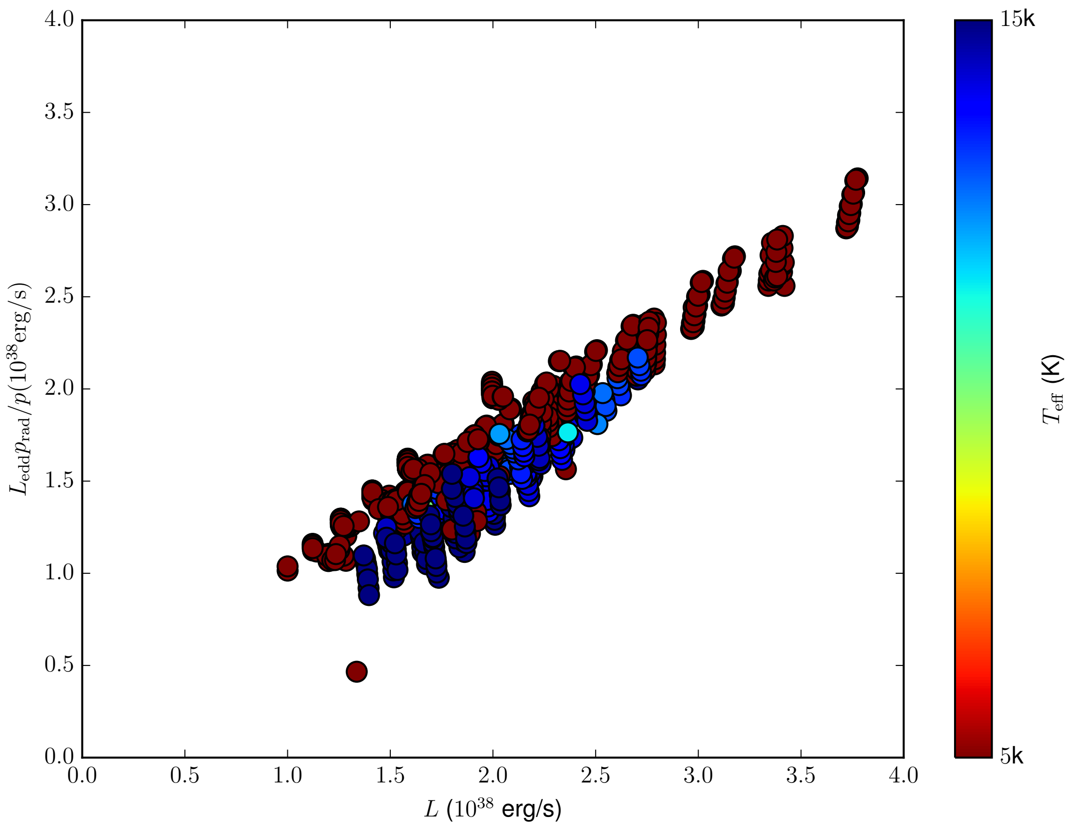


Figure 4.1: The expression $(1-\beta_i)L_{\mathrm{edd},i}$ as a function of total luminosity of the star, L_o, for each simulated envelope. The two quantities compare favorably for all the simulations, never differing by more than 17%. This makes for a convenient substitution in an analytic study of the problem of supergiant envelopes.

pressure evaluated at that boundary. Investigations of intermediate-mass red giants (e.g., Sugimoto and Fujimoto, 2000) find Equation 4.1 to be appropriate for such stars with large core luminosities. In application to the envelopes of more massive stars, Gräfener *et al.* (2012) investigated the envelope inflation of Wolf-Rayet stars and Luminous Blue Variables. The work in this paper falls in between the two extremes of Sugimoto and Fujimoto (2000) and Gräfener *et al.* (2012).

A comparison of L_o and $(1-\beta_i)L_{\mathrm{edd},i}$ is presented in Figure 4.1. The two quantities compare favorably at the low luminosity end, where the hydrogen burning shell has not completely taken over the luminosity of the star. For the stars with higher luminosity, the total luminosity drops below the rescaled Eddington luminosity by about 17%, but it is close enough to get some traction on the analytic understanding of the system. That this quantity would be so well correlated with the total luminosity of the system is not entirely surprising. This value of the luminosity is what is needed to precisely balance outward pressure with gravitational acceleration, so it is fitting that the total luminosity would match this value where the luminosity is generated.

4.3.2 Envelope Structure

Most of the envelope simulations fall into two categories, as can be seen in their effective temperatures in Figure 4.2: those with effective temperatures above about 10^4 K (blue supergiants) and those below (red supergiants). In general, the red supergiants are primarily convective throughout the upper envelope whereas blue

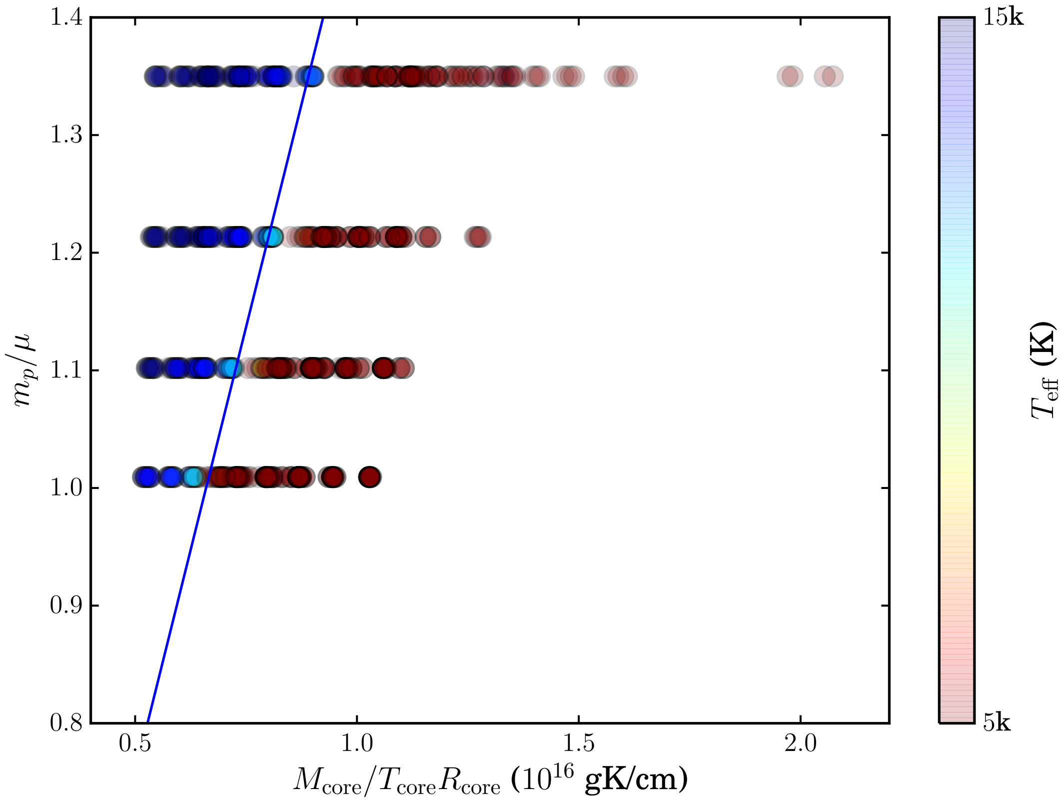


Figure 4.2: The effective temperature of each simulated envelope is plotted with respect to the mean molecular weight and the gravitational potential of the core divided by the temperature at the core/envelope boundary. This choice of axes clearly separates the red and blue supergiants along a line. The line shown is given by Equation 4.9, which analytically predicts the red to blue supergiant transition.

supergiants are primarily radiative. The separation between these two populations appears distinctly when plotted in the context of the composition and gravitational potential of the core. The general dependence on composition and gravitational potential was initially discussed in Lauterborn *et al.* (1971); however, a more analytic understanding is undertaken here.

4.4 Discussion

A purely theoretical understanding of the red and blue supergiant transition is constructed by investigating the radiative transport through the envelope and determining where convection sets in. The transition to convective transport happens when the opacity becomes dominated by bound-free interactions, at around 5×10^5K.

4.4.1 Analytical Model

An analytical model is constructed for the transition condition for red and blue supergiants. The envelope simulations show that the model depends effectively on two parameters: the mass of the helium core and the hydrogen fraction of the inner envelope.

Section 4.3.1 shows that the luminosity is given very nearly by Equation 4.1. The structure of a relaxed star is the quantity of interest, so the luminosity outside the hydrogen burning shell is assumed uniform, and M, κ, and β are assumed to change only slightly across the burning shell.

Photon diffusion yields (e.g., Equation 3-55 from Clayton, 1969)

$$\frac{dT}{dr} = -\frac{3}{16\pi ac}\frac{\kappa\rho L}{r^2T^3}. \tag{4.4}$$

Equation 4.1 can be substituted for L. In these simulations, β_i is on average 0.76 ± 0.09, meaning that the radiative pressure and ideal gas pressure are comparable, with the ideal gas pressure slightly more dominant. For simplicity, the pressure of the envelope is assumed to be half gas and half radiation pressure, $p \approx 2\rho kT/\mu$. Thus,

$$(1-\beta) \equiv \frac{p - p_{\text{ideal}}}{p} = \frac{p_{\text{rad}}}{p} \approx \frac{\mu a T^4}{2\rho kT}, \tag{4.5}$$

since the radiation pressure, $p_{\text{rad}} = aT^4$, comprises the rest of the pressure in a non-degenerate gas. Using Equation 4.1 with the above expression for $1-\beta$, Equation 4.4 becomes

$$\frac{dT}{dr} = -\frac{3G\mu M_i}{8kr^2}, \tag{4.6}$$

which is an integrable function of radius as long as μ does not change appreciably. For the simulations in this work, a constant μ is guaranteed; however, for a more elaborate model, Equation 4.6 might need to be solved numerically for a given μ profile. Note that the effect of opacity has vanished due to contributions of photon diffusion and the expression for the Eddington luminosity.

The solution, assuming μ is uniform, is just

$$T - T_i = \frac{3G\mu M_i}{8k}\left(\frac{1}{r} - \frac{1}{R_i}\right), \tag{4.7}$$

where T_i and R_i are the temperature and the radius at the core-envelope boundary. In these simulations, T_i (which is not an input to the calculations) is well constrained at 4.5×10^7K by a feedback mechanism limiting the strength of nuclear fusion.

If the temperature is constrained not to pass into the bound-free regime by twice the core radius, the condition for being a blue supergiant is as follows:

$$T_{\mathbf{BF}} - T_i > \frac{3G\mu M_i}{16k}\frac{1}{R_i}, \tag{4.8}$$

where $T_{\mathrm{BF}} \approx 5 \times 10^5$K is the temperature at which bound–free opacity becomes dominant. Since $T_{\mathbf{BF}}$ is two orders of magnitude smaller than T_i, this can be expressed as the simpler

$$T_i > \frac{3G\mu M_i}{16k}\frac{1}{R_i}. \tag{4.9}$$

The choice of twice the core radius is somewhat arbitrary; it compares favorably with the simulation results. However, if the radius choice is instead an infinite radius (i.e., the requirement that a blue supergiant must never have a convective envelope), the expression is instead

$$T_i > \frac{3G\mu M_i}{8k}\frac{1}{R_i}, \tag{4.10}$$

which is comparable.

If the temperature does not fall below this threshold, the envelope will remain in the regime where electron scattering dominates the opacity, and the envelope will not become convective; i.e., the star will not become a red supergiant. Looking again at Figure 4.2, this expression cleanly separates the two populations according to red and blue supergiants.

Thus, the physics of the situation are as follows. An increase in the gravitational potential of the core increases the luminosity above the hydrogen shell in order to support the shell material. A greater flux transported by photon diffusion requires a steeper temperature gradient, which results in crossing the bound-free transition sooner. Thus, a dense core leads to a star tending more toward a red supergiant.

The dependence on composition can be explained in a similar manner. As Walmswell *et al.* (2015) showed, it is the mean molecular weight of the material that matters and not because of its impact on the opacity, as that was shown to be irrelevant in Equation 4.6. Rather, it is the contribution to the pressure that determines the effect. A smaller ideal gas pressure (i.e., a larger μ) results in a larger luminosity needed to support the shell material. Thus, for the same reason as for the gravitational potential, a large mean molecular weight leads to a star tending more toward a red supergiant.

4.4.2 Comparison with Full Star Evolution

Our analytical model can be compared with the simulations from Section 3.5 in Figure 4.3, where the same quantities have been plotted as in Figure 4.2; however, here they are for full stellar evolution calculations taken halfway through core helium burning. The simulations from Section 3.5 were all 15 solar mass stars with strong variations in overshooting convection and semi-convection, so they provide an extensive sample of the possible results of stars in this mass regime. The results of the full star calculations still agree remarkably well with the analytic model above.

As has been shown, the temperature profile through the radiative part of the envelope is a simple expression of the temperature of the hydrogen burning shell (approximately fixed by the nuclear physics), the gravitational potential of the core, and the mean molecular weight just above the core. With a known temperature profile, determining whether a star is a red or blue supergiant is a simple matter.

This result sheds some light on the nature of blue loops. A single star in Figure 4.3 would evolve from left to right as hydrogen shell burning increases the total mass of the helium core. For a star to undergo a blue loop, it must begin with a relatively large mean molecular weight. The high mean molecular weight would put this star strongly into the red supergiant regime, meaning that the radius that the temperature drops below the bound-free opacity transition (and therefore into the convective regime) is quite deep in the star. Thus, the convection zone will bring low mean molecular weight material from the surface to near the core quickly, substantially lowering the mean molecular weight, and shifting the track rapidly upward in Figure 4.3. This moves the star into the blue supergiant regime, where changes in mean molecular weight happen more gradually due to the lack of convective interaction. Once a blue supergiant, the star continues to transform by increasing the mass of its helium core until it moves back across the transition.

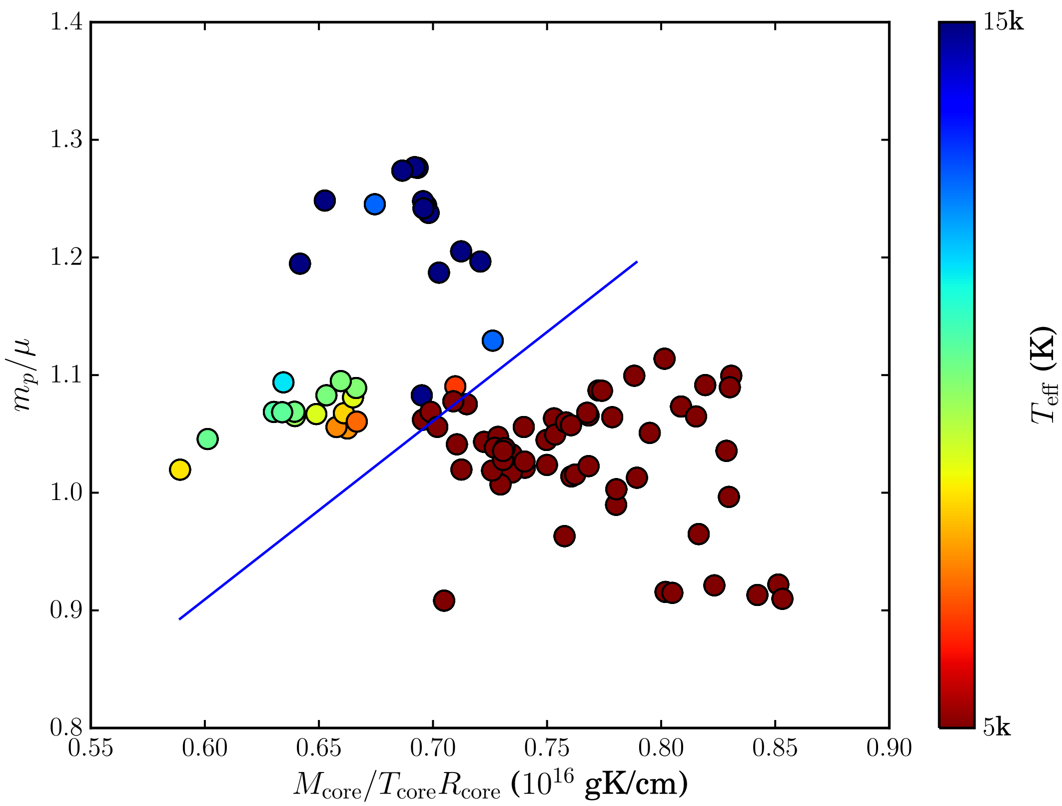


Figure 4.3: Full stellar evolution calculations are plotted in the same manner as in Figure 4.2. Here, each point represents the state of the star halfway through core helium burning. The prediction of this work for the supergiant transition still appears to adequately represent the transition; however, it does appear that the slope of the prediction is somewhat steeper than is observed. This could be explained by the choice of $\beta \approx 0.5$, which directly determines the slope of the predicted line.

