A theory of non-local mixing-length convection – III. Comparing theory and numerical experiment

Scott A. Grossman

Canadian Institute for Theoretical Astrophysics, 60 St George Street, Toronto, Ontario M5S 1A7, Canada

Accepted 1995 September 14. Received 1995 September 11; in original form 1994 September 30

ABSTRACT

We solve the non-local convection equations. The solutions for four model problems are compared with results of generalized smoothed particle hydrodynamics (GSPH) simulations. In each case we test two closure schemes: (1) where third moments are defined by the diffusion approximation; and (2) where the full third moment equations are used and fourth moments are defined by a modified form of the quasi-normal approximation. In overshooting models, the convective flux becomes negative shortly after the stability boundary. The negative amplitude remains small, and the temperature gradient in the overshooting zone has nearly the radiative value. Turbulent velocities decay by a factor of e after $(0.5-1.5)\ell$, depending on the model, where $\ell$ is the mixing length. Turbulent viscosity is more important than negative buoyancy in decelerating overshooting fluid blobs. These predictions are consistent with helioseismology.

The equations and the GSPH code use the same physical approximations, so it was anticipated that, if the closures for high-order moments are accurate enough, solutions for the low-order moments will automatically agree with the GSPH results. Such internal consistency holds approximately. Unexpectedly, however, the best second moments are found with the first closure scheme, and the best third moments are obtained with the second. The relationship among moments from the solution of the equations is not the same as the relationship found by GSPH simulations. In particular, if we use the best fourth moment closure model suggested by Paper II, we cannot obtain steady-state solutions, but adding a sort of diffusion for stability helps.

Key words: convection – hydrodynamics – turbulence – stars: interiors.

1 INTRODUCTION

The turbulent mixing of material in a convective region into an adjacent region of stability is called convective overshooting. In the broadest terms, convection and convective overshooting have two major consequences for stars. (1) Convection makes the temperature gradient in a star shallower than the radiative value, and usually very close to the adiabatic gradient. Convective overshooting also may modify the temperature gradient in the overshoot region, with implications for the hydrostatic structure of stars. There continues to be disagreement in the literature about whether the overshooting zone is nearly radiative or nearly adiabatic. (2) The composition in convective and overshooting regions is homogenized. In stars with convective cores, the lifetimes of the various nuclear burning phases can be altered significantly by overshooting, and, in stars with convective envelopes, undershooting that reaches into the interior can modify the surface composition. The consequences for stellar evolution if overshooting is significant (say at least a few tenths of a pressure scaleheight) have been investigated extensively (e.g., Bertelli, Bressan & Chiosi 1985; Maeder & Meynet 1989). Nevertheless, whether or not overshooting is, in fact, important remains uncertain.

Papers addressing convective overshooting have been written for more than three decades, and still there are major qualitative disagreements among authors regarding the extent and importance of overshooting. Several authors have concluded that overshooting is significant (e.g., Shaviv & Salpeter 1973; Maeder 1975; Bressan, Bertelli & Chiosi 1981; Zahn 1991; Xiong & Chen 1992; Roxburgh 1978, 1989, 1992), whereas others have come to the opposite conclusion (Travis & Matsushima 1973; Langer 1986). The
difficulty in resolving this issue theoretically is that convective overshooting is a completely non-local problem. That is, the behaviour of the overshooting region depends sensitively on the behaviour of the adjacent convective region. The mixing-length theory (Bohm-Vitense 1958) usually used to describe stellar convection is strictly local. It is not a simple matter to decide which authors are more likely to be correct.

In fact, Renzini (1987) has criticized most of the above works, on both sides of this issue, for internal inconsistencies and unjustified assumptions.

As a result of theoretical uncertainties, most stellar models are computed using local convection and no overshooting. When overshooting is included (e.g. Doom 1982a,b; Maeder & Meynet 1988; Chin & Stothers 1991), it usually is included by using an overshooting distance parametrized by $a_{\text{over}} = d_{\text{over}}/H\nu$, the ratio of the overshooting distance to the pressure scaleheight at the Schwarzschild stability boundary.

Since a variety of observational consequences of overshooting are known (see Stothers 1991) for a summary of many observational tests, it is possible, in principle, to calibrate the overshooting parameter, without an understanding of the physics of non-local convection. Many authors have attempted such a calibration, but, as with the theoretical work, there is disagreement among authors here also. Several comparisons between cluster data and theoretical Hertzsprung–Russell (HR) diagrams favour an intermediate amount of overshooting ($d_{\text{over}} = (0.2–0.3) H\nu$) from stellar cores (Maeder & Mermilliod 1981; Mermilliod & Maeder 1986; Chiosi et al. 1989); and undershooting from convective envelopes may be required also (Alongi et al. 1991). Other authors conclude that stellar evolution can be understood without overshooting (Stothers 1991; Stothers & Chin 1992), and set an upper limit of $d_{\text{over}} < 0.2 H\nu$. Even if the empirical calibrations eventually converged to a widely accepted value appropriate to certain regions of certain kinds of stars, overshooting distances probably would depend on the particular conditions of the star and probably would not apply to both core and envelope convection in all types of stars. Thus a theory for non-local convection ultimately will be required.

This paper is the third in a series devoted to the development of a theory of non-local convection. In Paper I of this series (Grossman, Narayan & Arnett 1993), we developed a Boltzmann transport theory for the evolution of turbulent fluid elements and derived the equations for the hydrodynamic evolution of high-order correlations of velocity and temperature. In that work, the state variables of each fluid blob were its vertical position $z$, vertical velocity $v$ and temperature $T$. The ensemble of fluid blobs was described by the absolute distribution function $f_k(t, z, v, T)$ or, equivalently, by the relative distribution function $f_k(t, z, w, \theta)$ of perturbations $w = v - \bar{v}$, $\theta = T - \bar{T}$. (Bars over variables indicate ensemble averages.) The distribution function evolved according to a Boltzmann-like equation, given by

$$\frac{\partial f_k}{\partial t} + \frac{\partial}{\partial z} (\bar{v} f_k) + \frac{\partial}{\partial w} (w f_k) + \frac{\partial}{\partial \theta} (\theta f_k) = \Gamma,$$

where $\bar{w}$ and $\bar{\theta}$ represent the dynamical equations for the evolution of a single fluid blob and $\Gamma$ is a collision term. Equation (1) connects convection theories of the ballistic particle type (theories that integrate $\bar{w}$ and $\bar{\theta}$) to those that use a hydrodynamic approach. The derivation assumes that fluid blobs are in pressure balance with the local mean fluid.

In Paper I, the effects of turbulent viscosity and turbulent diffusion entered the $\bar{w}$- and $\bar{\theta}$-equations as eddy-damping terms. We assumed that turbulent fluid blobs travel their characteristic distance of a mixing length $\ell$ with characteristic turbulent velocity $\alpha \nu$ before giving up their excess momentum and heat to the ambient fluid. Hence turbulent viscosity and diffusion coefficients are defined as $\gamma_{\text{turb}} \sim \chi_{\text{turb}} \sim \ell \nu$. Since we account for turbulent losses on the left-hand side of the transport equation (1), we set the collision term on the right-hand side to zero. By taking $w$- and $\theta$-moments of the transport equation, we derived all moment equations up to third order. In an alternative formulation of this problem, presented in detail in Appendix B, we treat the effects of turbulent damping in the collision term instead of the dynamical equations for $\bar{w}$ and $\bar{\theta}$.

The moment equations form an unclosed hierarchy, and thus require closure relations if they are to be solved. In Paper II of this series (Grossman & Narayan 1993), we simulated non-local convection in a one-dimensional box using an algorithm we call 'generalized smoothed particle hydrodynamics' (GSPH). This code used the same equations for the evolution of velocities and temperatures, $\bar{w}$ and $\bar{\theta}$, as the moment theory, thereby simulating the same physics. We presented results of four simulations. Two were for 'homogeneous' convection, that is, for fluids convective throughout, one of which was in the regime of efficient convection and the other in the regime of inefficient convection. Two were overshooting simulations, where fluids made the transition from instability to stability at the centre of the box. The unstable region in one case was in the regime of efficient convection and in the other in the regime of inefficient convection. We investigated the non-local behaviour of convecting fluids, emphasizing the physics of overshooting regions. The relationships of third and fourth moments to lower order moments were studied, in order to discover useful closure approximations.

In this paper we solve the moment equations of Paper I using closure relations suggested by the GSPH simulations of Paper II. We regard the GSPH results as data to be modelled by a correct analytic description of non-local convection. Hence we compare the steady-state solutions of the moment equations with the results of the four GSPH simulations presented in Paper II.

The equations are solved using two different closure schemes. The first, used previously in the astrophysical literature by Xiong (1980, 1981, 1989), closes the equations at the third moments using the diffusion approximation. Xiong’s equations have been solved in detail for the convection and overshooting zones of the Sun (Unno, Kondo & Xiong 1985; Unno & Kondo 1989; Xiong & Chen 1992) and other stars of various sorts (Xiong 1985, 1986, 1990).

The second closure scheme closes the equations at the fourth moments using a modified version of the quasi-normal approximation. To our knowledge, hydrodynamic equations closed at this high order have never before been solved in the astrophysical literature. Canuto (1992, 1993) wrote equations to the same high order using the quasi-normal closure scheme. They have been solved for boundary conditions appropriate for Earth’s convective boundary layer (Canuto et al. 1994). One key difference between this
problem and astrophysical convection is that planetary convection can be described as an initial-value problem, and the equations can be integrated outward. The convective and total fluxes are computed in this integration. In stellar convection, the total flux is specified, and boundary conditions must be placed on both the inner and outer boundaries of the fluid. Thus solving the astrophysical problem is more complicated. Also, astrophysical convection occurs in zones many pressure scaleheights deep, whereas in the Earth convection occurs over a fraction of a scaleheight (in a region $\sim 1$ km thick compared with a scaleheight $\sim 7$ km), and is driven not by the superadiabatic temperature gradient but by the temperature gradient. Nevertheless, the favourable agreement between the computed moments and their observed behaviour in this context lends support to our approach to non-local convection.

In this paper, we consider not only two different closure schemes for our set of moment equations, but also, in fact, two different sets of moment equations. The first set comprises simply the equations of Paper I. The second is derived using the alternative collisional treatment of turbulent losses. The differences in the equations seem minor, but solutions show significant quantitative differences. The latter theory is somewhat flexible and may ultimately prove to be a more useful set of equations.

The plan of the paper is as follows. In Section 2 we discuss the moment equations and the closure relations that we solve. We analyse the moment equations to make rough analytic estimates of overshooting distances. These predictions are compared with those of Zahn (1991) and with results inferred from helioseismology. In Section 3 we outline the numerical method used to solve the moment equations. The solutions for our four models are presented in Section 4, where we make a detailed comparison with the GSPH results. In Section 5 we discuss the alternative formulation of the moment equations, where turbulent losses originate as scattering by the Boltzmann collision term. We refer to these alternative equations as the scattering equations, in which we also have included turbulence in the horizontal dimensions. We discuss the closure relations used in the solution of these equations. In Section 6 the equations are solved and compared with results of new GSPH simulations. A discussion of the comparisons between theory and numerical simulation and conclusions appear in Section 7.