4.5 Conclusions

The red to blue supergiant transition for stars near 15 solar masses can be explained by looking only at the properties of the star near the hydrogen burning shell. The transition can be described simply with the following equality:

$$T_C = \frac{3G\mu M_C}{16k}\frac{1}{R_C}. \tag{4.11}$$

This describes the transition well for both the envelope calculations included here and for full star evolution calculations during helium burning.

In the future, this work can be expanded by investigating other mass regimes expected to undergo blue loops. In particular, the effects of mass loss on red and blue supergiants (Salasnich *et al.*, 1999; Meynet *et al.*, 2015) does not appear in this formalism and could be explored easily by expanding the range of M_o explored. With this new analytic framework, Supernova 1987A can be revisited in order to attempt to understand what effects could have led to its unexpected evolution and constrain whether a merger is indeed required for such an evolution.

Chapter 5

A Semi-Analytic Flux Model for Fingering Convection in Astrophysics

Having studied new prescriptions of semi-convection and overshooting convection from numerical simulations, attention is now drawn to a third minor mixing process of relevance to astrophysics, fingering convection. Here, we perform multi-dimensional simulations of fingering convection and derive a semi-analytical model that can be incorporated into stellar evolution codes. Though fingering convection is usually referred to in the context of low-mass stars, it can be relevant in massive stars as well, as is seen in Figure 3.1. The contents of this chapter come from Sections 1, 2, 3, 4, and 6 of Brown *et al.* (2013).

5.1 Introduction

The phenomenon of fingering convection (otherwise known as thermohaline convection) occurs in regions of stellar and planetary interiors where the mean molecular weight, μ, increases upwards while remaining stable to the Ledoux criterion. If not for the effects of diffusion, such a system would be stable to small perturbations. In the presence of diffusion, however, instability can occur. Indeed, when a high μ, high entropy parcel is displaced downward, heat rapidly leaks via thermal diffusion. Since compositional diffusion is much weaker, the parcel remains denser than its surroundings and proceeds to sink. The instability takes the form of finger-like structures and can significantly increase the flux of heat and composition above that of molecular or radiative diffusion.

Only in the last few years have numerical simulations of fingering convection approaching the astrophysical parameter regime become more readily available (e.g., Denissenkov, 2010; Traxler *et al.*, 2011a). Denissenkov (2010) ran 2D simulations to measure the aspect ratio of fingers and used the previous literature and a dimensional argument to find a prescription for mixing by fingering convection in red giant branch stars. He concluded that while this process could provide some of the required mixing discussed by Charbonnel and Zahn (2007), it alone was insufficient to account for the observed abundances of low-mass red giant branch stars. Traxler *et al.* (2011a) presented 3D numerical simulations of fingering convection and generated an empirical fit of their results to propose transport laws for fingering convection in astrophysics. Using their mixing model, Garaud (2011) studied the evolution of the surface metallicity of solar-type stars after planet infall and found that fingering convection would transport the material out of the convective zone too quickly to explain the planet-metallicity connection. This then suggests that planets can form more easily in high-metallicity proto-planetary disks. However, the results of Traxler *et al.* (2011a) also confirm the findings of Denissenkov (2010) on the red giant branch star abundances.

5.2 Code

As discussed in Traxler *et al.* (2011a), the characteristic length scale of the fingering instability in astrophysical objects is much smaller than a pressure scale height, which in turn is generally small enough to ensure that the curvature of the star plays little role in the system dynamics. We therefore use a Cartesian grid (x, y, z), with gravity given by $\mathbf{g} = -g\hat{\mathbf{z}}$ to model a small region of the star. In addition, the velocities of these systems are much slower than the sound speed, so we can treat the fluid with the Boussinesq approximation (Spiegel and Veronis, 1960). We assume that there is a background temperature profile, $T_0(z)$, and a mean molecular weight profile, $\mu_0(z)$, which depend only on z. If the region considered is small enough, these background profiles can be approximated by linear functions with uniform gradients, T_{0z} and μ_{0z}, respectively. We therefore assume that each profile is comprised of a linear background state and a perturbation, e.g. $T_{\rm tot} = zT_{0z} + T$.

The equation of state for the density perturbation, ρ, within the Boussinesq approximation, is

$$\frac{\rho}{\rho_0} = -\alpha T + \beta \mu, \tag{5.1}$$

where ρ_0 is the mean density of the region. This assumes that the density of the fluid in the region is roughly uniform, which is an excellent approximation at these scales, except in the effects of buoyancy, where the density variations are small enough to only constitute a first order correction to the background. The coefficients α and β are those of thermal expansion and compositional contraction and can be obtained by linearizing the equation of state around ρ_0. The remaining governing equations for a Boussinesq system are (Spiegel and Veronis, 1960)

$$\nabla \cdot \mathbf{v} = 0, \tag{5.2}$$

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} = -\frac{\nabla p}{\rho_0} + g(\alpha T - \beta \mu)\hat{\mathbf{z}} + \nu \nabla^2 \mathbf{v}, \tag{5.3}$$

$$\frac{\partial T}{\partial t} + \mathbf{v} \cdot \nabla T + w\left(T_{0z} - T_{\mathrm{ad}}\right) = \kappa_T \nabla^2 T, \text{ and} \tag{5.4}$$

$$\frac{\partial \mu}{\partial t} + \mathbf{v} \cdot \nabla \mu + w \mu_{0z} = \kappa_\xi \nabla^2 \mu, \tag{5.5}$$

where $\mathbf{v} = (u, v, w)$ is the fluid velocity, p is the pressure, ν is the kinematic viscosity, κ_T is the thermal diffusivity, κ_ξ is the compositional diffusivity, and T_{ad} is the background adiabatic temperature gradient, assumed constant within the considered domain. It should be noted here that ν, κ_T, and κ_ξ are assumed to be constant as additional requirements of the Boussinesq approximation. Other interesting dynamics may arise for variable ν, κ_T, and κ_ξ, but these are not addressed here.

Given a stable density gradient with opposing contributions from temperature and composition, two possibilities exist: one with $T_{0z} - T_{\rm ad} > 0$ (fingering convection) and one with $T_{0z} - T_{\rm ad} < 0$ (semi-convection). We then non-dimensionalize the governing equations, taking our length scale to be the anticipated finger width, $d = (\kappa_T \nu / g\alpha |T_{0z} - T_{\rm ad}|)^{1/4}$ (Stern, 1960; Kato, 1966). We scale time, temperature, and composition as $[t] = d^2/\kappa_T$, $[T] = (T_{0z} - T_{\rm ad})d$, and $[\mu] = (\alpha/\beta)(T_{0z} - T_{\rm ad})d$. With the Prandtl number defined as $\mathrm{Pr} \equiv \nu/\kappa_{\rm T}$ and the diffusivity ratio as $\tau \equiv \kappa_\mu/\kappa_T$, the non-dimensionalized governing equations take the following form:

$$\hat{\nabla} \cdot \hat{\mathbf{v}} = 0, \tag{5.6}$$

$$\frac{1}{\mathrm{Pr}} \left(\frac{\partial \hat{\mathbf{v}}}{\partial \hat{t}} + \hat{\mathbf{v}} \cdot \hat{\nabla} \hat{\mathbf{v}} \right) = -\hat{\nabla} p + (\hat{T} - \xi)\hat{\mathbf{z}} + \hat{\nabla}^2 \hat{\mathbf{v}}, \tag{5.7}$$

$$\frac{\partial \hat{T}}{\partial \hat{t}} + \hat{\mathbf{v}} \cdot \hat{\nabla} \hat{T} \pm \hat{w} = \hat{\nabla}^2 \hat{T}, \text{ and} \tag{5.8}$$

$$\frac{\partial \xi}{\partial \hat{t}} + \hat{\mathbf{v}} \cdot \hat{\nabla} \xi \pm \frac{\hat{w}}{R_0} = \tau \hat{\nabla}^2 \xi, \tag{5.9}$$

where the $\pm$ is positive for fingering convection and negative for semi-convection.

Equation 5.9 introduces a new parameter, R_0. This so-called "density ratio" is the ratio of the stabilizing entropy gradient to the destabilizing compositional gradient (Ulrich, 1972),

$$R_0 = \frac{\nabla - \nabla_{\rm ad}}{\frac{\phi}{\delta} \nabla_\mu} = \frac{\alpha (T_{0z} - T_{\rm ad})}{\beta \mu_{0z}}. \tag{5.10}$$

In the case of fingering convection, several limits appear from linear stability. If $R_0 < 1$, the destabilizing compositional gradient dominates, and the system is unstable to the Ledoux criterion. One expects the region to be fully mixed by overturning convection. For $R_0 \gg 1$, the stabilizing entropy gradient dominates, and the system is stable. The only possible transport is via diffusion. For intermediate values, $1 < R_0 < 1/\tau$, Baines and Gill (1969) showed that the system is unstable to fingering convection. As in Traxler *et al.* (2011a), we define a reduced density ratio,

$$r = \frac{R_0 - 1}{\tau^{-1} - 1}, \tag{5.11}$$

which remaps the "fingering regime" to the interval of $r \in [0, 1]$ regardless of the value of τ.

In the case of semi-convection, the inverse density ratio R_0^{-1} is more frequenty used. If $R_0^{-1} < 1$, the destabilizing entropy gradient dominates, and the system is unstable to the Ledoux criterion. For $R_0^{-1} \gg 1$, the stabilizing composition gradient dominates, and the system is stable. The only possible transport is through diffusion. For intermediate values, $1 < R_0^{-1} < (1 + \mathrm{Pr}) / (\tau + \mathrm{Pr})$, Walin (1964) showed that the system is unstable to semi-convection.

We modeled double-diffusive convection numerically, solving Equations 5.6 through 5.9 using the pseudo-spectral double-diffusive convection code from Traxler *et al.* (2011a,b) and Stellmach *et al.* (2011). All perturbations T, μ, and $\mathbf{v}$ were assumed to be triply periodic. In all simulations described below, unless specifically noted, the computational domain is taken to be a cube with side length $100d$. Since the wavelength of the fastest growing mode ranges from $6d$ to $15d$ in the parameter regime considered in this paper, this implies that there are at least 5 to 10 wavelengths of the fastest growing mode per $100d$. We tested the required resolution by inspecting the compositional and vorticity fields, which are in general the least resolved ones given the small values of Pr and τ selected.

5.3 Simulations

Traxler *et al.* (2011a) explored a significant sample of parameter space, with Pr and τ ranging from 1/3 down to 1/30, as seen in Table 5.1. However, they only tested a few different density ratios in the cases with Pr, $\tau \sim 1/30$. Furthermore, owing to computational limitations, they did not explore the low R_0 limit. We explored more comprehensively in R_0 the parameters chosen by Traxler *et al.* (2011a) and also ran simulations down to Pr, $\tau \sim 1/100$. It is currently computationally prohibitive to reduce Pr, τ much lower than this, particularly at low R_0, where the simulations become increasingly turbulent.

As in Traxler *et al.* (2011a), we report our results in the form of the Nusselt numbers, which are defined as the ratio of the total vertical flux to the diffused flux and expressed here in terms of the non-dimensional fields as

$$\mathrm{Nu_T} = 1 - \langle wT \rangle, \tag{5.12}$$

$$\mathrm{Nu}_\mu = 1 - \frac{R_0}{\tau}\langle w\mu \rangle. \tag{5.13}$$

Note that the form in terms of dimensional variables w and T is $\mathrm{Nu_T} = 1 - \langle \mathrm{wT} \rangle / \kappa_\mathrm{T}(\mathrm{T_{0z}} - \mathrm{T_{ad}})$. This definition of the Nusselt number describes the flux of the potential temperature and not that of temperature. The two are only equal when $T_\mathrm{ad} = 0$. For a more complete discussion, see Mirouh *et al.* (2012). The vertical turbulent fluxes, $\langle wT\rangle$ and $\langle w\mu\rangle$, are measured by taking the average of the products wT and $w\mu$ over the entire domain.

When we discuss layer formation, we will also use the ratio of the total non-dimensional fluxes,

$$\gamma = \frac{-1+\langle wT\rangle}{-\frac{\tau}{R_0}+\langle w\mu\rangle} = \frac{R_0\mathrm{Nu_T}}{\tau\mathrm{Nu}_\mu}, \tag{5.14}$$

and the ratio of the turbulent non-dimensional fluxes,

$$\gamma_{\mathrm{turb}} = \frac{\langle wT\rangle}{\langle w\mu\rangle} = \frac{R_0(\mathrm{Nu_T}-1)}{\tau(\mathrm{Nu}_\mu-1)}. \tag{5.15}$$

5.3.1 Typical and Atypical Simulations

In Figure 5.1, we present the temporal evolution of the thermal and compositional Nusselt numbers in a typical numerical simulation of fingering convection, taking $\mathrm{Pr} = 1/10$, $\tau = 1/30$, and $R_0 = 3$. This evolution begins with a period of exponential growth from $t = 0$ to $t = 160$. The observed growth rate is twice the growth rate of the fastest growing mode according to a linear analysis of the governing equations. This is as expected, as shown in Section 5.4.1. The transport peaks around $t = 165$, after which the fluxes of temperature and composition saturate (at a level slightly below the peak) and remain roughly constant for the remainder of the simulation.

In order to visualize the saturation process, we present snapshots of the compositional field of the simulation at three stages before and after saturation in Figures Figure 5.2(a), (b), and (c), marked with vertical lines in Figure 5.1. In Figure 5.2(a),

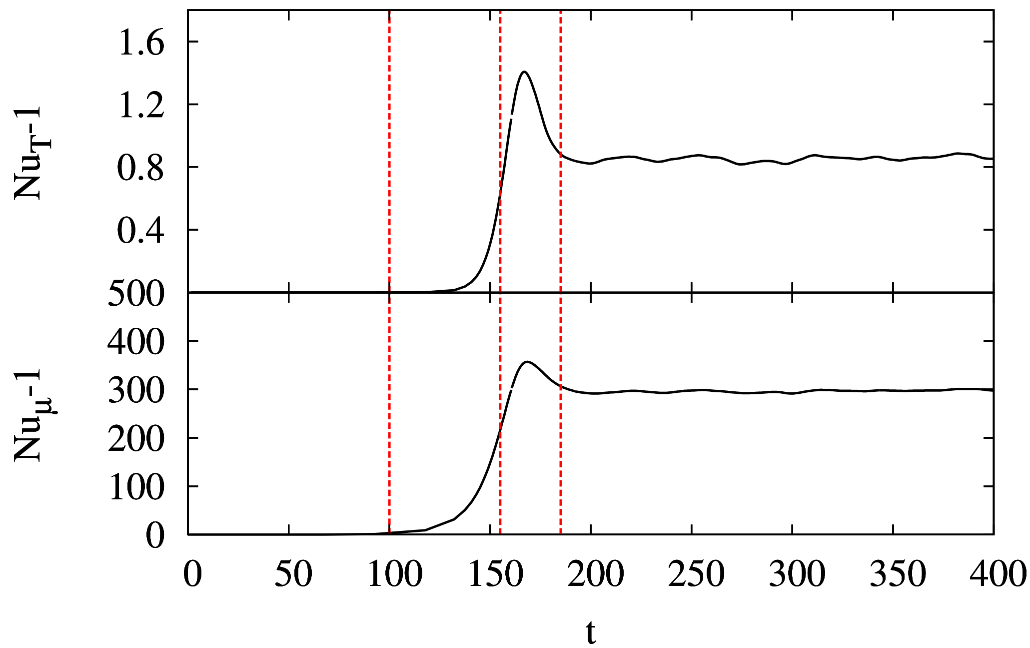


Figure 5.1: A typical example of the evolution of the thermal (top) and compositional (bottom) Nusselt numbers, $\mathrm{Nu_T}$ and Nu_μ (see Equation 5.12) in a simulation with $\mathrm{Pr} = 1/10$, $\tau = 1/30$ with $R_0 = 3$. Time is measured in units of the thermal diffusion time across a length d. We note that the Nusselt curves grow exponentially, peak, and fall to a quasi-steady equilibrium state, which we call the saturated regime. The vertical lines indicate the times of the snapshots in Figure 5.2.

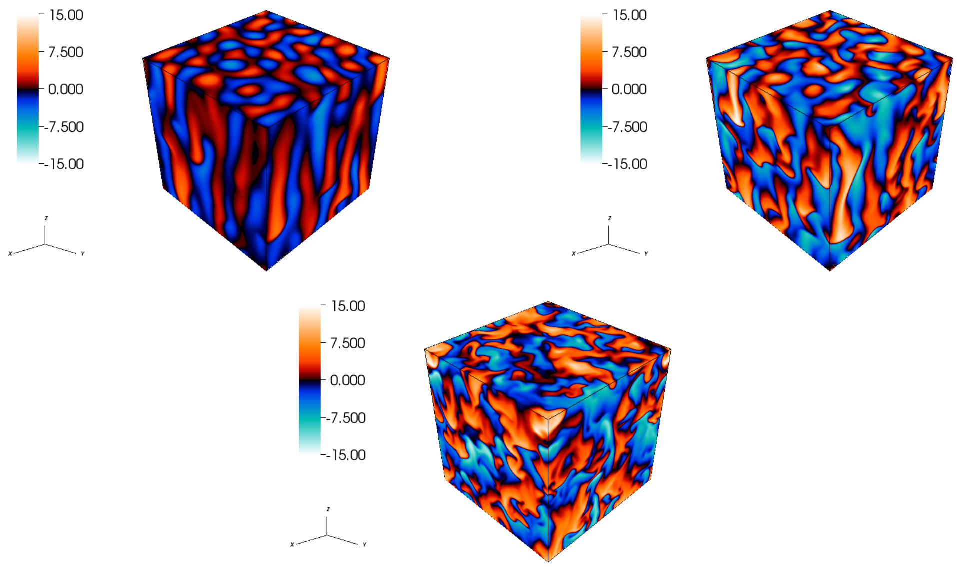


Figure 5.2: Snapshots of the compositional perturbation in the simulation shown in Figure 5.1 at the characteristic times marked in Figure 5.1. In Figure 5.2(a) at $t \approx 100$, prior to saturation, the composition field is dominated by tall elevator mode structures characteristic of linear fingering convection. In Figure 5.2(b) at $t \approx 155$, though the original elevator modes are still recognizable, they have been disrupted by secondary instabilities. At $t \approx 180$, Figure 5.2(c) illustrates the saturated regime discussed in Figure 5.1. Note that the elevator modes have been completely destroyed by the secondary instabilities, leaving a domain filled completely with homogeneous turbulence.

which shows the state just as the primary fingering instability begins to develop, we see tall and thin vertical structures, called elevator modes, with high μ plumes sinking and low μ plumes rising. In Figure 5.2(b), which is taken just before the peak in Figure 5.1, we recognize the same vertical structures but notice that smaller-scale instabilities have distorted them. The latter eventually destroy the elevator modes and the system achieves saturation in Figure 5.2(c). At this point, very little of the original elevator modes remain discernible, and quasi-steady, homogeneous turbulence dominates the dynamics.

While most simulations behave exactly as Figure 5.1, in some cases where R_0 is chosen to be very close to the onset of overturning convection (i.e. for small r, see Equation 5.11), the Nusselt numbers begin to increase gradually again after saturation. This behavior appears clearly in Figure 5.3, for which $\mathrm{Pr} = 1/10$, $\tau = 1/30$, and $R_0 = 1.1$. Two phases exist after saturation: from $t = 100$ to $t = 600$, the mean Nusselt numbers increase slowly and undergo quasi-periodic oscillations. We find that these oscillations have twice the buoyancy frequency, which is what we expect for gravity waves. The mean Nusselt numbers rise due to nonlinear effects in the waves. At $t = 600$, the transport increases. Meanwhile, the mean density profile in part of the domain inverts, suggesting the formation of a fully convective layer (see Section 5.4.2). This is reminiscent of the formation of gravity waves and layer development in oceanic fingering convection (Stellmach *et al.*, 2011). We will discuss these large-scale instabilities and related transport properties in Section 5.4.2.

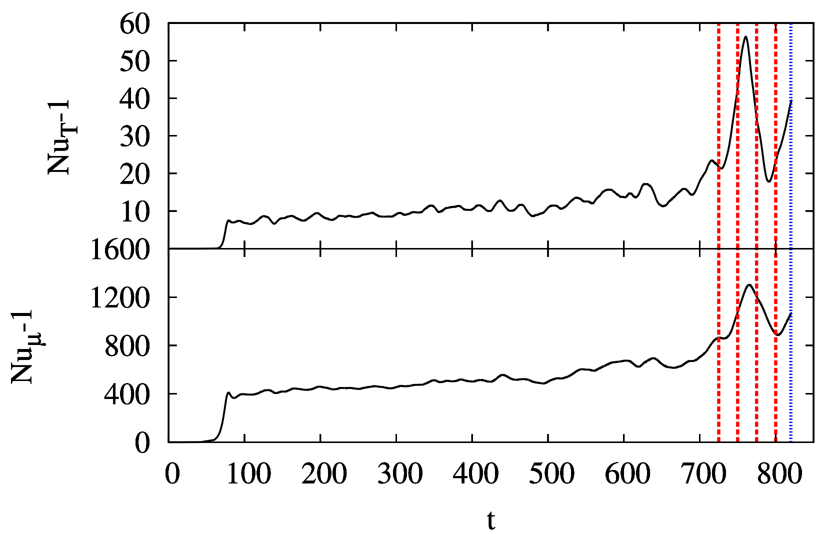


Figure 5.3: An atypical simulation showing the development and effect of large-scale instabilities on the Nusselt numbers, Nu_T and Nu_μ in a simulation with $\mathrm{Pr} = 1/10$, $\tau = 1/30$, and $R_0 = 1.1$. We note that this simulation behaves quite differently from the one shown in Figure 5.1. After the initial saturation at $t = 80$, both Nusselt numbers gradually increase until $t = 700$. They both exhibit quasi-periodic oscillations at twice the buoyancy frequency (though such oscillations are more apparent in the thermal Nusselt number), which suggests the development of gravity waves. At $t = 700$, the transport increases in a manner similar to the onset of layers in the oceanographic case (e.g. simulations by Stellmach *et al.*, 2011). The blue dotted line marks the time at which we plot a snapshot of the compositional perturbation in Figure 5.8, and the red dashed lines indicate the timesteps for which we calculate the density profiles in Figure 5.9.

5.3.2 Extraction of the Fluxes

To calculate the saturated fluxes in the homogenous phase of all simulations, i.e., prior to the onset of any large-scale dynamics discussed above, we used the following method. We identified the first local minimum in $\mathrm{Nu_T}(t)$ after saturation and set this to be the beginning of the saturated regime. We then fit the remainder of the time series to a linear function of time with linear regression. If the uncertainty in the time derivative is greater than its magnitude, the Nusselt curve after saturation is effectively flat and the mean flux is the time average in this interval with the error taken to be the root mean square variation in the fluxes. On the other hand, for a simulation like the one shown in Figure 5.3, where the flux saturates initially but gravity waves begin enhancing the transport again thereafter, we isolated the duration of the saturated regime by reducing the later end of the interval until the linear fit was effectively flat.

Table 5.1: Turbulent flux measurements and respective errors for low Pr, low τ simulations of thermohaline convection. (*) Data from Traxler *et al.* (2011a). (†) Data from Radko and Smith (2012) with a domain size of $38.2 \times 38.2 \times 76.3$ and a resolution of $256 \times 256 \times 512$.

Pr	τ	R_0	$\mathrm{Nu_T}$	σ_T	Nu_μ	σ_μ	Layers
1/3	1/3	1.02003*	31.0	4.0	120.0	10.0	N

		1.08012*	23.0	1.0	92.0	5.0	N
		1.10015*	20.0	1.0	83.0	5.0	N
		1.2003*	13.5	0.6	60.0	3.0	N
		1.30045*	9.9	0.5	46.0	2.0	N
		1.50075*	5.7	0.2	28.3	0.9	N
		2.0015*	2.16	0.05	9.5	0.3	N
		2.4021*	1.3	0.01	3.5	0.09	N
		2.8027*	1.037	0.003	1.33	0.02	N
		2.90285*	1.0159	0.0009	1.143	0.008	N
1/3	1/10	1.01	28.6	0.4	428.0	6.0	Y
		1.54*	9.0	0.2	197.0	4.0	N
		3.97*	1.62	0.02	38.7	1.0	N
		7.03*	1.075	0.003	7.6	0.2	N
		9.46*	1.0041	0.0004	1.41	0.04	N
1/10	1/3	1.10015*	8.0	0.5	33.0	2.0	N
		1.70105*	2.16	0.05	8.7	0.3	N
		2.30195*	1.24	0.01	2.9	0.1	N
		2.8027*	1.019	0.0007	1.17	0.006	N
		2.90285*	1.0065	0.0003	1.059	0.003	N
1/10	1/10	1.003	11.36	0.09	169.0	2.0	Y
		1.09*	8.3	0.7	143.0	9.0	N

		1.45*	4.3	0.1	86.0	3.0	N
		1.99*	2.5	0.05	54.0	2.0	N
		2.8*	1.73	0.02	34.8	1.0	N
		2.98*	1.62	0.01	31.3	0.6	N
		3.25*	1.51	0.02	27.4	0.8	N
		4.96*	1.148	0.003	11.6	0.2	N
		7.03*	1.032	0.001	3.9	0.1	N
		9.1*	1.0036	0.0001	1.36	0.01	N
1/10	1/30	1.1	7.6	0.1	500.0	7.0	Y
		1.5	4.24	0.06	347.0	5.0	N
		2.0	2.77	0.02	252.0	2.0	N
		3.0	1.85	0.02	174.0	3.0	N
		6.0	1.253	0.004	95.0	1.0	N
		8.0	1.1382	0.0006	66.0	0.3	N
		10.963*	1.059	0.002	32.4	0.7	N
		20.34*	1.0075	0.0005	7.1	0.4	N
1/10	1/100	1.1†	7.4812	-	1584.517	-	N
		1.3†	5.1979	-	1285.765	-	N
		1.5†	4.1962	-	1124.055	-	N
		1.7†	3.581	-	1031.832	-	N
		1.9†	2.914	-	884.519	-	N

		2.1†	2.6215	-	816.522	-	N
		2.3†	2.4437	-	789.866	-	N
		2.5†	2.2994	-	761.9	-	N
		2.7†	2.0685	-	684.936	-	N
		2.9†	1.9298	-	636.724	-	N
		6.0	1.316	0.002	416.0	2.0	N
		12.0	1.088	0.002	214.0	4.0	N
1/30	1/10	1.1	4.92	0.08	77.0	1.0	N
		1.5	2.26	0.04	37.8	0.9	N
		2.0	1.63	0.01	24.8	0.3	N
		3.97*	1.141	0.005	9.9	0.3	N
		7.03*	1.021	0.001	2.83	0.09	N
1/100	1/100	5.0	1.072	0.001	102.0	1.0	N
		10.0	1.02569	3×10^{-5}	62.9	0.2	N

5.3.3 Numerical Simulations

A compilation of our new results with published data from Traxler *et al.* (2011a) and Radko and Smith (2012) is given in Table 5.1. Figure 5.4 shows the corresponding values of the mean Nusselt numbers of both composition and temperature, in the homogenous saturated phase, as a function of the reduced density ratio r (see

Equation 5.11). As expected, we see that both $\mathrm{Nu_T} - 1$ and $\mathrm{Nu}_\mu - 1$ tend to 0 as $r \to 1$. In contrast, as $r \to 0$, both $\mathrm{Nu_T} - 1$ and $\mathrm{Nu}_\mu - 1$ increase rapidly, presumably approaching values of Rayleigh-Benard convection in a triply-periodic domain (Calzavarini *et al.*, 2005; Garaud *et al.*, 2010). As an additional note, it is clear from Figure 5.4 that Nu_μ does not depend on Pr and τ individually, but rather on their ratio. This effect had already been observed in Traxler *et al.* (2011a), and led them to propose their empirical scaling laws:

$$\mathrm{Nu_T} - 1 = \mathrm{Pr}^{1/2}\tau^{3/2}\mathrm{f(r)}, \tag{5.16}$$

$$\mathrm{Nu}_\mu - 1 = \sqrt{\frac{\mathrm{Pr}}{\tau}}g(r), \tag{5.17}$$

where $g(r)$ and $f(r)$ have the form $ae^{-br}(1-r)^c$, where for $f(r)$, $a = 264 \pm 1$, $b = 4.7 \pm 0.2$, $c = 1.1 \pm 0.1$, and for $g(r)$, $a = 101 \pm 1$, $b = 3.6 \pm 0.3$, $c = 1.1 \pm 0.1$.

As discussed in Section 1, the prescription given by Traxler *et al.* (2011a) agrees well with the data for high R_0; however, these scalings underestimate the transport at low R_0, particularly as Pr and τ decrease, as is illustrated in Figure 5.4. It is clear from Figure 5.4 that their prescription breaks down at very low r, particularly as Pr decreases. For this reason we now propose a theory that better models the data at both large and small r. We derive this new theory with physical justification.

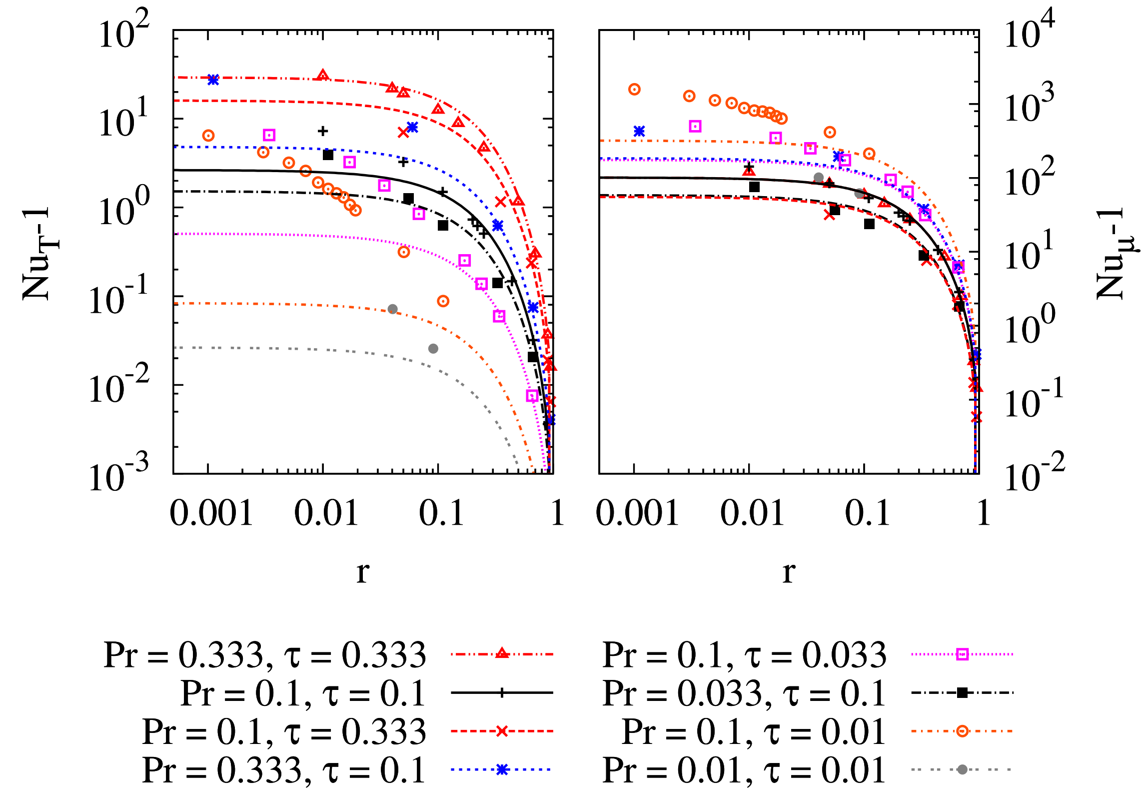


Figure 5.4: The Nusselt numbers as a function of the reduced density ratio, r, reported from the simulations in Table 5.1 (see references therein). Note that the compositional Nusselt number depends only on the ratio τ/Pr and r, as noted in Traxler *et al.* (2011a). For all simulations, the error bars are smaller than the size of the plotted symbols. We also include the model from Traxler *et al.* (2011a) as the accompanying curves. It is clear that the model works well for large r. However, at low r, the model can underestimate the Nusselt number by up to two orders of magnitude.