2 PAPER I EQUATIONS

2.1 The equations and closures

In Paper I we derived equations for the time evolution of the mean density $\rho$, mean vertical velocity $\bar{\theta}$, and mean temperature $T$ of a convecting fluid. In addition, we wrote evolution equations for correlations of $w$ and $\theta$, the velocity and temperature perturbations. A minimal convection theory requires at least the three second-moment equations for $w^2$, $w\theta$ and $\theta^2$. The equations include gradients of third moments. If these third-moment terms are dropped, standard local mixing-length theory is recovered, but to describe non-local convection these third moments must be kept. They can be defined either by non-trivial closure relations (unlike the trivial relations equating them to zero) or by the four third-moment equations for $w^3$, $w^2\theta$, $w\theta^2$ and $\theta^3$. In the latter case, the third-moment equations include gradients of fourth moments, for which we will require non-trivial closures.

The two non-trivial closure schemes, to which we refer as the 'Xiong solution' and the 'Full solution', are outlined below. The equations that we solve here are slightly simplified from those of Paper I and are presented in Appendix A. They are supplemented by an ideal-gas equation of state.

2.1.1 Xiong solution

A common third-moment closure is the diffusion or down-gradient approximation, which relates third moments to gradients of second moments. Although not invented by Xiong, they have been advanced in the astrophysical literature by him (Xiong 1980, 1981). They can be written as

$$w^3 = -c_1 \ell \sigma_u \nabla w^2,$$

$$w^2\theta = -c_2 \ell \sigma_u \nabla w\theta,$$

$$w\theta^2 = -c_3 \ell \sigma_u \nabla \theta^2,$$

where $\ell$ is the mixing length and $\sigma_u$ is the turbulent velocity dispersion. Averaged over all four simulations presented in Paper II, the constants $c_1$, $c_2$, and $c_3$ were calibrated to have values between 0.5 and 1.8. These optimal values, however, had dispersions from one simulation to the next of order 100 per cent. In the solutions below we have taken these constants to be equal to unity since the deviation from the optimal values apparently makes little difference in the second moments. Of course, the third moments are affected more directly by the values of these constants. Although a closure for $\theta^3$ is not required, we nevertheless compute it according to $\theta^3 = w^2\theta^2/3$ for comparison with the GSPH results.

2.1.2 Full solution

The Full solution includes the third-moment equations, which are closed at the level of the fourth moments. The closures can be written as

$$w^4 = \xi w^3 - d\sigma_u \nabla w^3,$$

$$w^3\theta = \xi w^2 \theta^2 - d\sigma_u \nabla w^2\theta,$$

$$w^2\theta^2 = \xi w\theta^3 - d\sigma_u \nabla w\theta^2,$$

$$w\theta^3 = \xi \theta^4 - d\sigma_u \nabla \theta^3.$$

Setting $\xi = 3$ in each of the first terms and dropping the second terms reproduce the standard quasi-normal approximation of mathematical hydrodynamics (e.g. Orszag 1977; Lesieur 1987). The second terms in these closures cause diffusion of third moments upon substitution into the third-moment equations. The parameter $d$ sets the diffusion rate. For most of our convection models (described below), we cannot obtain solutions without this term. If we set $d = 0$, physically positive quantities, such as $w^2$ or $\theta^2$, sometimes become negative as solutions evolve toward the steady state. It is likely that the final $w^3$ and $\theta^3$ would be negative if the solutions could reach a steady state, and that these unphysical solutions are not merely a transient state. If this happens,
our code eventually will fail to converge and will not reach a steady state. We find that using $d = 0.8$ allows steady-state solutions of acceptable qualitative agreement to our four standard problems. We do not present any results for a pure quasi-normal closure model. (In the few cases where we can obtain solutions with a pure quasi-normal closure, the solutions compare with the GSPH simulations much more poorly.) Note that a closure for $\theta^4$ is not required.

2.2 The scale of overshooting

Before solving the moment equations numerically, we make some estimates of the extent of convective overshooting based on analysis of the equations. To keep the problem tractable, we consider only the Xiong closure here. We simplify the second-moment equations to a point that captures only the essential non-local physics. We consider a quasi-homogeneous fluid of constant density $\rho$ and temperature $T$, but with a variable superadiabatic gradient $\Delta V T$. We assume that the turbulent velocity is sufficiently subsonic that gravity $g$ is balanced entirely by the thermal pressure gradient, and drop terms associated with viscous heating, since they are always small in our models. The steady-state second-moment equations from Appendix A become

$$\frac{\partial \omega^2}{\partial z} - \frac{2ga}{T} \omega + 2(A + B\sigma_{\omega}) \omega^2 = 0, \quad (4a)$$

$$\frac{\partial \omega^2}{\partial z} - \frac{ga}{T} \omega^3 + 2(A + B\sigma_{\omega} + D + E\sigma_{\omega}) \omega - \Delta V T \omega^2 = 0, \quad (4b)$$

$$\frac{\partial \omega^2}{\partial z} - 2\Delta V T \omega + 2(D + E\sigma_{\omega}) \omega^2 = 0. \quad (4c)$$

The constants are defined in equations (A11), but, briefly, $A$ and $D$ measure microscopic viscosity and radiative diffusion damping rates of momentum and heat excesses, and $B\sigma_{\omega}$ and $E\sigma_{\omega}$ are the damping rates by turbulent processes. The constant $a$ is the coefficient of thermal expansion, equal to unity for an ideal gas. The leading gradient terms are responsible for non-locality; without them, these equations would describe local mixing-length theory.

Let us consider a fluid with an unstable superadiabatic gradient $\Delta V T > 0$ adjacent to a stable region with $\Delta V T < 0$. In the unstable region, the second moments can be approximated by the values of the local theory, so that at the stability boundary the second moments can be written in terms of the local values at the stability boundary.

If we describe the decay of second moments in the overshoot zone with exponential dependences $\omega^2 \sim \exp(-z/H_{\omega})$, $\omega^3 \sim \exp(-z/H_{\omega^3})$, and $\omega \sim \exp(-z/H_{\omega})$, we estimate overshooting scales $H_{\omega^3} = 0.17$, $H_{\omega^2} = 0.05$, and $H_{\omega} = 0.05$, which compare favourably with the distances 0.17, 0.05 and 0.05 for the second moments to fall by a factor of $e$ beyond the stability transition. We note also that the local and non-local solutions are comparable at the stability transition, justifying our approximation that the second moments decay from the local values at the stability boundary.

In all cases of astrophysical relevance, the turbulent viscosity far exceeds the molecular viscosity, $B\sigma_{\omega}/A >> 1$, so that

$$H_{\omega^3} = H_{\omega^2} = \left[ \frac{\tau_{\omega}}{2(D + E\sigma_{\omega})} \frac{\Delta V T}{\Delta V T - \Delta V T_1} \right]^{1/2}. \quad (5)$$

The closer $\Delta V T_1$ is to zero, the shorter is the decay scale of moments involving $\theta$. The reason is that turbulent blobs moving adiabatically accumulate less temperature excess as they move in the unstable region, and hence lose what little they have accumulated more quickly in the stable region. If the Peclet number, the ratio of turbulent to diffusive damping rates, $E\sigma_{\omega}/D >> 1$, convection is efficient and

$$H_{\omega^3} = H_{\omega^2} = \left( \frac{\tau_{\omega}}{2} \right)^{1/2} \left( \frac{\Delta V T}{\Delta V T - \Delta V T_1} \right)^{1/2}. \quad (6)$$

If $E\sigma_{\omega}/D << 1$, convection is inefficient and

$$H_{\omega^3} = H_{\omega^2} = \left( \frac{E\sigma_{\omega}}{D} \right)^{1/2} \left( \frac{\Delta V T}{\Delta V T - \Delta V T_1} \right)^{1/2}. \quad (7)$$

Because the superadiabatic gradient does not enter into equation (4a), $\omega^2$ does not respond immediately to the change in the superadiabatic gradient. As we just showed, $\omega^3$ decays and an overshooting particle loses its buoyancy. When $\omega$ becomes small, the non-local term of equation (4a) must be balanced by the viscous damping term, so that

$$H_{\omega^3} = \left( \frac{\tau_{\omega}}{2} \right). \quad (8)$$

A typical overshooting scale might be estimated as the sum $H_{\omega^3} + H_{\omega^2}$. Of course, since the equations are non-linear, the decays are not simple exponentials, and these estimates are likely to be fairly crude.

To evaluate the usefulness of these quantitative estimates, in Fig. 1 we show the local and non-local non-linear solutions for two overshooting models. The regions of stability and instability are each characterized by approximately constant superadiabatic gradients. (The parameters of these models are the same as those of the overshooting models discussed below, except that the diffusion coefficient is constant in the upper and lower halves of the box and is discontinuous across the stability boundary.) The efficient overshooting model has a convective zone in the efficient regime. The solution to the local equations is shown as dotted lines. For the mixing length $\tau_{\omega} = 0.24$, we estimate $H_{\omega^3} \approx 0.12$, $H_{\omega^2} \approx 0.07$ and $H_{\omega} \approx 0.07$, which compare favourably with the distances 0.17, 0.05 and 0.05 for the second moments to fall by a factor of $e$ beyond the stability transition. We note also that the local and non-local solutions are comparable at the stability transition, justifying our approximation that the second moments decay from the local values at the stability boundary.

The inefficient overshooting model has a convective region in the inefficient regime. We estimate overshooting scales $H_{\omega^3} \approx 0.12$, $H_{\omega^2} \approx 0.07$ and $H_{\omega} \approx 0.07$, which compare well with the values 0.14, 0.03 and 0.03 for the real non-linear solutions. In the convective region, the difference between the local and non-local solutions is greater in this case because the convective flux is not constrained to be...
nearly the total flux. We note that the distance $H_{\omega}$ slightly underestimates the actual overshooting distance of $\omega^2$, whereas the sum $H_{\omega} + H_{\omega0}$ slightly overestimates it.

2.3 Zahn's analysis

We examine the analysis of Zahn (1991) within the framework of our moment theory and consider whether his results are consistent with ours. From the $T$-equation (A3), if there is no local heating or cooling and a constant flux flows through a fluid, changes in the radiative flux must be balanced by changes in the convective flux, $\nabla (\bar{K} \nabla T) = \nabla (\rho \bar{c}_p \bar{w} \theta)$, where $\bar{c}_p$ is the specific heat at constant pressure. Zahn assumes the overshooting region is nearly adiabatic, so that $\nabla T$ is approximately the constant, adiabatic value, $\nabla T_{ad}$. He also assumes that the outgoing convective flux falls to zero and then changes sign at the beginning of the overshoot region. Then, expanding to linear order, he obtains for the convective flux

$$F_{\text{conv}} = \bar{c}_p \bar{w} \theta \bar{v} \nabla \bar{K} \nabla T_{ad} \bar{\varepsilon},$$

where $F_{\text{conv}}$ is the convective flux, $\bar{c}_p$ is the average specific heat at constant pressure, $\bar{w}$ is the average velocity, $\bar{v}$ is the rms velocity, $\bar{K}$ is the average thermal conductivity, $\nabla$ is the gradient operator, $\nabla T_{ad}$ is the adiabatic temperature gradient, and $\bar{\varepsilon}$ is the rms temperature fluctuation.