5.4 Discussion

Recently, Radko and Smith (2012) proposed a theory that characterizes the saturation of the fingering instability and predicts the thermal and compositional fluxes at saturation. The fundamental idea of this theory is straightforward: the saturation is thought to be caused by a secondary instability of the elevator modes and occurs when the growth rate of the secondary instability, σ, is of the order of the growth rate of the elevator mode, λ. The latter can be obtained by linearizing the governing equations and deducing the properties of the fastest growing mode. Radko and Smith (2012) calculated the growth rate of the secondary instability using Floquet theory. Unfortunately, their procedure turns out to be computationally quite expensive. In stellar evolution codes, fluxes need to be calculated for different values of Pr, τ, and R_0 corresponding to each radial mesh point considered at each timestep. Using the Radko and Smith (2012) method would be prohibitive for this application.

We pursue a theoretical model that is effectively a simplified form of the Radko and Smith (2012) theory but has the advantage of being much more straightforward and can thus be used in real time in stellar evolution codes. In what follows, we first review the derivation of the growth rate of the fastest growing mode, λ, in Section 5.4.1 (Baines and Gill, 1969). In Section 5.4.2, we then propose a very simple analytical formula for the growth rate of the secondary instability, σ, and finally determine the Nusselt numbers expected when $\sigma \sim \lambda$.

5.4.1 Fastest Growing Modes

The properties of the fastest growing fingering mode can be determined by linearizing the governing equations and assuming exponential solutions of the form

$$q = \hat{q}e^{\lambda t + i\mathbf{k}\cdot\mathbf{x}}, \tag{5.18}$$

for each of the velocity, temperature, composition and pressure perturbations (Baines and Gill, 1969). This yields a cubic equation for the growth rate λ, with coefficients that depend on the wave vector, $\mathbf{k}$, and the governing non-dimensional parameters, Pr, τ, and R_0. It can be shown that all properties of the fastest growing mode (e.g., the velocity, temperature, and salinity fields) are independent of z, hence the term "elevator" mode used to describe them. Since our equations are symmetric in x and y, λ only depends on the magnitude l of the horizontal wavenumber and not on its direction.

The resulting cubic then reads (Baines and Gill, 1969):

$$\begin{aligned}
&\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0, \text{ where} \qquad (5.19)\\
&a_2 = l^2(1 + \mathrm{Pr} + \tau),\\
&a_1 = l^4(\tau\mathrm{Pr} + \mathrm{Pr} + \tau) + \mathrm{Pr}\left(1 - \frac{1}{\mathrm{R}_0}\right),\\
&a_0 = l^6\tau\mathrm{Pr} + \mathrm{l}^2\mathrm{Pr}\left(\tau - \frac{1}{\mathrm{R}_0}\right).
\end{aligned}$$

To identify the fastest growing mode, we maximize λ with respect to the wavenumber

l. This yields a quadratic for λ:

$$a_2\lambda^2 + a_1\lambda + a_0 = 0, \text{ where} \tag{5.20}$$

$$a_2 = 1 + \mathrm{Pr} + \tau,$$

$$a_1 = 2l^2(\tau\mathrm{Pr} + \tau + \mathrm{Pr}),$$

$$a_0 = 3l^4\tau\mathrm{Pr} + \mathrm{Pr}\left(\tau - \frac{1}{\mathrm{R}_0}\right).$$

Equations 5.19 and 5.20 can be solved numerically simultaneously to determine the wavelength and growth rate of the fastest growing mode. More interestingly for anyone interested in quick analytical estimates, it can be shown that in the limit of Pr, $\tau \ll 1$,

$$\lambda \approx \begin{cases} \sqrt{\mathrm{Pr}} & \text{if } r \ll \mathrm{Pr} \ll 1 \\ \sqrt{\frac{\mathrm{Pr}\tau}{r}} & \text{if Pr} \ll \text{r} \ll 1 \end{cases}, \tag{5.21}$$

and

$$l^2 \approx \frac{1}{\sqrt{1 + \tau/\mathrm{Pr}}}. \tag{5.22}$$

A detailed derivation of these asymptotic limits (and higher-order terms) is presented in Brown *et al.* (2013).

5.4.2 Estimating the Nusselt Number

Radko and Smith (2012) consider all possible sources of secondary instabilities for the elevator modes. Here, for simplicity, we restrict our analysis to instabilities arising from the shear between adjacent elevators and neglect viscosity. As we demonstrate below, this yields very satisfactory results, and allows for a much simpler solution to the problem.

The elevator modes are characterized by a vertically invariant flow field of the kind

$$\mathbf{U}(x,t) = w_0 e^{\lambda t} \sin(lx)\hat{\mathbf{z}} \equiv w_{\mathrm{E}}(t)\sin(lx)\hat{\mathbf{z}} \ , \tag{5.23}$$

where λ and l are now strictly defined as the growth rate and horizontal wavenumber of the fastest growing mode, introduced in the previous section. Without loss of generality, we have selected the horizontal phase of the mode to be $\sin(lx)$. This flow field is associated with a temperature and composition field

$$T_{\mathrm{E}}(x,t) = T_0 e^{\lambda t}\sin(lx) \text{ and } \mu_{\mathrm{E}}(x,t) = \mu_0 e^{\lambda t}\sin(lx) \ , \tag{5.24}$$

where w_0, T_0, and μ_0 are related via Equations 5.8 and 5.9:

$$\lambda T_0 + w_0 = -l^2 T_0 \Rightarrow T_0 = -\frac{w_0}{\lambda + l^2} \ , \tag{5.25}$$

$$\lambda\mu_0 + R_0^{-1} w_0 = -\tau l^2 \mu_0 \Rightarrow \mu_0 = -\frac{R_0^{-1} w_0}{\lambda + \tau l^2} \ , \tag{5.26}$$

since elevator modes are exact solutions of the full set of nonlinear equations. At early times, the velocity shear, $w_E(t)$, is weak and the elevator modes grow unimpeded. Using the definitions of the Nusselt numbers given in Equation 5.12 with Equations 5.23, 5.24, and 5.25, we find that

$$\begin{aligned} \mathrm{Nu_T(t)} &= 1 - \frac{1}{L_x L_y L_z}\int_x\int_y\int_z w_{\mathrm{E}}(t) T_{\mathrm{E}}(t)\sin^2(lx)\,dx dy dz \\ &= 1 - \frac{1}{2} w_{\mathrm{E}}(t) T_{\mathrm{E}}(t) = 1 + \frac{w_0^2}{2(\lambda + l^2)} e^{2\lambda t} \ . \end{aligned} \tag{5.27}$$

A similar expression can be obtained for $\mathrm{Nu}_\mu(\mathrm{t})$. This shows that the Nusselt numbers initially grow exponentially with a growth rate that is twice that of the fastest growing mode, as mentioned in Section 5.3.1, prior to saturation.

We now consider secondary shearing instabilities on the elevator mode flow. Since $\mathrm{Pr} \ll 1$, we assume that viscosity is negligible. In that case, the growth rate of the shearing modes, σ, is a simple increasing function of the velocity within the fingers. Since the latter increases exponentially, σ also increases, while the growth rate of the fingering instability remains constant. We therefore expect that there will be a time where the two are of the same order of magnitude. At this point, the elevator modes are destroyed and saturation occurs.

As in Radko and Smith (2012), we neglect the temporal variation of the primary elevator modes when evaluating the growth rate of the secondary shearing modes, and simply write

$$\mathbf{U}(x, z) = \hat{w}_{\mathrm{E}} \sin(lx)\hat{\mathbf{z}} \ . \tag{5.28}$$

The growth rate of the shearing instability can be found through dimensional analysis or through Floquet theory (see Brown *et al.*, 2013). The dimensional argument goes as follows: the only relevant velocity in the system is that of fluid within the fingers, $\hat{w}_{\mathrm{E}}$, and the only relevant length scale is $1/l$, so the only relevant growth rate is $\hat{w}_{\mathrm{E}} l$. Hence,

$$\sigma = K \hat{w}_{\mathrm{E}} l \ . \tag{5.29}$$

where K is a universal constant of proportionality. This coefficient should be independent of any system parameter (τ, Pr, R_0) or any property of the mode. All information about the latter is contained in l and $\hat{w}_{\rm E}$.

Assuming that saturation occurs when the shearing instability growth rate is of the order of the fingering growth rate, the saturation condition can be written mathematically as

$$\sigma = K\hat{w}_{\rm E}l = c\lambda, \tag{5.30}$$

where c is independent of the fundamental parameters of the system. This equation uniquely determines the velocity within the fingers at saturation to be

$$\hat{w}_{\rm E} = \frac{C\lambda\sqrt{2}}{l}, \tag{5.31}$$

where $C = c/K\sqrt{2}$ is another universal constant (to be determined). Equation 5.31 is similar to the one used by Denissenkov (2010) to estimate the effective diffusivity by dimensional analysis. Our own model, however, departs from his analysis at this point and produces a result that more accurately fits the results of simulations.

A given elevator mode with velocity $w_{\rm E}(x,z) = \hat{w}_{\rm E}\sin(lx)$ has a temperature profile $T_{\rm E}(x,z) = \hat{T}_{\rm E}\sin(lx)$ with $\hat{T}_{\rm E} = -\frac{\hat{w}_{\rm E}}{\lambda+l^2}$, for the same reasons that led us to Equation 5.25. Hence, at saturation,

$$\mathrm{Nu}_T = 1 + \frac{\hat{w}_{\rm E}^2}{2(\lambda+l^2)} = 1 + C^2\frac{\lambda^2}{l^2(\lambda+l^2)} \tag{5.32}$$

using Equation 5.31. And by similar arguments, it can be shown that

$$\mathrm{Nu}_\mu = 1 + \mathrm{C}^2\frac{\lambda^2}{\tau \mathrm{l}^2\left(\lambda+\tau \mathrm{l}^2\right)}, \tag{5.33}$$

where C is the same constant as in Equation 5.32. The physical interpretation of C is now clearer. Since C is related to the relative importance of the shearing and fingering instabilities at saturation, a large value of C indicates that the shearing instability cannot easily disturb the initial fingers, resulting in a large saturated flux. Conversely, a small value suggests that shear easily destroys fingers, resulting in a small saturated flux.

We compare this prescription with data from this study and those simulations from Traxler *et al.* (2011a) and Radko and Smith (2012) that approach the astrophysical regime. We fit C by using a chi-squared statistical test and find a best fit at $C = 7.0096$. As illustrated in Figure 5.5, for a value of $C = 7$ and $\tau \leq \mathrm{Pr}$, we find that our theoretical results match all simulations remarkably well. It is important to note that, while C needs to be fitted, it is a universal constant and cannot depend on Pr, τ, or R_0. The fact that we are able to use only one value of C to correctly fit the data suggests that the premises of this theory are correct.

In Figure 5.6, we show that in the opposite case ($\tau > \mathrm{Pr}$) we do not observe the same quality of fit. The full analysis by Radko and Smith (2012) is also unable to explain these results, so it is likely that saturation in this case operates somewhat differently. However, since stars are always in the former situation, the poorness of fit of the latter case is interesting but not relevant to our ultimate purpose.

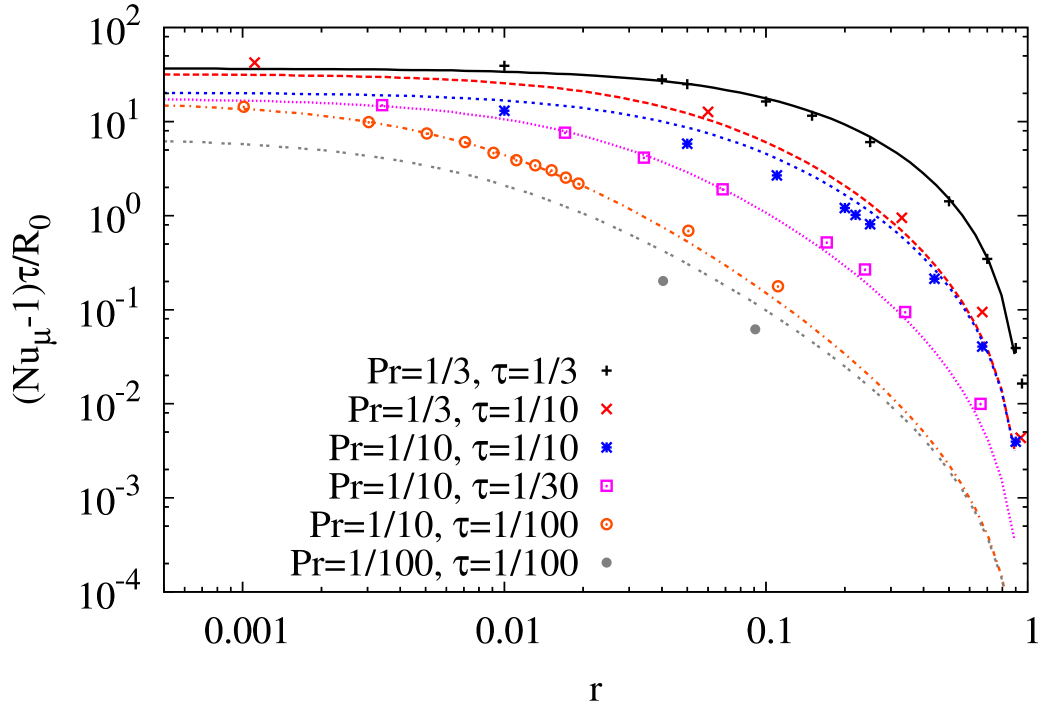


Figure 5.5: Comparison between the turbulent compositional flux observed in simulations (Table 5.1 and references therein), shown as symbols, to our theory (see Equation 5.33), shown as curves, for the simulations with $\tau \leq$ Pr. Since τ is always less than or equal to Pr in astrophysics, the cases considered here are physically relevant. The value of the unknown constant C merely shifts the theoretical model predictions vertically by the same amount for all curves, and we find that the best fit has $C \approx 7$. Using this value, we find that the measured flux in simulations fits the model remarkably well. We rescale the fluxes by τ/R_0 to separate the theoretical curves for ease of viewing.

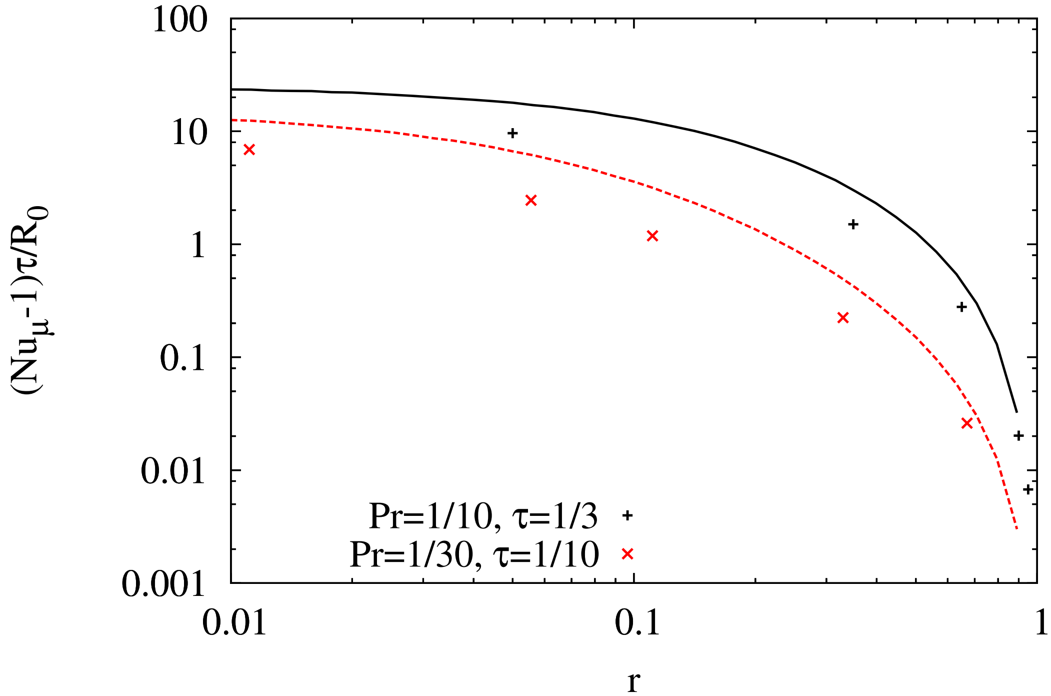


Figure 5.6: Same as Figure 5.5 but for cases with $\tau > \mathrm{Pr}$. The same quality of fit is not observed in these simulations, and the difference between the data and the model is of the order of a few. We rescale the fluxes by τ/R_0 to be consistent with Figure 5.5.

The Gamma-instability

Radko (2003) developed the γ-instability theory to explain the formation of thermohaline staircases in fingering regions of the ocean. The theory was found to agree with direct numerical simulations of staircase formation in two and three dimensions (Radko, 2003; Stellmach *et al.*, 2011) and has since become accepted in the physical oceanography community.

The original γ-instability theory is a linear mean-field instability theory. Radko (2003) first averaged the governing equations of fingering convection (see Equations 5.6 through 5.9) spatially, over length scales that span several fingers, and temporally, over multiple finger lifetimes. He then performed a linear stability analysis of the averaged, or mean-field, equations. He argued that the stability of the system depends on the quantity γ_{turb}, defined as the ratio of the turbulent thermal flux, $\text{Nu}_{\text{T}} - 1$, to the turbulent compositional flux, $\tau(\text{Nu}_\mu - 1)/\text{R}_0$. More specifically, he showed that homogeneous fingering convection is unstable to layering if and only if $\partial\gamma_{\text{turb}}/\partial R_0 < 0$. Traxler *et al.* (2011a) later improved this theory and showed that in the astrophysical case, this result still holds as long as γ is redefined as the ratio of the total fluxes, diffusive and turbulent;[1] in other words, a system is unstable to layering if and only if

$$\frac{\partial\gamma}{\partial R_0} < 0 \text{ with } \gamma = \frac{R_0\text{Nu}_{\text{T}}}{\tau\text{Nu}_\mu}. \tag{5.34}$$

[1] The discussion of Denissenkov (2010) on layer formation uses γ_{turb} rather than γ, which Traxler *et al.* (2011a) showed to be incorrect for the astrophysical case.

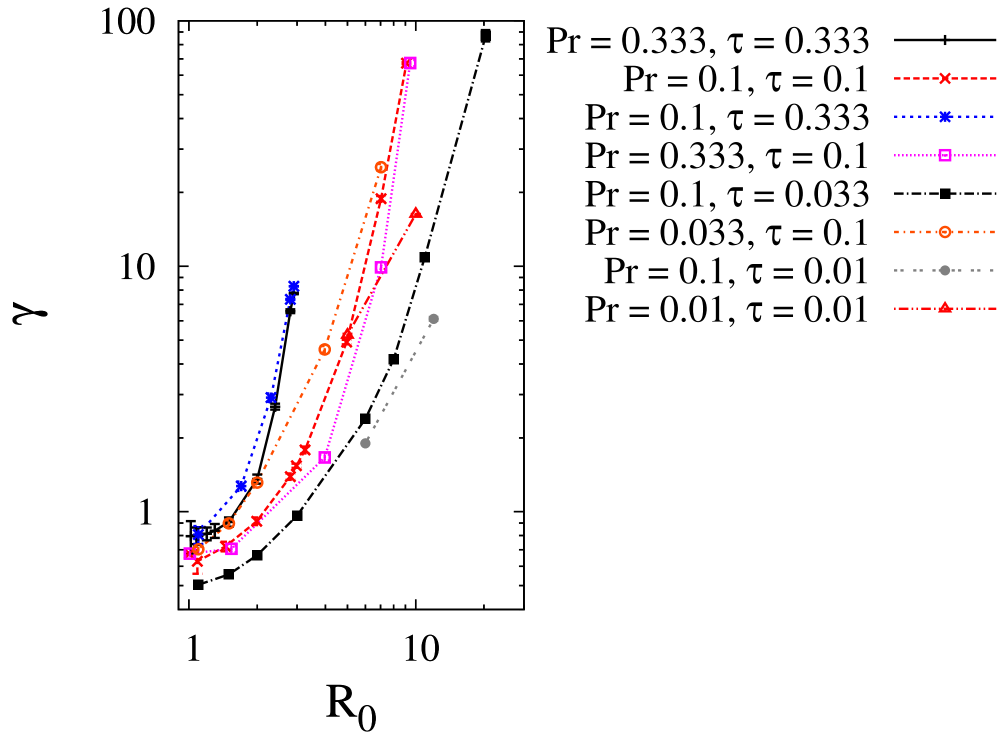


Figure 5.7: The total flux ratio, γ, using the same data as in Figure 5.4. Note how γ increases monotonically with R_0, which implies that thermocompositional staircases cannot form by the γ-instability.

Some of our simulations at very low r exhibit properties consistent with layering. In order to determine whether these layers form via the γ-instability, we plot in Figure 5.7 the variation of γ with R_0 for various values of Pr and τ. Unlike γ_{turb}, which does decrease for some r at low Pr, τ, the quantity γ always seems to increase. As discussed by Traxler *et al.* (2011a), this is because Nu_{T} and Nu_{μ} are both dominated by diffusive fluxes at low Pr and τ. The diffusive flux ratio, R_0/τ, increases strongly with R_0 because τ is asymptotically small and R_0 approaches $1/\tau$ as the system becomes increasingly stable. This effect dominates the total flux ratio. Since γ is always an increasing function of R_0, we conclude that the observed layers cannot be forming by the γ-instability. Thus, while the γ-instability can lead to the development of staircases in high Pr oceanic fingering, it is unlikely to be relevant for astrophysical fingering, confirming the results of Traxler *et al.* (2011a).

Simulations with Low Density Ratio

Despite the fact that γ is a strictly increasing function of r, we do observe layer formation in some simulations. We now study these results in more detail, focusing on the case presented in Figure 5.3, which has $\mathrm{Pr} = 1/10$, $\tau = 1/30$, and $R_0 = 1.1$ and shows strong evidence for the formation of a convective layer. Figure 5.8 shows a snapshot of the compositional perturbation at $t = 825$; large-scale plumes characteristic of overturning convection are clearly visible. We can check to see whether these plumes are associated with an inversion of the horizontally-averaged total density

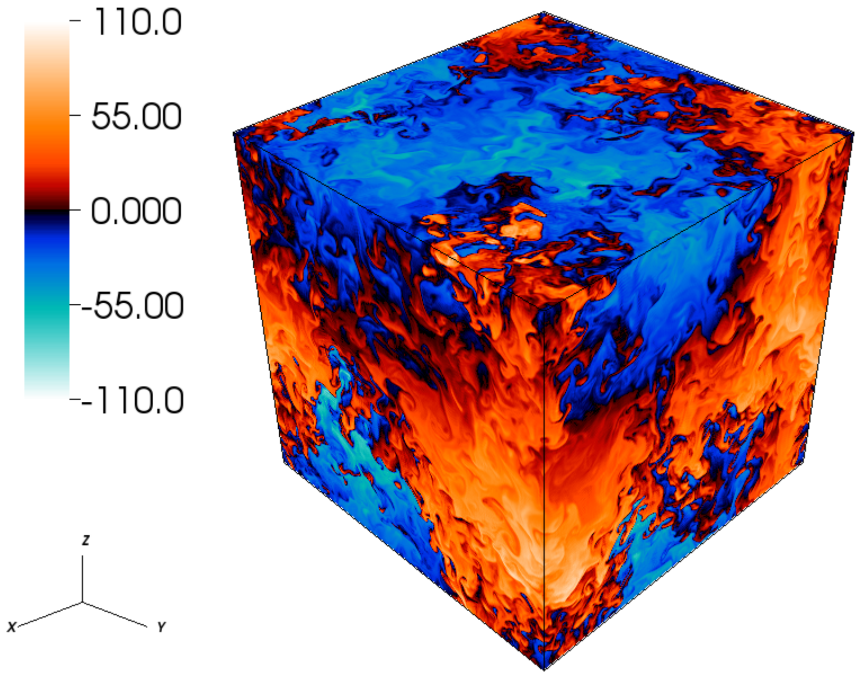


Figure 5.8: A snapshot of the compositional perturbation from the same simulation as in Figure 5.3 (Pr $= 1/10$, $\tau = 1/30$, and $R_0 = 1.1$) at $t = 825$. Large convective plumes are clearly visible and span a significant fraction of the compositional domain. The interface is harder to identify but can be seen near the bottom of the plot, where finger structures are still apparent. Note that it is far from "flat," and instead is highly distorted by the large-scale plumes in the convective layer.

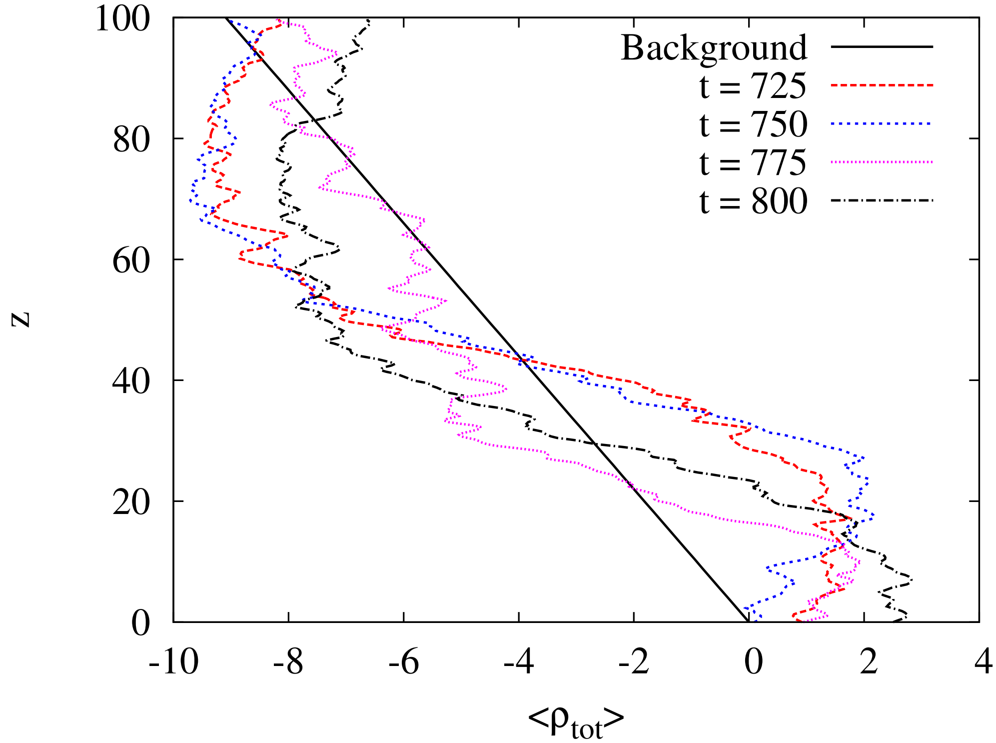


Figure 5.9: The horizontally-averaged density profile, $\langle \rho_{\text{tot}} \rangle$ (see Equation 5.35) as a function of z for the simulation shown in Figure 5.3 at the four times marked by vertical dashed lines. The background density is shown for reference as a solid line and is always decreasing with height. The sharp density transition in the lower part of the domain is the interface. At times $t = 725$, 750, and 800, there exist regions over which the density increases with height, which we call a density inversion. The only time without an inversion, $t = 775$, corresponds to the point where the Nusselt number is at its minimum (see Figure 5.3).

profile, which is given by

$$\langle \rho_{\rm tot} \rangle = -(1 - R_0^{-1})z - \langle T \rangle + \langle \mu \rangle. \tag{5.35}$$

Figure 5.9 shows $\langle \rho_{\rm tot} \rangle$ as a function of height for the simulation shown in Figure 5.3 at four different times selected during the later stages, where the Nusselt numbers increase significantly. The solid line indicates the linear background gradient for reference. Note how each of the four density profiles shows a sharp, stably stratified transition region between $z = 25$ and $z = 45$, the interface. Three of the four profiles also have a region where the total density increases with z, the convective layer. This demonstrates that the homogeneous fingering has indeed transitioned into a layered system with convective layers separated by stable interfaces. At $t = 725$, shown in Figure 5.9, the convective layer is just beginning to develop, and a slight inversion of the density profile appears in Figure 5.9. As the simulation approaches the stage of maximum transport at $t = 750$ (see Figure 5.3), the density inverts more strongly. The ensuing period of active mixing thickens the interface again and flattens the density profile. At $t = 775$, the convective mixing briefly shuts down. This process repeats, as can be seen at $t = 800$, when the density profile once is beginning to invert. Note that the interface appears to move up and down during this period, which we attribute to the advection of a large quantity of compositionally dense material by convective plumes.

This is not the only such simulation in which staircases form: this behavior exists at low r ($\lesssim 0.003$) for several values of Pr and τ, including $\mathrm{Pr} = 1/10$, $\tau = 1/10$ and

$\mathrm{Pr} = 1/3$, $\tau = 1/10$. In all these cases, the $\mathrm{Nu_T}$ and Nu_μ time series as well as the properties of the layered state are qualitatively similar to the one shown in Figure 5.3, Figure 5.8, and Figure 5.9.

Having established that this layering transition cannot be due to the γ-instability, we look instead at the original mechanism proposed by Stern (1969) (see also Stern *et al.*, 2001) for the formation of thermohaline staircases in the ocean. In addition to the γ-instability, fingering convection is also known to excite large-scale gravity waves through another mean-field instability, the "collective instability." Stern (1969) studied this effect, and showed that large-scale gravity waves can be excited under some circumstances and grow in amplitude exponentially. These waves transport momentum, heat, and heavy elements until they begin to "break," at which point the advected material mixes with the background. He went on to suggest that if the waves break on large enough scales, this could cause enough local mixing to trigger the formation of staircases.

In related numerical simulations at oceanographically relevant parameters ($\mathrm{Pr} = 7$, $\tau = 0.3$), Stellmach *et al.* (2011) found that gravity waves are indeed excited by the collective instability but do not trigger layers. Layer formation in their simulations—and thus presumably in the ocean—is caused by the γ-instability rather than the collective instability. Our simulations suggest, however, that the converse may be true in the astrophysical parameter regime. As discussed in Figure 5.3, gravity waves are observed to dominate the dynamics of the system between $t = 100$ and $t = 600$.