Figure 1. The superadiabatic gradient and second moments for an efficient and inefficient overshooting model. The transition from instability to stability is discontinuous, as assumed in the analysis of Section 2.2. The distances for the second moments to decay by a factor of $e$ beyond the stability transition at $z = 0.4$ are comparable to the distances estimated by equations (6), (7) and (9). Note that in the efficient model the local and non-local solutions have comparable amplitude in the convective region because the convective flux is constrained to be nearly the total flux. The comparison is not as favourable in the inefficient case, since there is no strong constraint on the convective flux. Here and throughout this paper, we take increasing $z$ and positive velocities in the direction opposing gravity.
which is negative since the temperature gradient is negative. We think that this is an inconsistency. The convective flux does indeed change sign slightly beyond the stability transition (see below), but it does this precisely because over-shooting blobs are moving against a stable temperature gradient. The convective flux cannot change sign if the temperature gradient remains adiabatic. Further, we find below that, when it does change sign, it does not obtain the large negative values required to maintain a nearly adiabatic temperature gradient, but remains relatively small.

According to Paper I (equation 5.1), the velocity of the overshooting blob evolves according to

$$\dot{w} = \frac{\partial w}{\partial z} \frac{ga}{T} - (A + Bw)w,$$

which has contributions arising from both buoyancy and viscosity. A large negative convective flux implies a large negative buoyancy, which Zahn regards as the primary reason that over-shooting blobs decelerate, and hence he neglects the viscosity term. Then, multiplying equation (11) by $w$, using the result of equation (10), and integrating the velocity until it decreases to zero, he obtains the overshooting distance

$$z_{\text{over}} \approx \left( \frac{w^0}{2} \right)^{3/2} \left( \frac{a}{T^0} \nabla \tilde{n} \tilde{T} \right)^{-1/2},$$

where $w^0$ is the characteristic convective velocity at the start of the overshooting zone, which can be estimated from the local mixing-length equations. Equation (12) is essentially Zahn’s equation (3.9), except for his inclusion of some dimensionless constants of order unity to account for the relation between $w\theta$ and $\tilde{w}\theta$.

Zahn’s reasoning is quite different from our preceding analysis, which presumed that the viscosity term, and not the buoyancy term, dominates the overshooting calculation. Although many authors have made the same presumption in the past as Zahn, Umezu (1992) demonstrates that the viscosity term is important for the computation of overshooting. Furthermore, the solutions in Fig. 1 verify that, although the buoyancy term in equation (11) becomes negative and contributes to the deceleration of velocities, it never over-whelms the viscosity term. The discrepancy between Zahn’s analysis and ours comes from his having a nearly adiabatic overshooting zone compared with our nearly radiative overshooting zone. Ultimately, we think that this can be traced to his neglect of the term $\nabla \tilde{n} \tilde{T}$ in the expansion of $\nabla \tilde{u}$.

Indeed, this term is dominant over the $\nabla \tilde{n} \tilde{T}$ term in the overshooting regions of our models below. Presumably this term is responsible for making the temperature gradient take nearly the radiative value, rather than the adiabatic value, in the overshooting region.

We compare predictions for a solar model based on Zahn’s and our analyses with data on the Sun’s overshooting (really undershooting) zone. Conviction in the solar convection zone is efficient, and the temperature gradient is nearly adiabatic. Using standard notation for the dimensionless logarithmic temperature gradient, we have \(\nabla \sim 10^{-6}\) in the convective zone of the Sun (Chaboyer, private communication, for a solar model computed using local mixing-length theory with $\ell = 2H_p$). Below the convection zone, the temperature gradient takes on the radiative value, making a transition to $\nabla \approx 0.2$ over about $0.8H_p$. Substituting these numbers into equation (6), we find that the convective flux decays on a scale $H_{\text{ad}} \sim 2 \times 10^{-3}H_p$. Since this is so much smaller than the $0.8H_p$ transition region, the temperature gradient here should depart very little from the radiative gradient. This is consistent with recent helioseismological observations which show that the extension of the adiabatic region below the solar convection zone is $< 0.1H_p$ and consistent with zero (Basu, Antia & Narasimha 1994; Monteiro, Christensen-Dalsgaard & Thompson 1994), and is in accord with the solar model computed using Xiong’s non-local theory (Xiong & Chen 1992). In contrast, Zahn predicts an extension to the region of nearly adiabatic convection of about half a pressure scaleheight below the stability transition, $H_{\text{ad}} \sim 0.5H_p$ (by Zahn’s equation 3.13).

Below the very narrow region of width $H_{\text{ad}}$ across which the temperature gradient becomes radiative rather than adiabatic, we predict a decay of the turbulent velocity on a scale of $H_{\text{ad}} - H_p$, comparable to the e-folding scale of 0.6$H_p$ found by Xiong & Chen (1992). The extent to which turbulent velocities overshoot the stability boundary bears on the amount of lithium and beryllium depletion in the solar atmosphere, but the precise extent of overshooting is not probed by helioseismology. Zahn predicts a decay on the scale $H_{\text{ad}} - H_p - 0.5H_p$, to within factors of order unity. Although comparable with our estimate, the physics leading to this estimate is quite different. In fact, since Zahn assumes that the thermal conductivity is linear at the stability transition and we make it discontinuous, Zahn’s analysis cannot address our simple model.

3 THE NUMERICAL METHOD OF SOLUTION

We use a relaxation method to solve the non-linear, coupled differential equations (Press et al. 1986). It has been necessary to treat the temporal integration and spatial gradients in non-trivial ways to obtain numerical solutions to the moment equations. Although our main interest is in the steady-state solutions of the equations, solutions usually cannot be obtained from crude initial guesses. However, following the time evolution from an initial guess to the steady state works well.

Convective time-scales are orders of magnitude longer than hydrodynamic time-scales. Thus relaxing a fluid to a convective steady state over many mixing times would be prohibitive if the time-step of the computation were Courant-limited. To avoid such limitations on the time-step, we treat the time integration implicitly (cf. Press et al. 1986). That is, if the $N$ time-dependent moments are represented by $f^n_i$, where the index $i$ refers to a particular moment and $n$ to the time-step, the moments are integrated using

$$f^{n+1}_i = f^n_i + y_i f^{n+1}_i \Delta t.$$

The source terms $y_i$ are, in general, functions of all the variables and their spatial derivatives. The integration is implicit because the updated values $f^{n+1}_i$ are used in the source term. For general non-linear differential equations, implicit integrations are not guaranteed to be stable for arbitrarily large time-steps. Our equations, however, seem...
stable for arbitrarily large time-steps if the solutions are sufficiently near the steady state, but not necessarily if the solutions are far from the steady state.

Stability of the integration does not imply accuracy, and to maintain reasonable accuracy we use a time-step such that no variable changes by more than 100 per cent in one step:

\[ \Delta t = \epsilon \min \left( \frac{f_i}{y_i} \right). \] (14)

The constant \( \epsilon \) usually need not be much less than unity to have both stability and accuracy, even when far from the steady state. [An exception to this time-step calculation is that, if the time-step is limited by a point in an overshooting zone where some variables become very small and equation (14) ill-defined, we take a time-step of \( \Delta t = \Delta z / a_w \), where \( \Delta z \) is the distance between grid-points.] As the solutions approach the steady state, the time-step increases. As a matter of standard practice, we integrate all solutions to a time of \( 10^{10} \). [For reference, a typical sound crossing time is of order unity.]

We adopt spatial boundary conditions that are consistent with the reflecting boundary conditions used in the GSPH simulations. In the GSPH simulations, particle velocities changed sign when they hit walls of the one-dimensional box, but retained the same temperature. In this case, moments that are odd in \( w \) are zero at the walls of the box, whereas moments that are even in \( w \) have gradients that are zero at walls. Each equation for an odd \( w \) moment has one boundary condition placed at one of the walls. Each equation for an even moment and the corresponding equation defining the gradient require two boundary conditions on the gradient, one at each wall. One consequence of applying these boundary rules rigorously (as is necessary for comparison with the GSPH simulations) is that we have adopted a gravitational acceleration \( g \) which goes to zero at the boundaries. We have boundary regions of 10 per cent of the width of the box over which the gravity falls from unity to zero (cf. section 3.3 of Paper II). Heating and cooling of the fluid occur over these same regions.

The relaxation routines of Press et al. (1986) spatially couple only two neighbouring points at a time, so that spatial derivatives must be taken as first differences. For a spatially centred differencing scheme, odd variables (ones that are zero at the boundaries) would be located at each grid-point, and even variables at half-grid-points, since the derivatives of even variables are odd and the derivatives of odd variables are even. Although the numerical routine does not allow this (since this differencing would couple three points), we can construct the differencing as if even variables are at the half-grid-points instead of the grid-points themselves. Boundary conditions do not present a problem since we never put boundary conditions on even variables. Without this trick to simulate spatial centring of the variables and to give second-order spatial accuracy, the solutions of the moment equations can develop sawtooth oscillations that ultimately prevent convergence. The solutions presented here have been solved on a grid of 50 eveny spaced points, with the two walls representing the first and last points.

The Xiong solution discussed above solves six time-dependent moment equations, and includes an equation of state, and three closure relations. Seven first derivatives are defined, as is the second derivative of \( v \), mostly because they are required for boundary conditions. Thus the minimum number of coupled equations is 17. The Full solution discussed above solves 10 time-dependent moment equations, includes an equation of state, and requires four closure relations, 10 first derivatives, and the second derivative of \( v \). The minimum number of equations is 26. For convenience we have defined a few auxiliary variables, so that we actually solve 29 equations simultaneously.

Our numerical solutions of the moment equations use the same dimensionless units as the GSPH simulations of Paper II. The units are defined by setting the acceleration of gravity \( g = 1 \), Newton’s constant \( G = 1 \), the mean density of the fluid averaged over the entire box \( \bar{\rho} = 1 \), and the ratio of Boltzmann’s constant to specific mass \( k_B/\mu = 1 \) (cf. appendix C of Paper II). In these units, the equation of state is \( P = \bar{\rho} T \).

4 COMPARISONS BETWEEN THEORY AND SIMULATIONS

Convection models have a heat source in the lower boundary region and a heat sink for cooling in the upper boundary region. The cooling rate is adjusted so that the temperature of the fluid is \( T = 1 \) at the bottom. There is about one pressure scaleheight across the box. If the thermal conductivity \( K \) is small enough, the fluid will convect. If \( K \) is within a factor of a few of the value for critical stability, convection is inefficient since radiative diffusion continues to carry a significant fraction of the energy flux. If \( K \) is much smaller, convection is efficient since convection carries nearly the entire energy flux.

Below, we present solutions for two models with constant \( K \) across the box, one in the efficient regime and the other in the inefficient regime. We call these ‘homogeneous models’ because of the constant \( K \), and there are no significant gradients of convective properties on scales less than a pressure scaleheight. By increasing \( K \) from a small value that causes instability to a larger value that gives stability, we construct overshooting models. We present two overshooting models, where the unstable regions correspond to the efficient and inefficient regimes. The stability transition is at the centre of the box, with the overshoot region in the upper half. The details of the particular parameters for each of these models can be found in sections 5.1–5.4 of Paper II.

4.1 ‘Homogeneous’ models

In Fig. 2 we compare the superadiabatic gradient \( \Delta \nabla T \) and the three second moments (actually the velocity and temperature dispersions \( \sigma_v \) and \( \sigma_T \) and the \( \vartheta \) correlation \( w_\vartheta = w_\vartheta(\sigma_v,\sigma_T) \)) of the two homogeneous GSPH simulations of Paper II with the Xiong and Full solutions. Fig. 3 compares the four third moments, and Figs 4 and 5 compare the four fourth moments, unnormalized and normalized, required in the Full solution.