Furthermore, the gradual increase in $\mathrm{Nu_T}$ and Nu_μ during that phase suggests that the waves have high enough amplitudes to interact nonlinearly with each other, and with the background, by contrast with the oceanographic case. Indeed, linear waves cause oscillations in $\mathrm{Nu_T}$ and Nu_μ without changing their mean values—only nonlinear waves can affect the mean. Since the waves in Figure 5.3 are strong enough to interact nonlinearly, they may also break and cause mixing on larger scales, following the original hypothesis proposed by Stern (1969).

Checking this hypothesis in detail is difficult with the available simulations. However, we can at least determine whether the observed waves do indeed appear in accordance with the collective instability theory. A system is unstable to the collective instability if the Stern number, in our notation defined as

$$A = \frac{(\mathrm{Nu_T} - 1)\left(\gamma_{\mathrm{turb}}^{-1} - 1\right)}{\mathrm{Pr}\left(1 - \mathrm{R}_0^{-1}\right)}, \tag{5.36}$$

exceeds order unity (Stern *et al.*, 2001). Note that this expression uses $\gamma_{\mathrm{turb}} = R_0(\mathrm{Nu_T} - 1)/\tau(\mathrm{Nu}_\mu - 1)$ instead of γ. For the simulation shown in Figure 5.3 and Figure 5.9, $\mathrm{Nu_T} - 1 = 6.55$ and $\gamma_{\mathrm{turb}} = 0.438$, so that the Stern number is 924. This implies that the system is indeed unstable to the collective instability.

More generally, we find that all the simulations in which we observe layer formation have Stern numbers of the order of 10^3 or more. However, not all such simulations go into layers, as can be seen in Figure 5.10, so it is likely that the condition for layer formation through the collective instability also depends on the value of r. It is also possible that these gravity waves grow in all simulations with $A > 1$, but since the

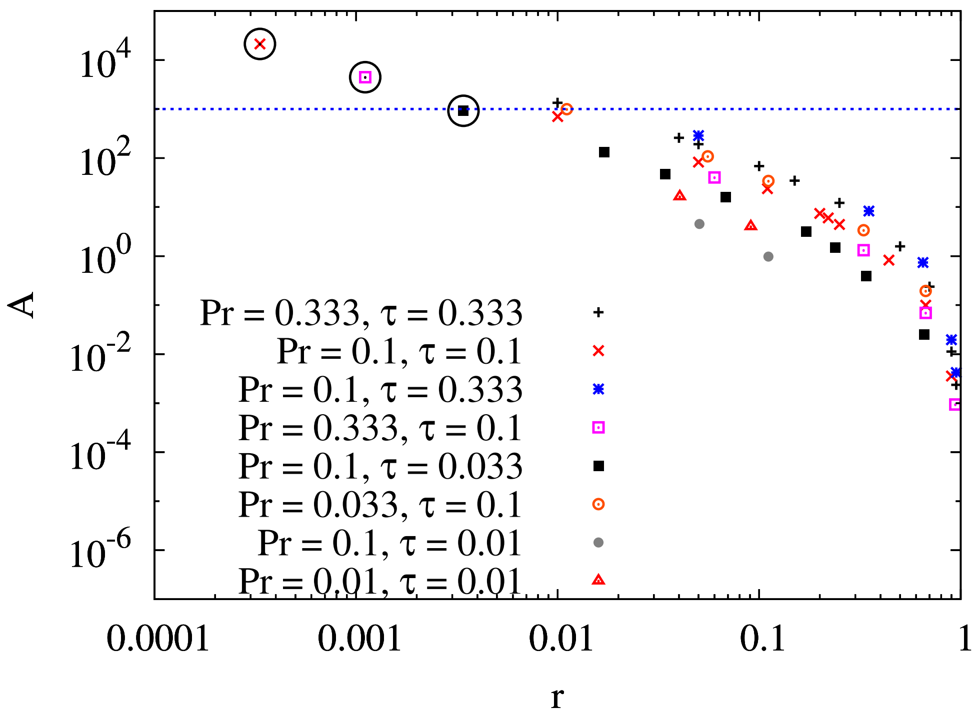


Figure 5.10: Stern number as a function of r, calculated from the data presented in Figure 5.4 (and Table 5.1 using equation Equation 5.36). Simulations that develop into layers are denoted with a large black circle. A Stern number exceeding order unity suggests the possibility of triggering of gravity wave growth but doesn't necessarily imply layer formation. Here we see that the only cases in which layers form occur when $A \gtrsim 10^3$, which is marked with a horizontal dotted line.

growth rate depends on the Stern number, gravity waves may be too weak to detect if A is too low. If this is the case, it will be difficult to see layer formation in models with smaller A because these simulations are expensive and difficult to integrate for long.

Much work still needs to be done to identify the parameters for which we expect layer formation in the astrophysical regime and to quantify transport in the layered case (see Wood *et al.* (2013) for related efforts in the case of semi-convection). We have only three simulations that develop into layers and thus require a more extensive sample of simulations at low r in order to better understand the conditions of layer formation and their properties as a function of Pr, τ, and r. In every simulation that exhibits layers, we find that a single convective plume spans the majority of the domain and do not see the development of multiple layers, as normally seen in simulations of the fingering regime in the oceanographic case (e.g., Stellmach *et al.*, 2011) and in the case of semi-convection (Rosenblum *et al.*, 2011; Mirouh *et al.*, 2012). To characterize this process, we must increase the size of the domain and extend the integration time to let the layers develop more naturally. This will allow us to describe more completely the size scales and transport of thermocompositional layers in astrophysical fingering convection.

5.5 Conclusions

We have developed a simple semi-analytical model for the vertical transport of heat and composition by homogeneous fingering convection that reproduces the results of numerical simulations from this study and from previous studies and can easily be implemented in a stellar evolution code. We have found that, under conditions relevant for stellar interiors, the turbulent heat transport is negligible, but compositional transport can be important. In particular, this model predicts more efficient transport for weakly stratified systems than was initially suggested in previous work by Traxler *et al.* (2011a).

We have also found the first evidence for layer formation by fingering convection in the astrophysical parameter regime. We have identified the collective instability as the origin of layer formation. By contrast, in the oceanographic and semi-convective cases, the γ-instability causes layer formation (e.g., Stellmach *et al.*, 2011; Mirouh *et al.*, 2012). Much work still needs to be done to characterize the conditions for layer formation and transport by layered convection in this case. With our limited computational resources, we have found that layer formation appears to require high Stern numbers (see Equation 5.36) to occur. If true, this would imply that layers are only likely to form in regions that are very close to being fully convective anyway. However, it remains to be seen whether lower Stern numbers can also result in layer formation. To study this will require additional simulations and longer integration times than have been covered in this paper. In addition to the conditions of layer formation, it is also important to develop a theoretical or empirical model for the transport of this case. Our simulations show that transport increases significantly when layers form. Concurrent work by Wood *et al.* (2013) reveals that thermal and compositional transport in the layered semi-convective case scales with the layer height to the power of 4/3. If this scaling also applies here, we may expect that very significant transport rates could occur for large enough layer heights. What determines the latter, however, will also require new theories.

Chapter 6

Prospects

The effects of fluid dynamics in stellar evolution remain an elusive part of astronomy; however, numerical simulations are beginning to bring light to these processes. Some potential advances are listed here, which are just becoming possible either because of recent advances in analytics or in computational power.

6.1 Fingering Convection

The major questions regarding astrophysical fingering convection are to do with the physical nature of the process, the quantitative mixing induced, and the effects on stellar evolution. The physical nature of the process, as discussed in Chapter 6.1, could either be homogenous turbulence or coherent convective layers. The work of Medrano *et al.* (2014) showed that layer formation in most stellar parameter regimes is unlikely. Thus, fingering convection is likely well described through the description of homogeneous turbulence discussed in Chapter 6.1. That being said, a true realization of fingering convection in stars may be complicated by rotation and gravity waves excited by convection. Simulations including such processes are now possible and their results could determine if any corrections may be needed for implementation into stellar codes.

With such results as of yet unavailable, the current homogeneous prescription for fingering convection has been compared to various facets of stellar observations. Explanations of the effects on planetary infall events have been mostly successful. Garaud (2011) indicated that infall events would be mixed too quickly via thermohaline convection to result in an observable excess of metallicity. Deal *et al.* (2015) used the model described here to successfully explain lithium depletion in a star after a suspected planetary infall event.

Another problem has also received recent attention: that of the development of an iron layer in the envelope of a star due to competing effects of gravitational settling and radiation pressure. Richard *et al.* (2001) initially discussed the possibility of such a layer becoming strong enough to incite convection, which could have asteroseismic consequences. Théado and Vauclair (2009) noted that such a layer would first become thermohaline unstable. Zemskova *et al.* (2014) simulated this region in a thin domain then applied the thermohaline mixing discussed here to this problem. They found that this prescription for thermohaline convection reproduced their simulations and that convection is indeed a possible outcome of such a layer developing. Asteroseismic observations of these stars could prove enlightening to the understanding of the nature of this iron layer.

The peculiar abundances of red giants—once attributed to fingering convection—have not been solved by these advances. Denissenkov (2010) found that extreme fingering convection could provide the additional mixing necessary to explain these abundances. However, Brown *et al.* (2013) showed that the predicted mixing of homogenous fingering convection would not be sufficient. As mentioned in Denissenkov and Merryfield (2011), toroidal magnetic fields may be substantial at the location of fingering convection in these stars. In addition, nuclear burning is also relevant near this region. Magnetohydrodynamics simulations of fingering convection or hydrodynamics simulations of fingering convection with nuclear fusion could provide additional insight to this problem.

6.2 Semi-Convection

The main goal of studies of astrophysical semi-convection remains to calculate accurate fluxes of composition and energy in stellar contexts and use these fluxes to determine observable properties of stellar evolution. The effects of semi-convection remain important in many aspects of stellar evolution but continue to be particularly important in massive stars. A number of major roadblocks continue to plague the field, including determining the correct physical behavior of these regions.

The physical nature of semi-convection has been thought to take the form of overstable oscillations for much of its history in stellar evolution; now, however, the correct picture seems to be stacks of convective layers. The simulations of Wood *et al.* (2013) provided flux prescriptions for these layers as a function of the layer height. This height remains unknown as the layers in these simulations merge over time until they fill the entire simulated domain. Future numerical work of such systems could elucidate the merger timescale and final properties of these layers, which would provide a much needed closure of these prescriptions for their use in stellar evolution codes.

As with fingering convection, semi-convection in the presence of rotation and strong gravity waves could be important. In particular, since semi-convection can occur near regions unstable to convection, gravity wave excitations may prove critical to the discussion of semi-convection in stars. The stability of layered semi-convection in the presence of such phenomena would serve to determine whether layered semi-convection is the appropriate realization of such regions in stars.

6.3 Convective Overshoot

As with semi-convection, the main goal of studying astrophysical convective overshoot is to determine the precise mixing properties of such regions within stellar contexts. Also like semi-convection, one of the main problems currently facing convective overshoot is that the nature of the process remains ambiguous, with the simulations of Meakin and Arnett (2007) suggesting entrainment of material into the convection zone and the simulations of Brummell *et al.* (2002) suggesting only slight mixing. Of course, stellar interiors may see the consequences of both realizations for different types of convection zones, as is mentioned by Viallet *et al.* (2015); however, much more work must be done before stellar astrophysicists can grasp what is occuring in these regions of stars.

Pushing the simulations of Brummell *et al.* (2002) to lower Pr, higher Ra, and higher resolution may provide some additional enlightenment. The inclusion of new physics into convective overshoot simulations would also complement the literature, such as convection driven by temperature- and composition-dependent opacities and by nuclear burning. Such direct numerical simulations are feasible and more physically comparable to the problem at hand. Initial simulations in the Boussinesq case would be valuable to gain physical insight before fully compressible simulations are attempted.

Additionally, a series of simulations similar to those of Brummell *et al.* (2002) but for outward-moving overshooting convection and/or spherical geometries would also be of substantial interest. The interest in outward-moving convective overshoot stems from core convection in massive stars, which is significant to their evolution. Many calculations (e.g., Hurlburt *et al.*, 1986) have shown that upward convective overshoot has different properties. In addition, because these convection zones are quite large, the geometric effects of the spherical problem are non-negligible, particularly for convective overshoot from the cores of massive stars and the envelopes of red supergiants.

6.4 Massive Star Evolution

Several major problems still exist in the context of massive star evolution. In particular, the inability to understand quantitatively the envelope behavior of supergiants remains a major issue. The more detailed understanding of the internal behavior from this work will assist in mitigating this problem; however, the uncertainties in mixing processes still contribute substantially. The supernova mechanism is another tricky area of stellar evolution, and much recent work has been devoted to determining the exact nature of supernovae through numerical simulations. The results of such multi-dimensional studies should provide answers regarding which stars actually explode. This will provide more accurate nucleosynthesis than the use of one-dimensional prescriptions, as was done in the work here. However, accurate pre-supernova models are necessary as starting conditions of these explosions.

With recent advances on the effects of minor mixing processes on the fluxes within stars, it would be prudent to complete a new set of massive star calculations to replace Woosley and Heger (2007) with pre-supernova models with more realistic physics. Such a new grid would have implications on many aspects of stellar evolution including more accurate supernova progenitors, nucleosynthesis, and remnant distributions. Given that the numerical tools have been extensively developed for the studies contained in Chapter 2 and in Sukhbold *et al.* (2016), a cursory examination would not be difficult. These new progenitor models could also be used in supernova explosion calculations to determine whether these substantive changes in the interiors of the stars could have any major impacts on supernova light curves.

Also of substantial interest with regards to stellar evolution are the new developments on the blue and red supergiant transition. A new grid of stellar models would allow for revisiting the problem of the blue-to-red supergiant ratio in galaxies to follow up on the work of Langer and Maeder (1995). In addition, this new formalism should allow for readdressing the question of SN 1987A, which was expected to be a red supergiant at core collapse but was instead blue. In particular, it would appear that the understanding of semi-convection and overshooting convection likely rule out most of the current single star models for the formation of the SN 1987A progenitor. Additional studies of the global effects of metallicity and rotation in the presence of these processes are still needed to completely rule out this possibility.

Much work is yet to be done in the field of massive stellar evolution. Hopefully, the models described here in addition to newer prescriptions from numerical simulations will begin to remove some of the parameterizations used in stellar evolution and eventually allow the field to come to a consensus on the appropriate physics that govern stars.

Chapter 7

Appendices

.1 The Superseded Implementation of the Semi-Convection Model

Due to the inverse dependence of $l_{\rm osht}$ on Pe, some numerical smoothing is needed to limit the maximum extent of the overshooting convection as Pe becomes small. The following expression was chosen to prevent mixing that would occur at a rate faster than the material could propagate through the star:

$$l_{\rm osht}^{\rm new} = l_{\rm osht} \frac{v_{\rm osht}}{v_{\rm osht} + l_{\rm osht}/\Delta t}, \tag{.1}$$

where Δt is the timestep taken. This reduces to the correct expression for $l_{\rm osht}$ when the overshoot length is not unreasonably large. This expression was used in the models run in this work; however, the new prescription described in Section 3.3.1 is more physical, and all future models will be run using that model.

Bibliography

Arlandini, C.; Käppeler, F.; Wisshak, K.; Gallino, R.; Lugaro, M.; Busso, M. and Straniero, O. Neutron Capture in Low-Mass Asymptotic Giant Branch Stars: Cross Sections and Abundance Signatures. *The Astrophysical Journal* **525**, 886 (1999).

Arnett, W. Supernova 1987A. *Annual Review of Astronomy and Astrophysics* **27**, 629 (1989).

Arnett, W. D. Analytic solutions for light curves of supernovae of Type II. *The Astrophysical Journal* **237**, 541 (1980).

Baines, P. G. G. and Gill, A. E. On thermohaline convection with linear gradients. *Journal of Fluid Mechanics* **37**, 289 (1969).

Basu, S. Seismology of the base of the solar convection zone. *Monthly Notices of the Royal Astronomical Society* **288**, 572 (1997).

Batchelor, G. K. Flow at Large Reynolds Number: Effects of Viscosity. In *An*

Introduction to Fluid Dynamics, Cambridge University Press, Cambridge, pp. 264–377 (2000).

Baturin, V. A. and Mironova, I. V. Smooth models of overshooting at the base of the solar convective zone. *Astrophysics and Space Science* **328**, 265 (2010).

Bertelli, G.; Bressan, A. and Chiosi, C. Models with convective overshoot: Tables of isochrones. *Astrophysics and Space Science* **169**, 105 (1990).

Bertelli, G.; Chiosi, C. and Bressan, A. G. Mass loss and overshooting in massive stars. *Astronomy & Astrophysics* (1981).

Brown, J. M.; Garaud, P. and Stellmach, S. Chemical Transport and Spontaneous Layer Formation in Fingering Convection in Astrophysics. *The Astrophysical Journal* **768**, 16 (2013).

Brown, J. M. and Woosley, S. E. NUCLEOSYNTHETIC CONSTRAINTS ON THE MASS OF THE HEAVIEST SUPERNOVAE. *The Astrophysical Journal* **769**, 99 (2013).

Brummell, N. H.; Clune, T. L. and Toomre, J. Penetration and Overshooting in Turbulent Compressible Convection. *The Astrophysical Journal* **570**, 825 (2002).

Buchmann, L. New Stellar Reaction Rate for 12C(alpha , gamma) 16O. *The Astrophysical Journal* **468**, L127 (1996).

Buonanno, R.; Corsi, C. E. and Pecci, F. F. The giant, asymptotic and horizontal branches of globular clusters. II. Photographic of the metal-poor clusters M 15, M 92 and NGC 5466. *Astronomy & Astrophysics* (1985).

Calzavarini, E.; Lohse, D.; Toschi, F. and Tripiccione, R. Rayleigh and Prandtl number scaling in the bulk of Rayleigh–Bénard turbulence. *Physics of Fluids* **17**, 055107 (2005).

Canuto, V. M. Stellar mixing. *Astronomy & Astrophysics* **528**, A76 (2011).

Carpenter, J. R.; Sommer, T. and Wüest, A. Simulations of a double-diffusive interface in the diffusive convection regime. *Journal of Fluid Mechanics* **711**, 411 (2012).

Castellani, V.; Chieffi, A.; Tornambe, A. and Pulone, L. Helium-burning evolutionary phases in population II stars. I Breathing pulses in horizontal branch stars. *The Astrophysical Journal* **296**, 204 (1985).

Castellani, V.; Giannone, P. and Renzini, A. Induced Semi-Convection in Helium-Burning Horizontal-Branch Stars II. *Astrophysics and Space Science* **10**, 355 (1971).

Cayrel, R.; Depagne, E.; Spite, M.; Hill, V.; Spite, F.; François, P.; Plez, B.; Beers, T.; Primas, F.; Andersen, J.; Barbuy, B.; Bonifacio, P.; Molaro, P. and Nordström,

B. First stars V - Abundance patterns from C to Zn and supernova yields in the early Galaxy. *Astronomy & Astrophysics* **416**, 1117 (2004).

Chabrier, G. The Galactic Disk Mass Function: Reconciliation of the Hubble Space Telescope and Nearby Determinations. *The Astrophysical Journal* **586**, L133 (2003).

Charbonnel, C. Clues for non-standard mixing on the red giant branch from 12C/13C and 12C/14N ratios in evolved stars. *Astronomy & Astrophysics* **282**, 811 (1994).

Charbonnel, C. and Zahn, J. P. Thermohaline mixing: a physical mechanism governing the photospheric composition of low-mass giants. *Astronomy & Astrophysics* **467**, L15 (2007).

Christensen-Dalsgaard, J.; Monteiro, M. J. P. F. G.; Rempel, M. and Thompson, M. J. A more realistic representation of overshoot at the base of the solar convective envelope as seen by helioseismology. *Monthly Notices of the Royal Astronomical Society* **414**, 1158 (2011).

Clayton, D. D. *Principles of Stellar Evolution and Nucleosynthesis*, vol. 163. McGraw-Hill, New York (1969).

Deal, M.; Richard, O. and Vauclair, S. Accretion of planetary matter and the lithium problem in the 16 Cygni stellar system. *Astronomy & Astrophysics* **584**, A105 (2015).

Deng, L.; Chiosi, C. and Bressan, A. Stellar evolution with turbulent diffusion. I. A new formalism of mixing. *Astronomy & Astrophysics* (1996).

Denissenkov, P. A. NUMERICAL SIMULATIONS OF THERMOHALINE CONVECTION: IMPLICATIONS FOR EXTRA-MIXING IN LOW-MASS RGB STARS. *The Astrophysical Journal* **723**, 563 (2010).

Denissenkov, P. A. and Merryfield, W. J. Thermohaline Mixing: Does It Really Govern the Atmospheric Chemical Composition of Low-Mass Red Giants? *The Astrophysical Journal Letters* **727**, L8 (2011).

Dohm-Palmer, R. C. and Skillman, E. D. The Ratio of Blue to Red Supergiants in Sextans A from [ITAL]Hubble Space Telescope[/ITAL] Imaging. *The Astrophysical Journal* **123**, 1433 (2002).

Doom, C.; de Greve, J. P. and de Loore, C. Stellar evolution in the upper Hertzsprung-Russell diagram. *The Astrophysical Journal* **303**, 136 (1986).

Eggleton, P. P. Composition Changes During Stellar Evolution. *Monthly Notices of the Royal Astronomical Society* **156**, 361 (1972).

Eggleton, P. P.; Dearborn, D. S. P. and Lattanzio, J. C. Deep Mixing of He-3: Reconciling Big Bang and Stellar Nucleosynthesis. *Science* **314**, 1580 (2006).

Ekström, S. and others. Grids of stellar models with rotation - I. Models from 0.8 to

120 Msun at solar metallicity (Z = 0.014). *Astronomy & Astrophysics* **537**, A146 (2012).

Fischer, D. A. and Valenti, J. The Planet-Metallicity Correlation. *The Astrophysical Journal* **622**, 1102 (2005).

Freytag, B.; Ludwig, H. and Steffen, M. Hydrodynamical models of stellar convection. *Astronomy & Astrophysics* **313**, 497 (1996).

Fryer, C. L. Mass Limits For Black Hole Formation. *The Astrophysical Journal* **522**, 413 (1999).

Fryer, C. L. and Heger, A. Core-Collapse Simulations of Rotating Stars. *The Astrophysical Journal* **541**, 1033 (2000).

Garaud, P. What Happened to the Other Mohicans? The Case for a Primordial Origin to the Planet-Metallicity Connection. *The Astrophysical Journal Letters* **728**, L30 (2011).

Garaud, P.; Ogilvie, G. I.; Miller, N. and Stellmach, S. A model of the entropy flux and Reynolds stress in turbulent convection. *Monthly Notices of the Royal Astronomical Society* **407**, 2451 (2010).

Gilroy, K. K. Carbon isotope ratios and lithium abundances in open cluster giants. *The Astrophysical Journal* **347**, 835 (1989).

Gräfener, G.; Owocki, S. P. and Vink, J. S. Stellar envelope inflation near the Eddington limit. Implications for the radii of Wolf-Rayet stars and luminous blue variables. *Astronomy & Astrophysics* **538**, A40 (2012).

Heger, A.; Fryer, C. L.; Woosley, S. E.; Langer, N. and Hartmann, D. H. How Massive Single Stars End Their Life. *The Astrophysical Journal* **591**, 288 (2003).

Herald, D.; McNaught, R. H.; Morel, M.; Madore, B.; Shelton, I.; Claria, J.; Davidsen, A.; Kimble, R.; Gregg, M.; Sonneborn, G. and Kirshner, R. Supernova 1987A in the Large Magellanic Cloud. *IAU Circ.* **4317**, 1 (1987).

Herwig, F. The evolution of AGB stars with convective overshoot. *Astronomy & Astrophysics* **360**, 952 (2000).

Horiuchi, S.; Beacom, J. F.; Kochanek, C. S.; Prieto, J. L.; Stanek, K. Z. and Thompson, T. A. The Cosmic Core-collapse Supernova Rate Does Not Match the Massive-star Formation Rate. *arXiv.org* **738**, 154 (2011).

Humphreys, R. M. and McElroy, D. B. The initial mass function for massive stars in the Galaxy and the Magellanic Clouds. *The Astrophysical Journal* **284**, 565 (1984).

Hurlburt, N. E.; Toomre, J. and Massaguer, J. M. Nonlinear compressible convection penetrating into stable layers and producing internal gravity waves. *The Astrophysical Journal* **311**, 563 (1986).

Hurlburt, N. E.; Toomre, J.; Massaguer, J. M. and Zahn, J.-P. Penetration below a convective zone. *The Astrophysical Journal* **421**, 245 (1994).

Iwamoto, K.; Brachwitz, F.; Nomoto, K.; Kishimoto, N.; Umeda, H.; Hix, W. R. and Thielemann, F.-K. Nucleosynthesis in Chandrasekhar Mass Models for Type IA Supernovae and Constraints on Progenitor Systems and Burning-Front Propagation. *The Astrophysical Journal Supplement* **125**, 439 (1999).

Jaeger, M.; Kunz, R.; Mayer, A.; Hammer, J. W.; Staudt, G.; Kratz, K. L. and Pfeiffer, B. ^{22}Ne(α,n)^{25}Mg: The Key Neutron Source in Massive Stars. *Physical Review Letters* **87**, 202501 (2001).

Kappeler, F.; Beer, H. and Wisshak, K. s-process nucleosynthesis-nuclear physics and the classical model. *Reports on Progress in Physics* **52**, 945 (1989).

Kasen, D. and Woosley, S. E. Type II Supernovae: Model Light Curves and Standard Candle Relationships. *arXiv.org* **703**, 2205 (2009).

Kato, S. Overstable Convection in a Medium Stratified in Mean Molecular Weight. *Publications of the Astronomical Society of Japan* (1966).

Kippenhahn, R.; Ruschenplatt, G. and Thomas, H. C. C. The time scale of thermohaline mixing in stars. *Astronomy & Astrophysics* **91**, 175 (1980).

Kunze, E. A review of oceanic salt-fingering theory. *Progress in Oceanography* **56**, 399 (2003).

Langer, N. Oxygen Isotopes from Massive Stars. In A.~Burkert, D.~H.~Hartmann, &. S., ed., *The History of the Milky Way and Its Satellite System* (1996), p. 169.

Langer, N.; Fricke, K. J. and El Eid, M. F. Evolution of massive stars with semiconvective diffusion. *Astronomy & Astrophysics* **145**, 179 (1985).

Langer, N. and Maeder, A. The problem of the blue-to-red supergiant ratio in galaxies. *Astronomy & Astrophysics* (1995).

Langer, N.; Sugimoto, D. and Fricke, K. J. Semiconvective diffusion and energy transport. *Astronomy & Astrophysics* **126**, 207 (1983).

Latour, J.; Toomre, J. and Zahn, J. P. Stellar convection theory. III - Dynamical coupling of the two convection zones in A-type stars by penetrative motions. *The Astrophysical Journal* **248**, 1081 (1981).

Lattimer, J. M. The Nuclear Equation of State and Neutron Star Masses. *Annual Review of Nuclear and Particle Science* **62**, 485 (2012).

Lattimer, J. M. and Prakash, M. Neutron Star Structure and the Equation of State. *The Astrophysical Journal* **550**, 426 (2001).

Lauterborn, D.; Refsdal, S. and Weigert, A. Stars with Central Helium Burning and the Occurrence of Loops in the H-R Diagram. *Astronomy & Astrophysics* **10**, 97 (1971).

Li, W.; Leaman, J.; Chornock, R.; Filippenko, A. V.; Poznanski, D.; Ganeshalingam, M.; Wang, X.; Modjaz, M.; Jha, S.; Foley, R. J. and Smith, N. Nearby supernova rates from the Lick Observatory Supernova Search - II. The observed luminosity functions and fractions of supernovae in a complete sample. *Monthly Notices of the Royal Astronomical Society* **412**, 1441 (2011).

Limongi, M. and Chieffi, A. The Nucleosynthesis of ^{26}Al and ^{60}Fe in Solar Metallicity Stars Extending in Mass from 11 to 120 M_{solar}: The Hydrostatic and Explosive Contributions. *The Astrophysical Journal* **647**, 483 (2006).

Limongi, M. and Chieffi, A. Presupernova evolution and explosion of massive stars. *Journal of Physics: Conference Series* **202**, 012002 (2010).

Lodders, K. Solar System Abundances of the Elements. Springer Berlin Heidelberg, Berlin, Heidelberg (2010).

Longland, R.; Iliadis, C. and Karakas, A. I. Publisher's Note: Reaction rates for the s-process neutron source ^{22}Ne + α [Phys. Rev. C 85, 065809 (2012)]. *Physical Review C* **86**, 019903 (2012).

Maeder, A. Stellar evolution III: the overshooting from convective cores. *Astronomy & Astrophysics* (1975).

Maeder, A. Stellar yields as a function of initial metallicity and mass limit for black hole formation **264**, 105 (1992).

Maeder, A. and Meynet, G. Grids of evolutionary models from 0.85 to 120M : observational tests and the mass limits. *Astronomy & Astrophysics* (1989).

McQuinn, K. B. W.; Skillman, E. D.; Dalcanton, J. J.; Dolphin, A. E.; Holtzman, J.; Weisz, D. R. and Williams, B. F. OBSERVATIONAL CONSTRAINTS ON RED AND BLUE HELIUM BURNING SEQUENCES. *The Astrophysical Journal* **740**, 48 (2011).

Meakin, C. A. and Arnett, W. D. Turbulent Convection in Stellar Interiors. 1. Hydrodynamic Simulation. *The Astrophysical Journal* **667**, 448 (2007).

Medrano, M.; Garaud, P. and Stellmach, S. DOUBLE-DIFFUSIVE MIXING IN STELLAR INTERIORS IN THE PRESENCE OF HORIZONTAL GRADIENTS. *The Astrophysical Journal* **792**, L30 (2014).

Merryfield, W. J. Hydrodynamics of semiconvection. *The Astrophysical Journal* **444**, 318 (1995).

Meylan, G. and Maeder, A. Comparisons of the HR diagrams of the youngest clusters in the Galaxy, the LMC and SMC - Evidence for a large MS widening. *Astronomy & Astrophysics* **108**, 148 (1982).

Meynet, G.; Chomienne, V.; Ekström, S.; Georgy, C.; Granada, A.; Groh, J.; Maeder, A.; Eggenberger, P.; Levesque, E. and Massey, P. Impact of mass-loss on the evolution and pre-supernova properties of red supergiants. *Astronomy & Astrophysics* **575**, A60 (2015).

Mirouh, G. M.; Garaud, P.; Stellmach, S.; Traxler, A. L. and Wood, T. S. A NEW MODEL FOR MIXING BY DOUBLE-DIFFUSIVE CONVECTION (SEMI-CONVECTION). I. THE CONDITIONS FOR LAYER FORMATION. *The Astrophysical Journal* **750**, 61 (2012).

Moll, R.; Garaud, P. and Stellmach, S. A New Model for Mixing by Double-diffusive Convection (Semi-convection). III. Thermal and Compositional Transport through Non-layered ODDC. *The Astrophysical Journal* **823**, 33 (2016).

Moore, K. and Garaud, P. Main Sequence Evolution with Layered Semiconvection. *The Astrophysical Journal* **817**, 54 (2016).

Musman, S. Penetrative convection. *Journal of Fluid Mechanics* **31**, 343 (1968).

Muthsam, H. J.; Goeb, W.; Kupka, F.; Liebich, W. and Zoechling, J. A numerical study of compressible convection. *Astronomy & Astrophysics* **293**, 127 (1995).

Noguchi, T. and Niino, H. Multi-layered diffusive convection. Part 1. Spontaneous layer formation. *Journal of Fluid Mechanics* **651**, 443 (2010).