The Xiong third moments, derived from closure equations (2) and shown in Fig. 3, are mostly in rough qualitative agreement with the GSPH simulations in the interior of the box. \( w^3 \) has the right qualitative shape in both cases, but about half the GSPH value since we use \( c = 1 \) instead of the optimum 1.8. The boundary regions of the remaining third moments are grossly in error in the efficient model, and this is reflected in \( \Delta \nabla T \) of Fig. 2 as well. Despite these discrepancies, the second moments of the efficient model, including
Figure 2. The superadiabatic gradient and the three second moments for the efficient and inefficient homogeneous convection models. Shown are the velocity and temperature dispersions, $\sigma_w$ and $\sigma_\theta$, and the velocity-temperature correlation, $w\theta_n = w\theta / \sigma_w \sigma_\theta$. The brackets on $w\theta$ indicate the same ensemble average as overbars in the text, and the subscript $n$ indicates that $w\theta$ has been normalized. The Full solutions have a more local character than the Xiong solutions.

We consider now the Full solution. When the $w\theta$ correlation is unity, as is seen to be the case in Fig. 2 (excepting the boundary regions), the quasi-normal closure predicts normalized fourth moments equal to $3$. As seen in Fig. 5, the GSPH simulations show that the normalized fourth moments actually vary from about 4 at the bottom of the box to about 2 at the top (boundary regions excepted, see also fig. 15 of Paper II), and, in fact, including the diffusion terms in

\[ \langle w\theta_n^4 \rangle = 3 \]

Normalization means dividing by powers of $\sigma_w$ and $\sigma_\theta$ to get non-dimensional ratios. For a pure quasi-Gaussian closure with $\xi = 3$, $\delta = 0$ in equations (3), $w_\theta^4 = 3$, $w^2 w_\theta^2 = 3 w_\theta$, $w^2 \theta^2 = 1 + 2 w_\theta$, and $w\theta^2 = 3 w_\theta$. Thus most fourth moments are not exactly 3 if the $w\theta_n$ correlation is not unity.
equations (3) does precisely this. The unnormalized fourth moments of the Full solution in Fig. 4 are in reasonable quantitative agreement with the GSPH results.

In Fig. 3 the third moments from the Full solution more nearly correspond to the GSPH results, both quantitatively and qualitatively, than the Xiong model. Oddly, the second moments are not also superior, and, in fact, are generally inferior in the boundary regions. The diffusion terms in closure equations (3) have had the effect of suppressing the amplitudes of third moments, improving agreement in Fig. 3. The smaller third moments, however, make the second moments of Fig. 2 more local in character, making agreement worse.

Without the diffusion term in the closure relations, the amplitudes of the third moments become several times larger, making the gradients of third moments correspondingly larger. These gradients impact the second moments by causing depressions adjacent to the boundary regions, especially at the right-hand boundary. In the efficient model, the depression has a depth of about 25 per cent. In the inefficient model, the depression grows to 100 per cent and \( w^2 \) tries to become negative. (Note that hints of such depressions can be seen in \( \sigma_w \),) This is clearly unphysical, and we cannot obtain steady-state solutions for the efficient model except by controlling the amplitudes of third moments using the diffusion term.

It appears that modelling the fourth-moment closures well does not mean that the second moments will be in good agreement with the simulations. It seems that the relations among the second, third and fourth moments predicted by...
the moment equations are not the same as the relations predicted by the GSPH simulations. Indeed, the second moments agree best when the third moments do not agree well.

4.2 Overshooting models

The overshooting models provide better tests of non-local behaviour since convective properties change on a scale shorter than a pressure scaleheight. The second, third and fourth moments for our efficient and inefficient overshooting models are shown in Figs 6–9.

We consider first the Xiong solutions. In Fig. 7 we see that the Xiong closure captures most qualitative features of the GSPH simulations. In particular, the peak values of the third moments occur at about the right place in the efficient model, although the peaks are systematically displaced to the right in the inefficient model. In all cases, however, with the exception of $\bar{\omega}^3$ in the efficient model, the quantitative agreement is rather poor, especially in the lower boundary region. The agreement of the second moments is better, as was the case with the homogeneous models. Of greatest physical relevance is the behaviour of $\sigma_\omega$ and $\bar{\omega}_\theta$ in the overshoot region. The extension of mixing into the overshooting zone and the reversal of the convective flux are predicted well in both cases. Note, however, that the superadiabatic gradient in the convective half of the fluid is not predicted well because of problems in the boundary region. Both the GSPH
and Xiong solutions have $\sigma_\theta$ decreasing faster in the overshooting zone than $\sigma_\omega$, as predicted by equations (6), (7) and (9).

We now consider the Full solutions. In Paper II we showed that the quasi-normal approximation represents the fourth moments reasonably well in the unstable region of the overshooting models, but, in overshooting zones, fourth moments are more complicated, as seen in Fig. 9 (see also fig. 19 of Paper II). Near the stability transition, the normalized fourth moments grow as the $\omega$- and $\theta$-distributions become skewed by non-local effects. Moments even in $\theta$ grow and then decay in the overshooting zone. Moments odd in $\theta$ change sign in the overshooting zone, after about 0.6 for the efficient overshooting model and almost immediately for the inefficient overshooting model. It is remarkable that including the diffusion term in equations (3) reproduces many features in Fig. 9 (although many details are in quantitative error), even though the diffusion term was justified as a numerical necessity, without a physical basis. In Fig. 8, since amplitudes are small in the overshooting zone, the appearance of errors in the overshooting zone is suppressed, although, for the inefficient overshooting model, errors in the unstable region are large.

In the efficient overshooting model the third-moment solutions in Fig. 7 generally agree better with the GSPH results than with the Xiong solution. As with the homogeneous models, the suppressed third-moment amplitudes make the efficient second moments in Fig. 6 more local in charac-

Figure 5. The normalized fourth moments for the efficient and inefficient homogeneous models. For reference, the quasi-Gaussian (Q.-G.) closure is shown.
Figure 6. The superadiabatic gradient and the three second moments for the efficient and inefficient overshooting models. The spikes near the walls seen in some moments are artefacts of the GSPH algorithm and are not physical.

Indeed, the second moments of the Xiong solution seem superior to those of the Full solution. It appears that better agreement between the Full and GSPH second moments could be achieved by reducing the magnitude of the diffusion terms in equations (3). Of course, the third moments would then be worse.

For the inefficient overshooting model, the third moments of the Full solution are systematically too big. The second moments, however, are in reasonable agreement, and the scale of overshooting predicted by the slope of \( \sigma_z \) is approximately correct. Increasing the magnitude of the diffusion terms in equations (3) could improve agreement of the third moments, but would degrade the second moments.

If we solve the equations using a pure quasi-normal closure, that is, without the diffusion of third moments, the resulting third-moment curves have qualitatively similar shapes, but roughly twice the amplitude. The larger third-moment gradients try to make \( w^2 \) and \( \theta^2 \) negative near the left boundary, causing the numerical scheme to fail. Without the diffusion term, we cannot obtain steady-state solutions for either the efficient or inefficient overshooting models.

Modelling the fourth moments reasonably well, as in the efficient overshooting model, has not caused the third moments to agree with the GSPH results to a comparable degree. Likewise, modelling the fourth moments very badly, as in the inefficient overshooting model, has not caused the...
third moments to be comparably bad. Even though the detailed agreement of the third moments is not great for the Xiong or Full solution, the agreement of the second moments is considerably better. As with the homogeneous models above, the relations among the GSPH moments are not the same as the relations predicted by the Xiong or Full solution.

4.3 Comparison with ballistic trajectories

Theoretical investigators of overshooting have most often performed a ballistic particle sort of analysis, where the equations of motion are integrated until the velocity of an overshooting particle is zero (Shaviv & Salpeter 1973 is a classic example). We do the same here to highlight a fundamental difference between this and the moment equation approach. The equations of motion are

\[ \dot{z} = \omega, \]
\[ \dot{\omega} = g a \theta \dot{T} - (A + B |\omega|) \omega, \]
\[ \dot{\theta} = \Delta \nabla T \omega = (D + E |\omega|) \theta. \]

We have replaced the turbulent damping rates \( B a \omega \) and \( E a \omega \) with \( B |\omega| \) and \( E |\omega| \), since \( \omega \) represents a characteristic turbulent velocity in the ballistic calculation. Furthermore, there is no stable attractor in a convective region unless the equations are made non-linear in this way. The constant \( a \) is the coefficient of thermal expansion defined in Appendix A.

We use the background values of \( \bar{T} \) and \( \Delta \nabla T \) computed for...
the Full solutions, and use as initial values of $w$ and $\theta$ the local solutions of $\alpha_w$ and $\alpha_\theta$.

Results of these integrations for both the efficient and inefficient overshooting models are shown in Fig. 10. In each case we show two integrations, one starting at $z=0.4$ as the solid curves and another starting at $z=0.3$ as the dotted curves. In the top panels we show the evolution in $w$-$e$ phase space. The particles enter the overshoot region with positive $w$ and $e$. In the case of the efficient overshooting model, the particles spiral toward the origin of phase space at the Brunt-Väisälä frequency, $\omega=(ga|\Delta VT|/T)^{1/2}$, with an amplitude decaying at the turbulent damping rate. In the inefficient overshooting model, damping by radiative diffusion is faster than turbulent damping (i.e. $D>|\omega_\nu|$), with a rate comparable to the Brunt-Väisälä frequency, so that particles evolve directly to the origin.

The remaining panels in Fig. 10 show $w$ versus $z$ and $\theta$ versus $z$. In ballistic calculations of overshooting (e.g. Shaviv & Salpeter 1973; Maeder 1975; Langer 1986; Zahn 1991; Schmitt, Rosner & Bohn 1994), the overshooting distance usually is defined by the location where $w$ first reaches zero. If we take the stability transition at $z=0.35$ for the efficient overshooting model and $z=0.4$ for the inefficient overshooting model, as suggested by the non-local $\Delta VT$ curves, overshooting distances are $d_{\text{over}}=0.6\ell$ in both cases. Reversal of the convective flux is predicted where $\theta$ becomes negative, after about $0.3\ell$ and $0.5\ell$ in the efficient and inefficient cases.
The scale for $w$ to reach zero is comparable to the scale for $a_w$ to decay by a factor $e$. The non-local solutions for $a_w$, however, approach zero asymptotically, but never reach it within a finite distance, as the ballistic calculations do. Although an $e$-folding distance might represent a realistic scale for the decay of turbulent velocities, since convective fluids generally mix so much faster than nuclear time-scales, overshooting may be important for many $e$-folding distances. To define overshooting distances, it will be necessary to compare the mixing time-scales with nuclear evolution time-scales. Defining a hard edge to overshooting where $w=0$ is probably misleading, unless overshooting is negligible and convection is essentially local.

5 EQUATIONS MODIFIED FOR SCATTERING

Rather than deal with turbulent losses using eddy viscosity and eddy diffusion rates in the dynamical equations for $w$ and $\theta$, we can treat turbulent losses as a scattering process in our Boltzmann transport formulation of convection. In addition, although we have previously neglected turbulence in the horizontal dimensions, we include it now. Thus we consider the evolution of particles in a $w-\theta$ phase space, where $\theta$ is a horizontal velocity perturbation. The velocity dispersion is now defined as $\sigma = (w^2 + 2u^2)^{1/2}$, and the ensemble of particles is described by the distribution function $f_B(t, z, w, u, \theta)$.

© 1996 RAS, MNRAS 279, 305–336
Figure 10. The ballistic evolution of particles overshooting in a fluid described by the superadiabatic gradient of the Full solution. Plots of $\theta$ versus $w$, $w$ versus $z$ and $\theta$ versus $z$ are shown for both the efficient and inefficient overshooting models. In each case two integrations are shown. The solid curves are for integrations starting at $z=0.4$, at the stability transition, where $w$ and $\theta$ have already started to decay. The dotted curves are for integrations beginning at $z=0.3$. These particles actually penetrate slightly farther in the efficient overshooting model.