Nomoto, K. and Hashimoto, M.-a. Presupernova evolution of massive stars. *Physics Reports* **163**, 13 (1988).

O'Connor, E. and Ott, C. D. Black Hole Formation in Failing Core-Collapse Supernovae. *The Astrophysical Journal* **730**, 70 (2011).

Paczyński, B. Evolution of Single Stars. II. Core Helium Burning in Population I Stars. *Acta Astronomica* **20**, 195 (1970).

Pagel, B. E. J.; Simonson, E. A.; Terlevich, R. J. and Edmunds, M. G. The primordial helium abundance from observations of extragalactic H II regions **255**, 325 (1992).

Palacios, A.; Meynet, G.; Vuissoz, C.; Knödlseder, J.; Schaerer, D.; Cerviño, M. and Mowlavi, N. New estimates of the contribution of Wolf-Rayet stellar winds to the Galactic ^{26}Al. *Astronomy & Astrophysics* **429**, 613 (2005).

Paxton, B.; Cantiello, M.; Arras, P.; Bildsten, L.; Brown, E. F.; Dotter, A.; Mankovich, C.; Montgomery, M. H.; Stello, D.; Timmes, F. X. and Townsend, R. Modules for Experiments in Stellar Astrophysics (MESA): Planets, Oscillations, Rotation, and Massive Stars **208**, 4 (2013).

Pignatari, M.; Gallino, R.; Heil, M.; Wiescher, M.; Käppeler, F.; Herwig, F. and Bisterzo, S. The Weak s-Process in Massive Stars and its Dependence on the Neutron Capture Cross Sections. *The Astrophysical Journal* **710**, 1557 (2010).

Prantzos, N. Can nucleosynthesis determine the mass limit for stellar black hole formation? **284**, 477 (1994).

Prantzos, N. Radioactive ^{26}Al and ^{60}Fe in the Milky Way: Implications of the RHESSI detection of ^{60}Fe. *Astronomy & Astrophysics* **420**, 1033 (2004).

Pumo, M. L.; Contino, G.; Bonanno, A. and Zappalà, R. A. Uncertaities due to convection and production of s-nuclei in massive stars. *Memorie della Societa Astronomica Italiana* p. 4 (2011).

Radko, T. A mechanism for layer formation in a double-diffusive fluid. *Journal of Fluid Mechanics* **497**, 365 (2003).

Radko, T. and Smith, D. P. Equilibrium transport in double-diffusive convection. *Journal of Fluid Mechanics* **692**, 5 (2012).

Rempel, M. Overshoot at the Base of the Solar Convection Zone: A Semianalytical Approach. *The Astrophysical Journal* **607**, 1046 (2004).

Richard, O.; Michaud, G. and Richer, J. Iron Convection Zones in B, A, and F Stars. *The Astrophysical Journal* **558**, 377 (2001).

Rogers, T. M. and Glatzmaier, G. A. Penetrative Convection within the Anelastic Approximation. *The Astrophysical Journal* **620**, 432 (2005).

Rogers, T. M.; Glatzmaier, G. A. and Jones, C. A. Numerical simulations of penetration and overshoot in the sun. *The Astrophysical Journal* **653**, 765 (2006).

Rosenblum, E.; Garaud, P.; Traxler, A. L. and Stellmach, S. TURBULENT MIXING AND LAYER FORMATION IN DOUBLE-DIFFUSIVE CONVECTION: THREE-

DIMENSIONAL NUMERICAL SIMULATIONS AND THEORY. *The Astrophysical Journal* **731**, 66 (2011).

Roxburgh, I. W. A Note on the Boundary of Convective Zones in Stars. *Monthly Notices of the Royal Astronomical Society* **130**, 223 (1965).

Saio, H.; Nomoto, K. and Kato, M. Why did the progenitor of SN 1987A undergo the blue-red-blue evolution? *The Astrophysical Journal* **331**, 388 (1988).

Sakashita, S.; Ôno, Y. and Hayashi, C. The Evolution of Massive Stars. I. *Progress of Theoretical Physics* **21**, 315 (1959).

Salasnich, B.; Chiosi, C. and Bressan, A. Evolution of Massive Stars under New Mass-Loss Rates for RSG: Is the mystery of the missing blue gap solved? *Astronomy & Astrophysics* **342**, 131 (1999).

Salpeter, E. E. The Luminosity Function and Stellar Evolution. *The Astrophysical Journal* **121**, 161 (1955).

Santos, N. C.; Israelian, G. and Mayor, M. The metal-rich nature of stars with planets. *Astronomy & Astrophysics* **373**, 1019 (2001).

Saslaw, W. C. and Schwarzschild, M. Overshooting from Stellar Convective Cores. *The Astrophysical Journal* **142**, 1468 (1965).

Scalo, J. M. The stellar initial mass function. *Fundamentals of Cosmic Physics (ISSN 0094-5846)* **11**, 1 (1986).

Schmitt, J. H. M. M.; Rosner, R. and Bohn, H. U. The overshoot region at the bottom of the solar convection zone. *The Astrophysical Journal* **282**, 316 (1984).

Schmitt, R. W. Flux Measurements on Salt Fingers at an Interface. *Journal of Marine Research* **37**, 419 (1979).

Schmitt, R. W. The characteristics of salt fingers in a variety of fluid systems, including stellar interiors, liquid metals, oceans, and magmas. *Physics of Fluids (ISSN 0031-9171)* **26**, 2373 (1983).

Schmitt, R. W. Enhanced Diapycnal Mixing by Salt Fingers in the Thermocline of the Tropical Atlantic. *Science* **308**, 685 (2005).

Schürmann, D.; Gialanella, L.; Kunz, R. and Strieder, F. The astrophysical S factor of $^{12}C(\alpha,\gamma)^{16}O$ at stellar energy. *Physics Letters B* **711**, 35 (2012).

Schwab, J.; Podsiadlowski, P. and Rappaport, S. Further Evidence for the Bimodal Distribution of Neutron-star Masses. *The Astrophysical Journal* **719**, 722 (2010).

Schwarzschild, M. and Härm, R. Evolution of Very Massive Stars. *The Astrophysical Journal* **128**, 348 (1958).

Shaviv, G. and Salpeter, E. E. Convective Overshooting in Stellar Interior Models. *The Astrophysical Journal* **184**, 191 (1973).

Shirtcliffe, T. G. L. Transport and profile measurements of the diffusive interface in double diffusive convection with similar diffusivities. *Journal of Fluid Mechanics* **57**, 27 (1973).

Singh, H. P.; Roxburgh, I. W. and Chan, K. L. Three-dimensional simulation of penetrative convection-penetration above a convection zone. *Astronomy & Astrophysics* (1994).

Smartt, S. J. Progenitors of Core-Collapse Supernovae. *arXiv.org* **47**, 63 (2009).

Smartt, S. J. Abstract. *Publications of the Astronomical Society of Australia* **32**, e016 (2015).

Sofia, S. and Chan, K. L. Turbulent compressible convection in a deep atmosphere. II - Two-dimensional results for main-sequence A5 and F0 type envelopes. *Astrophysical Journal* **282**, 550 (1984).

Spiegel, E. A. and Veronis, G. On the Boussinesq Approximation for a Compressible Fluid. *The Astrophysical Journal* **131**, 442 (1960).

Spruit, H. C. The rate of mixing in semiconvective zones. *Astronomy & Astrophysics* **253**, 131 (1992).

Spruit, H. C. Semiconvection: theory. *Astronomy & Astrophysics* **552**, A76 (2013).

Steiner, A. W.; Hempel, M. and Fischer, T. Core-collapse Supernova Equations of State Based on Neutron Star Observations. *The Astrophysical Journal* **774**, 17 (2013).

Stellmach, S.; Traxler, A. L.; Garaud, P.; Brummell, N. H. and Radko, T. Dynamics of fingering convection. Part 2 The formation of thermohaline staircases. *Journal of Fluid Mechanics* **677**, 554 (2011).

Stern, M. E. The "Salt-Fountain" and Thermohaline Convection. *Tellus A* **12**, 172 (1960).

Stern, M. E. Collective instability of salt fingers. *Journal of Fluid Mechanics* **35**, 209 (1969).

Stern, M. E.; Radko, T. and Simeonov, J. Salt fingers in an unbounded thermocline. *Journal of Marine Research* **59**, 355 (2001).

Stothers, R. and Chin, C. W. Stellar evolution at high mass with semiconvective mixing according to the Schwarzschild criterion. *The Astrophysical Journal* **204**, 472 (1976).

Stothers, R. B. and Chin, C.-W. Blue loops during core helium burning as the consequence of moderate convective envelope overshooting in stars of intermediate to high mass. *The Astrophysical Journal* **374**, 288 (1991).

Straus, J. M.; Blake, J. B. and Schramm, D. N. Effects of convective overshoot on lithium depletion in main-sequence stars. *The Astrophysical Journal* **204**, 481 (1976).

Sugimoto, D. and Fujimoto, M. Y. Why Stars Become Red Giants. *The Astrophysical Journal* **538**, 837 (2000).

Sukhbold, T.; Ertl, T.; Woosley, S. E.; Brown, J. M. and Janka, H. T. CORE-COLLAPSE SUPERNOVAE FROM 9 TO 120 SOLAR MASSES BASED ON NEUTRINO-POWERED EXPLOSIONS. *The Astrophysical Journal* **821**, 38 (2016).

Sukhbold, T. and Woosley, S. E. The Compactness of Presupernova Stellar Cores. *The Astrophysical Journal* **783**, 10 (2014).

Théado, S. and Vauclair, S. Radiative accelerations, accumulation of iron and thermohaline convection inside stars. *Astrophysics and Space Science* **328**, 209 (2009).

Timmes, F. X. and Arnett, W. D. The Accuracy, Consistency, and Speed of Five Equations of State for Stellar Hydrodynamics. *The Astrophysical Journal Supplement* **125**, 277 (1999).

Timmes, F. X.; Woosley, S. E. and Weaver, T. A. Galactic chemical evolution: Hydrogen through zinc. *Astrophysical Journal Supplement Series (ISSN 0067-0049)* **98**, 617 (1995).

Traxler, A. L.; Garaud, P. and Stellmach, S. Numerically Determined Transport Laws for Fingering ("Thermohaline") Convection in Astrophysics. *The Astrophysical Journal Letters* **728**, L29 (2011a).

Traxler, A. L.; Stellmach, S.; Garaud, P.; Radko, T. and Brummell, N. H. Dynamics of fingering convection. Part 1 Small-scale fluxes and large-scale instabilities. *Journal of Fluid Mechanics* **677**, 530 (2011b).

Turner, J. S. The coupled turbulent transports of salt and and heat across a sharp density interface. *International Journal of Heat and Mass Transfer* **8**, 759 (1965).

Turner, J. S. Salt fingers across a density interface. *Deep Sea Research and Oceanographic Abstracts* **14**, 599 (1967).

Turner, J. S. and Stommel, H. A NEW CASE OF CONVECTION IN THE PRESENCE OF COMBINED VERTICAL SALINITY AND TEMPERATURE GRADIENTS. *Proceedings of the National Academy of Sciences* **52**, 49 (1964).

Twarog, B. A. and Wheeler, J. C. Primary nucleosynthesis in the galactic disk. *The Astrophysical Journal* **261**, 636 (1982).

Twarog, B. A. and Wheeler, J. C. A reanalysis of galactic primary nucleosynthesis. *The Astrophysical Journal* **316**, 153 (1987).

Ulrich, R. K. Thermohaline Convection in Stellar Interiors. *The Astrophysical Journal* **172**, 165 (1972).

van den Bergh, S. The Galaxies of the Local Group (to be Concluded). *Journal of the Royal Astronomical Society of Canada* **62**, 145 (1968).

Vauclair, S. Metallic Fingers and Metallicity Excess in Exoplanets' Host Stars: The Accretion Hypothesis Revisited. *The Astrophysical Journal* **605**, 874 (2004).

Veronis, G. Penetrative Convection. *The Astrophysical Journal* **137**, 641 (1963).

Viallet, M.; Meakin, C. A.; Prat, V. and Arnett, W. D. Toward a consistent use of overshooting parametrizations in 1D stellar evolution codes. *Astronomy & Astrophysics* **580**, A61 (2015).

Walin, G. Note on the stability of water stratified by both salt and heat. *Tellus A* **16**, 389 (1964).

Walmswell, J. J.; Tout, C. A. and Eldridge, J. J. On the blue loops of intermediate-mass stars. *Monthly Notices of the Royal Astronomical Society* **447**, 2951 (2015).

Wang, W.; Harris, M. J.; Diehl, R.; Halloin, H.; Cordier, B.; Strong, A. W.; Kretschmer, K.; Knödlseder, J.; Jean, P.; Lichti, G. G.; Roques, J. P.; Schanne, S.; von Kienlin, A.; Weidenspointner, G. and Wunderer, C. SPI observations of the diffuse ^{60}Fe emission in the Galaxy. *arXiv.org* **469**, 1005 (2007).

Weaver, T. A.; Zimmerman, G. B. and Woosley, S. E. Presupernova evolution of massive stars. *The Astrophysical Journal* **225**, 1021 (1978).

Wiktorowicz, G.; Belczynski, K. and Maccarone, T. Black hole X-ray transients: The formation puzzle. *Binary Systems* p. 37 (2014).

Wood, T. S.; Garaud, P. and Stellmach, S. A NEW MODEL FOR MIXING BY DOUBLE-DIFFUSIVE CONVECTION (SEMI-CONVECTION). II. THE TRANSPORT OF HEAT AND COMPOSITION THROUGH LAYERS. *The Astrophysical Journal* **768**, 157 (2013).

Woosley, S. E. and Heger, A. Nucleosynthesis and Remnants in Massive Stars of Solar Metallicity. *Physics Reports* **442**, 269 (2007).

Woosley, S. E.; Heger, A. and Weaver, T. A. The evolution and explosion of massive stars. *Reviews of Modern Physics* **74**, 1015 (2002).

Woosley, S. E. and Kasen, D. SUB-CHANDRASEKHAR MASS MODELS FOR SUPERNOVAE. *The Astrophysical Journal* **734**, 38 (2011).

Woosley, S. E. and Weaver, T. A. The Evolution and Explosion of Massive Stars. II. Explosive Hydrodynamics and Nucleosynthesis. *The Astrophysical Journal Supplement* **101**, 181 (1995).

Xiong, D.-R. The evolution of massive stars using a non-local theory of convection. *Astronomy & Astrophysics* **167**, 239 (1986).

Xiong, D.-R. and Deng, L. The structure of the solar convective overshooting zone. *Monthly Notices of the Royal Astronomical Society* **327**, 1137 (2001).

Xiong, D.-R. and Deng, L. Lithium depletion in late-type dwarfs. *Monthly Notices of the Royal Astronomical Society* **395**, 2013 (2009).

Xiong, D.-R.; Kondo, M.-A. and Unno, W. The surface lithium and beryllium depletions of the sun. *Acta Astronomica Sinica (ISSN 0001-5245)* **32**, 333 (1991).

Zahn, J. P. Convective Overshooting: Physical Properties and Seismic Evidence (invited review). *Radial and Nonradial Pulsations as Probes of Stellar Physics* **259**, 58 (2002).

Zaussinger, F. and Spruit, H. C. Semiconvection: numerical simulations. *Astronomy & Astrophysics* **554**, A119 (2013).

Zemskova, V.; Garaud, P.; Deal, M. and Vauclair, S. Fingering convection induced by atomic diffusion in stars: 3D numerical computations and applications to stellar models. *arXiv.org* (2014).

Zhang, Q. S. TESTING THE CORE OVERSHOOT MIXING DESCRIBED BY A TURBULENT CONVECTION MODEL ON THE ECLIPSING BINARY STAR HY VIR. *The Astrophysical Journal* **761**, 153 (2012).

Zhang, W.-Q.; Woosley, S. E. and Heger, A. Fallback and Black Hole Production in Massive Stars. *The Astrophysical Journal* **679**, 639 (2008).