The alternative treatment of turbulent losses in the collision term allows for greater flexibility through the particular choices we make in modelling the scattering process. Here we assume that scattering is isotropic in velocity, so that a single parameter $\xi$ defines the width of a Gaussian scattering function for both $w$ and $u$. A parameter $\sigma_w$ defines the width of a Gaussian scattering function for $\theta$. In Paper I we carried out this alternative derivation to the local level only. This approach was used by Narayan, Loeb & Kumar (1994) to study causal diffusion and by Kumar, Narayan & Loeb (1995) to study the interaction of convection and rotation. In Appendix B of this paper, we write the complete moment equations, up to third order. We see that, in the appropriate limit, the second-moment equations are identical to the previous formulation, but the third- and higher-moment equations are necessarily different.

We solve these alternative equations using the same closure schemes as above, but supplemented with a few more closure relations for new moments of the horizontal velocity perturbation $u$. We shall refer to these alternative equations as the Xiong and Full solutions, as before.

(1) Xiong solution. If we close the equations at the level of the third moments using the diffusion approximation, in addition to equations (2) we require another closure,
\[ \overline{wu^2} = -c_4 \sigma_w \nabla \overline{u^2}. \]  
We take the constant $c_4 = 1$, like the other constants in equation (2).

(2) Full solution. The Full solution requires additional fourth-moment closures. Using the same modified form of the quasi-normal approximation as in equations (3), we have
\[ \overline{w^2 u^2} = \xi \overline{w^2} \overline{u^2} / 3 - d \sigma_w \nabla \overline{wu^2}, \]  
\[ \overline{wu^2 \theta} = \xi \overline{w \theta} \overline{u^2} / 3 - d \sigma_w \nabla \overline{wu^2 \theta}, \]  
\[ u^4 = \xi u^2, \]  
\[ \overline{u^4} = \xi \overline{u^2} \overline{\theta^2} / 3. \]
In writing these relations we have used $\overline{\theta} = \overline{wu} = 0$, since there is no preferred horizontal direction. We include diffusion terms on equations (17a) and (17b), since these fourth moments are associated with the turbulent transport term of third-moment equations. The closures of equations

©1996 RAS, MNRAS 279, 305–336
(17c) and (17d) are required for certain terms not directly associated with a particular third moment. As before, we use $\zeta = 3$ for quasi-normal closure, and $d = 0.8$.

The Xiong solution involves seven time-dependent moment equations, and requires an equation of state, four closure relations and 10 derivatives for boundary conditions, making the minimum problem one of 22 coupled equations. The Full solution uses 13 time-dependent moment equations, an equation of state, eight closures, and 18 derivatives for boundary conditions, making 40 equations in all. Because we have added a few auxiliary variables for convenience, we actually solve 43 equations simultaneously.

6 COMPARISONS BETWEEN THE SCATTERING THEORY AND SIMULATIONS

We present results for models using precisely the same four sets of parameters as above. We also present results of simulations from a version of the GSPH code modified to be consistent with this alternative formulation of the moment theory. The new equations and the corresponding modifications to the GSPH code are discussed in detail in Appendix B. We mention the main modifications to the GSPH code here. The evolutionary equations for particles no longer have the eddy damping terms, and particles have a probability of scattering with a characteristic time-scale $\ell/\alpha$. If a particle

---

**Figure 11.** The superadiabatic gradient and four second moments for the efficient and inefficient homogeneous models. Both the moment equations and the GSPH code have been modified to treat turbulent losses as scatterings of turbulent particles. Note that both the vertical and horizontal velocity dispersions, $\sigma_w$ and $\sigma_u$, are shown in the same panels. The set of lines for $\sigma_w$ are always below those for $\sigma_u$. 

©1996 RAS, MNRAS 279, 305–336
does scatter, it does so randomly into distributions of $w$, $u$ and $\theta$ described by parameters $\xi$ and $\sigma$.

To understand the physical implications for various $\xi, \sigma$, we performed GSPH simulations for a range of these parameters. If $\xi = \sigma = 0$, there are large peaks at the origin of the phase-space distribution, since all particles eventually are scattered there. Then any convective flux must be carried by broad tails of the distribution function $f_R(z, w, u, \theta)$. In this case, normalized fourth moments are of order 50, instead of closer to the quasi-normal value of 3. Not only do the GSPH results seem unreasonable if $\xi = \sigma = 0$, but we cannot obtain steady-state solutions to the moment equations in this case. Certain variables which must be positive tend to become negative as the moment equations evolve toward the steady state.

If we use $\xi = \sigma = 0.9$, the normalized fourth moments generated by GSPH simulations are more reasonable, having values of order 10 or less in convective regions. Furthermore, we can obtain steady-state solutions of the moment equations for all four of our model problems. Hence we adopt $\xi = \sigma = 0.9$ in the following calculations. Below we compare these solutions with results of new GSPH simulations. We also consider whether these comparisons are more or less favourable than the comparisons of Section 4.

6.1 'Homogeneous' models

The second moments of the homogeneous models are compared with the new GSPH results in Fig. 11. At the cost of some confusion, we have plotted both the vertical and
horizontal velocity dispersions, $\sigma_w$ and $\sigma_u$, in the same panels. The curves for $\sigma_w$ are always the lower set. The third moments are shown in Fig. 12. The moments $wu^2$ and $u^2\theta$ are not shown to avoid overwhelming confusion. They are always much smaller than third moments with $u$ replaced by $w$. The fourth moments are shown in Figs 13 and 14, but here too we have not plotted moments involving $u$.

The GSPH third moments in Fig. 12 are qualitatively similar to those of Fig. 3. In the efficient model, the Xiong third moments show significant qualitative failures, particularly associated with the boundary regions. The boundary problem is seen also in the $\Delta VT$ curve in Fig. 11. Despite these differences, the second moments, including the boundary regions, are in good qualitative agreement with the simulations. The most noteworthy quantitative difference is that the velocity dispersions $\sigma_w$ and $\sigma_u$ are predicted to be 25 per cent too big. We note that the velocity–temperature correlation is less than unity, as predicted by equation (B25) which gives $w\theta/\sigma_w\sigma_\theta = 0.55$ in the local limit. In the inefficient model, the third moments are in excellent agreement with the simulations and are clearly better than those of Fig. 3. Nevertheless, the agreement of the second moments is not improved by a corresponding degree. Again the velocity dispersions are systematically too big, and $\sigma_u$ does not have quite the right shape. We note also that, although equation (B26) predicts that $w\theta/\sigma_w\sigma_\theta = 1$ in the inefficient regime, the

![Efficient Convection](image-url)

**Figure 13.** The four dominant fourth moments of the efficient and inefficient homogeneous models. We have not shown the moments $w^2u^2$, $wu^2\theta$, $u^2\theta$ and $u^2\theta^2$ to avoid confusion in the figure.

©1996 RAS, MNRAS 279, 305–336

© Royal Astronomical Society • Provided by the NASA Astrophysics Data System
computed value is reduced somewhat, because the large value of $\xi$ has moved this inefficient model closer to the efficient regime (see equation B27).

We now consider the Full solutions. In Figs 13 and 14, the predicted fourth moments are generally smaller than those of the GSPH simulations. The quantitative agreement is, however, worse than in the corresponding Figs 4 and 5. In Fig. 12, certain third moments are improved over the Xiong solution, particularly in the efficient model. The quality of third moments is comparable to that of those in the corresponding Fig. 3. As usual, the second moments in Fig. 11 exhibit too much local character.

We find that third moments can be much improved if $d = 0.2$ in the closure relations instead of $d = 0.8$, restoring some of the non-locality to the second moments. With this change, the fourth moments hardly change at all. Apparently small and subtle changes in the fourth moments can have much more dramatic effects on second and third moments. Big changes in the fourth moments, making them appear in much better agreement, would lead to much worse solutions of the lower moments.

6.2 Overshooting models

Solutions of the scattering theory with $\xi = \sigma = 0.9$ for the overshooting models are shown in Figs 15-18, where we present also the second, third and fourth moments of new GSPH simulations.
We discuss first the Xiong solutions. The Xiong third moments in Fig. 16 capture the behaviour of the GSPH simulations only roughly. As we observed with the original version of the theory in Section 4, however, the agreement of the second moments is better. Nevertheless, we do note a few important discrepancies. The velocity dispersions $\sigma_v$ and $\sigma_w$ are somewhat too big in the convective half of the fluid, and $\sigma_w$ decays too fast in the overshooting zone. In the efficient overshooting model, the superadiabatic gradient is in error in the left boundary region. In the inefficient overshooting model, the shape of the $\omega \theta$ correlation is not captured well, which can be attributed to having $\sigma_\theta$ too small in the overshooting region.

In both overshooting models, $\sigma_\theta$ decays faster in the overshooting region than does $\sigma_v$. The reason is that negative buoyancy and turbulent viscosity decelerate vertical motions, while only turbulent viscosity acts horizontally. Thus the ratio of kinetic energy in the horizontal motions to kinetic energy in vertical motions, $u^2/w^2$, increases in the overshooting zone.

The fourth moments of the Full solution in Figs 17 and 18 are not in good agreement with the GSPH simulations. In
Figure 16. The four dominant third moments of the efficient and inefficient overshooting models. Again we have omitted the curves of third moments involving $\mu$.

Despite these large discrepancies of fourth moments, in Fig. 16 the third moments of the Full solution are in better agreement, although not better than for the corresponding Fig. 7. In the efficient overshooting model, they are clearly superior to the Xiong third moments, and, in the inefficient overshooting model, agreement is comparable (the amplitudes are slightly worse, but the location of the peaks of the curves is better). As usual, the second moments in Fig. 15 have too much local character, as evidenced especially in the efficient overshooting panels. This suggests that decreasing the magnitude of the diffusion term on the fourth-moment closures could improve agreement at the second-moment level. Indeed, if we use $d = 0.2$ instead of $d = 0.8$, we find $\sigma_{\phi}$ and $\sigma_{\theta}$ in better agreement in the inefficient overshooting model. However, the normalized moments do show many qualitatively correct features. In the efficient overshooting model, the difference with $\mu^4$ is particularly severe. The solutions for the inefficient overshooting model are worse.

We note that the GSPH $\mu^4$-curve hardly decays in the overshooting region of the efficient overshooting model. In the overshooting region, $\sigma_{\phi}$ decreases, thus increasing the characteristic time-scale for scattering. Overshooting particles with the largest $\mu$ and $\theta$ can cross the overshooting zone without scattering, accelerating for most of the distance, whereas more typical particles do scatter to smaller amplitudes. Although $\sigma_{\phi}$ is decreasing, the distribution of $\mu$ has an envelope that actually gets broader in the overshooting zone, indicative of broad wings in the distribution of $\mu$ that maintains the magnitude of $\mu^4$ even as $\mu^2$ decreases.
model (although the qualitative shape of the normalized $w\theta$-curve and the amplitudes of the third moments are slightly worse), but we cannot get a converged solution to the efficient overshooting model.

We can describe third moments better using the Full solution, rather than the Xiong closures. As before, however, good third moments do not necessarily imply good second moments. Indeed, there are substantial differences between the Xiong and Full third moments, but the differences between the Xiong and Full second moments are not as great.

7 DISCUSSION AND CONCLUSIONS

We have solved the moment equations describing stellar convection for two, somewhat different, formulations of non-local mixing length theory. In the first formulation, velocity and temperature perturbations were damped at an eddy diffusion rate defined by the mixing length. The second formulation treated turbulent dissipation as a scattering process, where the scattering time-scale is determined by the time required to cross the distance of a mixing length. In both cases we solved the equations using two different closure schemes, called the Xiong and Full solutions. The Xiong solution closes the equations at the third moments using the diffusion approximation. The Full solution closes the equations at the fourth moments with a modified version of the quasi-normal approximation. These various solutions were computed for four models: efficient and inefficient homogeneous convection, and efficient and inefficient overshooting.
Regardless of the closure scheme, certain features of our non-local solutions seem to hold generally. In an overshooting zone, shortly beyond the stability boundary, the convective flux becomes negative. Hence more energy must be transported outward by radiation than in local convection, and, in the overshooting region where $\Delta VT < 0$, the temperature gradient becomes somewhat steeper and more nearly adiabatic. In our solutions, the convective flux does not obtain large negative values in the overshooting zone, and the departure of the temperature gradient from the local radiation value is small. Thus overshooting hardly affects the hydrostatic structure of a star. This conclusion is supported by recent helioseismological results, but is contrary to the claim of many authors that the overshooting zone is nearly adiabatic. If the overshooting zone were nearly adiabatic, overshooting blobs moving against a barely stable temperature gradient could not obtain large temperature deficits that give a large negative flux, as would be required for big departures from the radiative temperature gradient.

In the overshooting zone, the turbulent velocities do not decay as fast as the convective flux. Solutions of the Paper I equations have a turbulent velocity $\sigma_v$ that decays by a factor $e$ in a distance $0.5t - 0.8t$, depending on the model. The scattering equations give a decay scale about twice as large. Turbulent viscosity is important in damping turbulent velocities; the deceleration of negative buoyancy is less important. This is different from most previous overshooting calculations, which regard negative buoyancy as the
dominant or only mechanism for deceleration. These results about the convective flux and large extent of overshooting are qualitatively consistent with those of Xiong (1985) and Xiong & Chen (1992).

Although the moment equations predict a characteristic distance for the decay of turbulent velocity, the turbulent velocity approaches zero asymptotically and the overshooting distance is formally infinite. This is fundamentally different from ballistic particle theories that compute an unambiguous, finite overshooting distance. This difference is a consequence of considering the entire ensemble of turbulent blobs simultaneously, rather than evolving only one blob in a static background. The ballistic calculation does seem useful for defining characteristic scales, although the definite boundaries on the extent of overshooting may be misleading. It is possible that turbulent mixing may be more rapid than relevant nuclear burning time-scales for several scale distances, so that overshooting distances cannot be defined without considering evolutionary time-scales for detailed stellar models. Such considerations go beyond the scope of the present work.

In addition to simply solving the equations, we wanted to demonstrate the internal consistency of our theory by showing that the solutions of the moment equations reproduce the results of GSPH simulations. Since the GSPH code and moment equations rely on the same approximations and same picture of mixing-length convection, we anticipated that, if the closure relations for high-order moments were good enough, then the solutions of the lower moments automatically would agree with the GSPH moments. We have demonstrated broad qualitative agreement, and the quantitative solutions are perhaps reasonable zeroth-order approximations. Nevertheless, detailed comparison with the GSPH results reveals many shortcomings. In particular, fitting high-order moments well apparently does not mean that the lower order moments will exhibit agreement of comparable quality. Conversely, low-order moments may agree well when high-order moments are badly in error.

As a general rule, the Xiong solution predicts second moments better, with third-moment agreement not as good. The Full solution predicts third moments better, but the second moments show inferior agreement. Furthermore, there is no strong reason to prefer the solutions of the Paper I equations solved in Section 4 or solutions of the scattering equations solved in Section 6. Since the second moments \( \omega^2 \) and \( \omega \theta \) are most important for constructing stellar models, we conclude that the Xiong closures perform impressively well.

It is surprising to us that the quality of internal consistency is not much better. Whatever the degree of internal consistency, however, the mixing-length approximations are severe, and an external consistency with the real world is not guaranteed. The GSPH and moment solutions eventually should be compared with and calibrated by three-dimensional hydrodynamic simulations. It is now well-known that compressible convection exhibits qualitative features that the moment theory cannot describe. In particular, convective flows are characterized by deeply penetrating plumes of material, with broader, gentler upflows (Cattaneo et al. 1991; Hossain & Mullan 1991; Hurlburt, Toomre & Massaguer 1986; Stein & Nordlund 1989). Although this horizontal structure is beyond our ability to predict, we do predict the velocity asymmetries that lead to a downward-directed kinetic energy flux in convective regions. Simulations of overshooting (Singh, Roxburgh & Chan 1994) and under­

In summary, we have a theory of non-local convection with enough qualitative success that it may represent a significant improvement over previous theories, but it clearly has significant shortcomings too. The quantitative disagreement with the GSPH results may have two different origins. (1) The closure approximations may not adequately describe the relations among high- and low-order moments. (2) The moment equations themselves may be flawed. We consider these two possibilities in turn.

The closure approximations we use, for both the Xiong solution and the Full solution, are reasonable approximations to the actual high-order moments, as demonstrated in Paper II, but they are not the best relations suggested by the GSPH simulations. Our preferred third-moment closure from Paper II are an alternative form of the diffusion approximation. Those alternative closures, however, do not have the required boundary symmetries, and are not compatible with the method of solution in this paper. The best fourth-moment closure from Paper II related fourth moments to second moments using the quasi-normal approximation, with perturbations by third moments. We have not been able to solve any problems with these optimal closures, and have not been able to solve some with the much simpler pure quasi-normal approximation. Failure to obtain solutions usually means that physically positive second moments, \( \omega^2 \) or \( \theta^2 \), become negative and the numerical method eventually fails to converge. (We note that there is nothing in the mathematical nature of the moment equations that enforces \( \omega^2 \) or \( \theta^2 \) to be positive, or the correlation \( |\omega\theta|/\sigma_\omega\sigma_\theta \) to be less than unity.)

We found that adding a diffusion term to the quasi-normal closure prevents time evolution to unphysical solutions. Regrettably, such a diffusion term was not investigated in Paper II, and we can think of no physical argument justifying it, except that it seems required to obtain numerical solutions. Nevertheless, the third-moment diffusion term seems to introduce qualitatively desirable features into the fourth moments, giving better results than a pure quasi-Gaussian closure. We speculate that, just as each Xiong closure contains only one out of many terms of the complete third-moment closure (given by solution of the third-moment equations), yet gives reasonable results, so too does the third-moment diffusion term, being only one of many (given by solution of the fourth-moment equations), give reasonable results. The third-moment diffusion term has the effect of constraining the amplitude of the third moments, thereby suppressing non-local effects and giving physical solutions. Since we cannot obtain solutions using closures that we claimed in Paper II were very good, we do not attribute the discrepancies between the GSPH results and the solutions directly to the closure relations.

The other possibility, that there is some problem with the moment equations, seems very possible. In this paper, we presented solutions for two different formulations of the
moment theory. Although the two sets of equations are very similar (mainly only certain coefficients are different), there are significant quantitative differences between the two sets of solutions. We have no physical basis for preferring one formulation over the other, which suggests that the coefficients of the various terms of the moment equations may be subject to eventual refinement. Finally, when the equations fail to converge, the problem usually originates in or adjacent to the boundary regions. This may indicate that the main problem with the equations is in the choice of boundary conditions, which were adopted because no other set of self-consistent boundary conditions was apparent for the GSPH code.

Finally, we have not yet computed models that describe realistic stars. The key modification will be the use of realistic opacities in the computation of a self-consistent thermal conductivity $\overline{K}$. It is reasonable to expect, however, that the qualitative features discussed here, namely a nearly radiative overshooting zone of significant extent, will remain true in more detailed models. Although we are in agreement with Xiong (1985) and Xiong & Chen (1992) on these points, we are in conflict with the conclusions of Stothers (1991) and Stothers & Chin (1992), who argue that overshooting is not required to understand various features of stars and star clusters. We have not included the effects of rotation (cf. Kumar et al. 1995) or magnetic fields in the physics of our moment equations, and, if Stothers’ conclusions prove true, then these simplifications may explain the discrepancy. At this time, however, the empirically determined extent of overshooting remains controversial. The results of this series of papers provide a way to make predictions for many different contexts, and it is possible that authors who favour large overshooting and those who favour small overshooting may both be right if they are considering convection in different contexts.

Acknowledgments

I thank Ramesh Narayan for his participation in this long-term project. With respect to this paper, certain key ideas about the numerical method and ideas about the scattering formulation of the moment equations are due to him. I am grateful to James Chiang for offering his careful reading of and comments on the manuscript, and to Brian Chaboyer for providing me with a state-of-the-art solar model.

References

Alongi M., Bertelli G., Bressan A., Chiosi C., 1991, A&A, 244, 95
Basu S., Antia H. M., Narasimha D., 1994, MNRAS, 267, 209
Bertelli G., Bressan A. G., Chiosi C., 1985, A&A, 150, 33
Böhm-Vitense E., 1958, Z. Astrophys., 46, 108
Bressan A. G., Bertelli C., Chiosi C., 1981, A&A, 102, 25
Canuto V. M., 1992, ApJ, 392, 218
Canuto V. M., 1993, ApJ, 416, 331
Canuto V. M., Minotti F., Ronchi C., Ympa R. M., 1994, J. Atmos.
Sci., 51, 1605
Cattaneo F., Brummell N. H., Toomre J., Malagoli A., 1991, ApJ, 370, 282
Chin C.-w., Stothers R. B., 1991, ApJS, 77, 299
Chiosi C., Bertelli G., Meylan G., Ortolani S., 1989, A&A, 219, 167
Doom C., 1982a, A&A, 116, 303
Doom C., 1982b, A&A, 116, 308
Grossman S. A., Narayan R., 1993, ApJS, 89, 261 (Paper II)
Grossman S. A., Narayan R., Arnett D., 1993, ApJ, 407, 284
Hessain M., Mullan J. D., 1991, ApJ, 380, 631
Hurlburt N. E., Toomre J., Massaguer J. M., 1986, ApJ, 311, 563
Hurlburt N. E., Toomre J., Massaguer J. M., Zahn J.-P., 1994, ApJ, 421, 245
Kumar P., Narayan R., Loeb A., 1995, ApJ, 453, 480
Langer N., 1986, A&A, 164, 45
Lester M., 1987, Turbulence in Fluids. Martinus-Nijhoff, Boston
Maeder A., 1975, A&A, 39, 303
Maeder A., Mermilliod J.-C., 1981, A&A, 93, 136
Maeder A., Meynet G., 1988, A&AS, 76, 411
Maeder A., Meynet G., 1989, A&A, 210, 155
Mermilliod J.-C., Maeder A., 1986, A&A, 158, 45
Monteiro M. J. P. F. G., Christensen-Dalsgaard J., Thompson M. J., 1994, A&A, 283, 247
Narayan R., Loeb A., Kumar P., 1994, ApJ, 431, 359
Orszag S. A., 1977, in Balian R., Peube J.-L., eds, Fluid Dynamics.
Gordon & Breach, New York, p. 235
Press W. H., Flannery B. P., Teukolsky S. A, Vetterling W. T., 1986,
Numerical Recipes. Cambridge Univ. Press, Cambridge
Renzini A., 1987, A&A, 188, 49
Roeburgh I. W., 1978, A&A, 65, 281
Roeburgh I. W., 1989, A&A, 211, 361
Roeburgh I. W., 1992, A&A, 266, 291
Schmitt I. H. M. Rosner R., Bohn H. U., 1994, ApJ, 282, 316
Shaviv G., Salpeter E. E., 1973, ApJ, 184, 191
Singh H. P., Roeburgh I. W., Chan K. L., 1994, A&A, 281, L73
Stein R. F., Nordlund A., 1989, ApJ, 342, L95
Stothers R. B., 1991, ApJ, 383, 820
Stothers R. B., Chin C.-w., 1992, ApJ, 390, 136
Travis, L. D., Matsushita S., 1973, ApJ, 180, 975
Umezu M., 1992, MNRAS, 258, 107
Unno W., Kondo M.-a., 1989, PASJ, 14, 197
Unno W., Kondo M.-a., Xiong D.-r., 1985, PASJ, 37, 235
Xiong D.-r., 1980, Chinese Astron., 4, 234
Xiong D.-r., 1981, Sci. Sinica, 24, 1406
Xiong D.-r., 1985, A&A, 150, 133
Xiong D.-r., 1986, A&A, 167, 239
Xiong D.-r., 1989, A&A, 213, 176
Xiong D.-r., 1990, A&A, 232, 21
Xiong D.-r., Chen Q. L., 1992, A&A, 254, 362
Zahn J.-P., 1991, A&A, 252, 179

Appendix A: The Simplified Paper I Moment Equations

Paper I contains a detailed derivation of the moment equations. In this appendix we present the slightly simplified version of them that we solve in Section 4. Since we are concerned mainly with steady-state solutions, we drop terms that include the mean flow velocity $\overline{v}$ (except for the viscosity term in the momentum equation A2), including all advection terms. The thermal conductivity $\overline{K}$ and heating rate $\overline{Q}$ may be functions of vertical position $z$, but do not have any explicit dependence on thermodynamic variables (so that we can ignore terms with $\overline{K}$ and $\overline{Q}$). The equations are as follows.

©1996 RAS, MNRAS 279, 305-336
Zeroth-moment equation:

\[
\frac{\partial \bar{\rho}}{\partial t} + \frac{\partial}{\partial z} (\bar{\rho} \bar{v}) = 0.
\]  \hspace{1cm} (A1)

\(v\)-moment equation:

\[
\frac{\partial \bar{v}}{\partial t} + g + \frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (P + \bar{\rho} \bar{w}^2) - C \frac{\partial^2 \bar{v}}{\partial z^2} = 0.
\]  \hspace{1cm} (A2)

\(T\)-moment equation:

\[
\frac{\partial \bar{T}}{\partial t} - \frac{1}{\bar{\rho} c_p} \frac{\partial}{\partial z} \left( K \frac{\partial \bar{T}}{\partial z} \right) + \frac{1}{\bar{\rho} c_p} \frac{\partial}{\partial z} (c_p \bar{T}) \bar{w} - \frac{\alpha^2}{\bar{\rho} c_p T} \frac{\partial P}{\partial z} \bar{w} - \frac{1}{c_p} (A + B \sigma_w) \bar{w}^2 - \frac{\bar{Q}}{\bar{\rho} c_p} = 0.
\]  \hspace{1cm} (A3)

\(w\)-moment equation:

\[
\frac{\partial \bar{w}}{\partial t} + \frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^2) + \frac{2}{\bar{\rho} T} \frac{\partial P}{\partial z} \bar{w}^2 + 2 (A + B \sigma_w) \bar{w}^2 = 0.
\]  \hspace{1cm} (A4)

\(w\theta\)-moment equation:

\[
\frac{\partial \bar{w}\theta}{\partial t} + \frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^2 \theta) - \frac{2}{\bar{\rho} T} \frac{\partial P}{\partial z} \bar{w}^2 \theta + 2 (A + B \sigma_w) \bar{w}^2 \theta = 0.
\]  \hspace{1cm} (A5)

\(\theta^2\)-moment equation:

\[
\frac{\partial \bar{\theta}^2}{\partial t} + \frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^2 \theta^2) - 2 \Delta \nabla \bar{T} \bar{w} \theta + 2 (D + E \sigma_w) \bar{\theta}^2 + \frac{2}{\bar{\rho} c_p T} \frac{\partial P}{\partial z} \bar{w} \theta^2 - \frac{1}{c_p} (A + B \sigma_w) \bar{w}^2 \theta = 0.
\]  \hspace{1cm} (A6)

\(w^3\)-moment equation:

\[
\frac{\partial \bar{w}^3}{\partial t} + \frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^3) - \frac{3}{\bar{\rho} T} \frac{\partial P}{\partial z} \bar{w}^2 \theta + 3 (A + B \sigma_w) \bar{w}^2 - \frac{3}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^3) = 0.
\]  \hspace{1cm} (A7)

\(w^2\theta\)-moment equation:

\[
\frac{\partial \bar{w}^2 \theta}{\partial t} + \frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^2 \theta) + \frac{2 \bar{\theta}}{\bar{\rho} T} \frac{\partial P}{\partial z} \bar{w}^2 \theta + 2 (A + B \sigma_w) \bar{w}^2 \theta - \frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^2 \theta) - \frac{\partial \bar{\theta}}{\partial z} (\bar{\rho} \bar{w}^2 \theta) - \frac{2 \bar{w} \theta}{\bar{\rho} T} \frac{\partial P}{\partial z} \bar{w} \theta + \frac{2}{\bar{\rho} c_p T} \frac{\partial P}{\partial z} (\bar{w}^2 \theta - \bar{w}^2 \bar{w} \theta)
\]  \hspace{1cm} (A8)

\(w^3\theta\)-moment equation:

\[
\frac{\partial \bar{w}^3 \theta}{\partial t} + \frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^3 \theta) + \frac{3 \bar{\theta}}{\bar{\rho} T} \frac{\partial P}{\partial z} \bar{w}^2 \theta + 3 (A + B \sigma_w) \bar{w}^2 \theta - \frac{2 \bar{w} \theta}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^3 \theta) - \frac{\partial \bar{\theta}}{\partial z} (\bar{\rho} \bar{w}^3 \theta) - \frac{3 \bar{\theta} \bar{\theta}}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^3 \theta) = 0.
\]  \hspace{1cm} (A9)

\(\theta^3\)-moment equation:

\[
\frac{\partial \bar{\theta}^3}{\partial t} + \frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{\theta}^3) - 3 \Delta \nabla \bar{T} \bar{w} \theta + 3 (D + E \sigma_w) \bar{\theta}^3 - \frac{3 \bar{\theta} \bar{\theta}}{\bar{\rho} c_p T} \frac{\partial P}{\partial z} (\bar{w}^2 \theta^2 - \bar{w}^2 \bar{w} \theta^2) - \frac{3 \bar{\theta} \bar{\theta}}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w} \theta) = 0.
\]  \hspace{1cm} (A10)
In these equations, $g$ is the acceleration of gravity, $c_p$ is the specific heat at constant pressure, and $a = -(\partial \ln \rho / \partial \ln T)_p$ is the coefficient of thermal expansion. The rms turbulent velocity is $\sigma_w = \langle w^2 \rangle^{1/2}$, and the superadiabatic gradient $\Delta T = (\partial \ln T_p / \partial z - \partial T / \partial z)$. The coefficients $A$, $B$, $C$, $D$ and $E$ depend on several, possibly different, length-scales. For simplicity, we make the standard mixing-length assumption that turbulent fluid blobs have the same horizontal and vertical dimensions, $l_H$ and $l_v$, which also are equal to the turbulent damping scales of $w$ and $\theta$, $\ell_w$ and $\ell_\theta$. Thus, in terms of a single mixing-length $\ell$, they are defined as

$$A = 10 v_{mic} / 3 \ell^2,$$

$$B = 2 / \ell,$$

$$C = 4 v_{mic} / 3,$$

$$D = 3 K / \rho c_p \ell^2,$$

$$E = 2 / \ell.$$ 

The coefficient $C$ appears only in the momentum equation (A2), where it is responsible for damping background motion. The combinations $B \sigma_w$ and $E \sigma_w$ define the rates of turbulent (or eddy) damping of $w$ and $\theta$, while $A$ and $D$ define their damping rates by microscopic processes (i.e. by viscosity and radiative diffusion). The ratios of turbulent to microscopic diffusion rates are defined by the Reynolds number $Re = B \sigma_w / A$ and the Peclet number $Pe = E \sigma_w / D$.

**APPENDIX B: THE ALTERNATIVE SCATTERING EQUATIONS**

In this appendix we outline an alternative derivation of the moment equations. In this formulation of the Boltzmann transport theory, turbulent losses are included through the collision term of the transport equation. In Paper I we derived only the local equations for this alternative formulation of the theory. Those results are extended to the non-local level of the theory here. It proves easy to include turbulence in the horizontal as well as vertical dimensions, so we consider a three-dimensional phase space $v_x - v_z - T$, where $v_x$ and $v_z$ are horizontal and vertical velocities, respectively. The distribution function $f_A(t, z, v_x, v_z, T)$ evolves according to

$$\frac{\partial f_A}{\partial t} + \frac{\partial}{\partial v_x} (v_x f_A) + \frac{\partial}{\partial v_z} (v_z f_A) + \frac{\partial}{\partial T} (T f_A) = \Gamma^+ - \Gamma^-.$$  

(B1)

The perturbations with respect to the mean background are $w = v_z - \bar{v}_z u = v_x - \bar{v}_x$ and $\theta = T - \bar{T}$. Since no bulk forces act horizontally, $\bar{v}_x = 0$. The three-dimensional velocity dispersion is $\sigma = (w^2 + 2 u^2)^{1/2}$. We do not require a horizontal spatial coordinate since there are no horizontal gradients in our problem.

The collision term on the right-hand side of equation (B1) is divided into destruction and creation functions. Particles lose their identity as a result of interactions with the rest of the fluid on a characteristic time-scale $\ell / \sigma$ in mixing-length convection, so we model the destruction term as

$$\Gamma^- = 2 B f_A.$$  

(B2)

The constant $B$ sets the rate of collisional scattering. Fluid blobs are not destroyed in collisions, but are scattered to new locations in $v_x - v_z - T$ phase space. Thus the creation function must have the form

$$\Gamma^+ = 2 B f_0,$$  

(B3)

where $f_0$ describes the distribution of velocities and temperatures into which fluid blobs get scattered. The flexibility of this formulation of the theory lies in our freedom to model the function $f_0$.

In the derivation presented here, we parametrize the creation function $f_0$ by its second moments according to

$$\langle w^2 \rangle_0 = \langle u^2 \rangle_0 = \frac{\xi^2 \sigma^2}{3},$$

(B4a)

$$\langle \theta^2 \rangle_0 = \sigma^2 \theta^2;$$

(B4b)

$$\langle w \theta \rangle_0 = \langle w u \theta \rangle_0 = 0,$$  

(B4c)

where $\xi$ and $\sigma$ are new parameters of the theory. Equation (B4a) asserts that particles are scattered isotropically, with a velocity dispersion that is some fraction of the initial velocity dispersion. Equation (B4b) makes the temperature dispersion of scattered particles a fraction of the initial temperature dispersion. The remaining second moments and all higher moments of $f_0$ are chosen...
to be zero. As long as $\xi < 1$, $\sigma \prec 1$, the scattering term will damp turbulent velocity and temperature excesses, and not enhance them.

Here we present mainly the results of the derivations, since a detailed description of the mathematical formalism and underlying philosophy can be found in Paper I.

B1 The $v_z$, $v_x$, $T$ and first-moment equations

The dynamical equations that describe the evolution of a fluid blob are

$$v_z = -g - \frac{1}{\rho} \frac{\partial p}{\partial z} \frac{\alpha \theta}{\rho T} Aw + C \frac{\partial \vec{v}_z}{\partial z^2},$$

(B5)

$$v_x = -Au,$$

(B6)

$$\dot{T} = \frac{\alpha}{\rho c_p} \frac{dP}{dt} \left[ 1 + \frac{\alpha \theta}{T} \right] + \frac{1}{\rho c_p} \frac{\partial}{\partial z} \left( K \frac{\partial T}{\partial z} \right) - D \theta + K_r \frac{\partial^2 T}{\partial z^2} \theta + \frac{1}{c_p} (A + B \sigma_\sigma) w^2 + 2u^2 - \frac{1}{c_p} B \xi^2 \sigma^2 + \frac{C}{c_p} \left( \frac{\partial \vec{v}_z}{\partial z} \right)^2 + \frac{\dot{Q}_x}{\rho c_p} + \frac{\dot{Q}}{\rho c_p}.$$  (B7)

These equations are very similar to the dynamical equations of our original formulation, with the significant difference that there are no turbulent damping terms. Whereas previously vertical velocity perturbations were damped by both microscopic and turbulent viscosity with a term of the form $(A + B \sigma_\sigma) w$, in equation (B5) we have only the microscopic term $Aw$, as similarly in the horizontal velocity equation (B6). Also, whereas temperature perturbations were damped by microscopic (radiative) and turbulent diffusion with a term $(D + E \sigma_\sigma) \theta$, in equation (B7) only the microscopic term $D \theta$ is used. Equation (B7) also has been modified to include correctly the viscous heating associated with the collision term.

The GSPH code, described in detail in Paper II, can accommodate this alternative treatment of turbulent losses easily. The one-dimensional code described in Paper II has been expanded to include horizontal velocities, and the dynamical equations have been modified from the original relations to these alternative ones. Although there was no scattering in the original code, we have added a routine to scatter particles. Particles scatter on a characteristic time-scale $\tau = (2B \sigma)^{-1}$, so the probability for a particle to scatter in a time-step $\Delta t$ is $\Delta t/\tau$. If a particle scatters, it is reassigned a temperature and horizontal and vertical velocities randomly selected from Gaussian distributions with dispersions given by equations (B4a) and (B4b).

The equations for the mean flow are quite similar to the original formulation. Indeed, the continuity equation for $\bar{\rho}$ and the moment equation for $\bar{v}_z$ remain unchanged from Paper I. The temperature equation is

$$\frac{D T}{D t} - \frac{\alpha}{\rho c_p} \frac{DP}{Dt} - \frac{1}{\rho c_p} \frac{\partial}{\partial z} \left( K \frac{\partial T}{\partial z} \right) - \frac{1}{c_p} (A + B (1 - \xi^2) \sigma) \dot{\theta} - \frac{C}{c_p} \left( \frac{\partial \bar{v}_z}{\partial z} \right)^2 - \frac{\dot{Q}}{\rho c_p} - \frac{1}{\rho c_p} \frac{\partial}{\partial z} (\bar{\rho} \bar{v}_x \bar{w} \theta) - \frac{\alpha^2}{\rho c_p} \frac{\partial P}{\rho c_p} \frac{\partial \bar{w} \theta}{\partial z} = 0.$$  (B8)

The only modification is in the term describing the viscous heating resulting from three-dimensional turbulence. The mean horizontal velocity equation is

$$\frac{D \bar{v}_x}{D t} + \frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \bar{w} \bar{u}) = 0,$$

(B9)

but $\bar{v}_x \bar{w}$ is clearly zero since there is no preferred horizontal direction. Thus $\bar{v}_x$, is a constant, which we take to be zero.

B2 The $w$, $u$, $\theta$ and higher-moment equations

We derive the higher-moment equations by considering the evolution of the perturbations $w$, $u$ and $\theta$. These evolution equations are

$$w = -\frac{\alpha \theta}{\rho T} \frac{\partial P}{\partial z} \left[ A + \frac{\partial \bar{v}}{\partial z} \right] + \frac{1}{\rho} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^2),$$

(B10)

$$\dot{u} = -Au,$$

(B11)

$$\dot{\theta} = \Delta \nabla T w + \frac{\alpha^2}{\rho c_p T} \left( \frac{\partial P}{\rho c_p} \frac{\partial \bar{w} \theta}{\partial z} - D \theta \right) + \frac{1}{\rho c_p} \frac{\partial}{\partial z} (\bar{\rho} \bar{w} \theta) + \frac{1}{c_p} (A + B \sigma_\sigma) (w^2 + 2u^2 - \sigma^2).$$  (B12)
The essential modification of the equations from their counterparts in Paper I is the omission of turbulent viscosity and turbulent diffusion losses.

The second-moment equations are the following.

\[ w^2 \text{-moment equation:} \]
\[
\frac{Dw^2}{Dt} + \frac{1}{\rho} \frac{\partial}{\partial z} (\rho w^2 \theta) - \frac{2}{\rho T} \frac{\partial P}{\partial z} w^2 + 2 \left( A + \frac{\partial \bar{v}_z}{\partial z} + 2 \sigma \right) w^2 - \frac{2}{3} \sigma^3 = 0.
\]

\[ w\theta \text{-moment equation:} \]
\[
\frac{Dw\theta}{Dt} + \frac{1}{\rho} \frac{\partial}{\partial z} (\rho w^2 \theta) + \frac{\alpha}{\rho T} \frac{\partial P}{\partial z} \theta^2 + 2 \sigma \left( A + \frac{\partial \bar{v}_z}{\partial z} + 2 \sigma \right) w^2 \theta - \Delta V w^2 \theta + \frac{\alpha^2}{\rho c_p T} \frac{DP}{\partial z} + \frac{K_T}{\rho c_p} \frac{\partial^2 \theta}{\partial z^2} + \frac{Q_T}{\partial c_p} \frac{\partial P}{\partial z} w^2 \theta
\]
\[
-\frac{1}{c_p} (A + \sigma \bar{w}^2) (w^2 + 2 w u^2) = 0.
\]

\[ \theta^2 \text{-moment equation:} \]
\[
\frac{D\theta^2}{Dt} + \frac{1}{\rho} \frac{\partial}{\partial z} (\rho w^2 \theta) - 2 \Delta V w^2 \theta + 2 [D + \sigma \bar{u}^2 (1 - \sigma^2)] \theta^2 - 2 \left( \frac{\alpha^2}{\rho c_p T} \frac{DP}{\partial z} + \frac{K_T}{\rho c_p} \frac{\partial^2 \theta}{\partial z^2} + \frac{Q_T}{\partial c_p} \frac{\partial P}{\partial z} \theta^2
\]
\[
-2 \frac{\alpha^2}{\rho c_p T} \frac{\partial P}{\partial z} w^2 \theta
\]
\[
-\frac{2}{c_p} (A + \sigma \bar{w}^2) (w^2 \theta + 2 w u^2) = 0.
\]

\[ u^2 \text{-moment equation:} \]
\[
\frac{Du^2}{Dt} + \frac{1}{\rho} \frac{\partial}{\partial z} (\rho u^2 v) + 2 (A + \sigma \bar{u}^2) u^2 - \frac{2}{3} \sigma^3 = 0.
\]

Since there is no preferred horizontal direction, the remaining second moments, \( \bar{u} \bar{w} \) and \( \bar{u} \bar{\theta} \), are zero in steady state, as are all moments that are odd in \( u \), and hence we do not bother to write these two moment equations. In the limit that \( \xi = \sigma_2 = 0 \), these equations reduce to exactly the second-moment equations of the original theory, verifying that the two alternative treatments of turbulent losses are equivalent at the local level. In the more general case that \( \xi \) and \( \sigma \) are non-zero, the horizontal velocity dispersion is non-zero and the equations are modified to accommodate three-dimensional turbulence.

The third-moment equations are the following.

\[ w^1 \text{-moment equation:} \]
\[
\frac{Dw^1}{Dt} + \frac{1}{\rho} \frac{\partial}{\partial z} (\rho w^1 \theta) + \frac{3 \alpha}{\rho T} \frac{\partial P}{\partial z} w^1 \theta + \left( 3 A + 3 \sigma \frac{\partial \bar{v}_z}{\partial z} + 2 \sigma \right) w^1 \theta - \frac{3}{\rho} \frac{\partial \bar{w}^2}{\partial z} (\rho \bar{w}^2) = 0.
\]

\[ w^2 \theta \text{-moment equation:} \]
\[
\frac{Dw^2 \theta}{Dt} + \frac{1}{\rho} \frac{\partial}{\partial z} (\rho w^2 \theta) + \frac{\alpha}{\rho T} \frac{\partial P}{\partial z} \theta^2 + \left( 2 A + D + 2 \frac{\partial \bar{v}_z}{\partial z} + 2 \sigma \right) w^2 \theta - \frac{2}{\rho} \frac{\partial \bar{w}^2 \theta}{\partial z} (\rho \bar{w}^2 \theta) - \Delta V w^2 \theta - \frac{\alpha^2}{\rho c_p T} \frac{DP}{\partial z} + \frac{K_T}{\rho c_p} \frac{\partial^2 \theta}{\partial z^2} + \frac{Q_T}{\partial c_p} \frac{\partial P}{\partial z} w^2 \theta
\]
\[
+ \frac{Q_T}{\partial c_p} \frac{\partial P}{\partial z} \frac{w^2 \theta}{\partial z} + \frac{\alpha^2}{\rho c_p T} \frac{DP}{\partial z} \frac{w^2 \theta}{\partial z} - \frac{\alpha^2}{\rho c_p} \frac{DP}{\partial z} (w^2 \theta - w^2 \bar{w} \theta) - \frac{\alpha}{\rho T} \frac{\partial \bar{w}^2 \theta}{\partial z} (\rho \bar{w}^2 \theta) - \frac{1}{c_p} (A + \sigma \bar{w}^2) (w^2 \theta + 2 w u^2) - \bar{w} u^2 = 0.
\]

\[ w^3 \text{-moment equation:} \]
\[
\frac{Dw^3}{Dt} + \frac{1}{\rho} \frac{\partial}{\partial z} (\rho w^3 \theta) + \frac{\alpha}{\rho T} \frac{\partial P}{\partial z} \theta^3 + \left( A + 2 D + 2 \frac{\partial \bar{v}_z}{\partial z} + 2 \sigma \right) w^3 \theta - \frac{3}{\rho} \frac{\partial \bar{w}^2}{\partial z} (\rho \bar{w}^2 \theta) - 2 \Delta V w^3 \theta - \frac{3}{\rho c_p T} \frac{DP}{\partial z} + \frac{K_T}{\rho c_p} \frac{\partial^2 \theta}{\partial z^2} + \frac{Q_T}{\partial c_p} \frac{\partial P}{\partial z} w^3 \theta
\]
\[
+ \frac{Q_T}{\partial c_p} \frac{\partial P}{\partial z} \frac{w^3 \theta}{\partial z} - \frac{3}{\rho c_p T} \frac{DP}{\partial z} \frac{w^3 \theta}{\partial z} - \frac{3}{\rho c_p T} \frac{DP}{\partial z} (w^3 \theta - w^3 \bar{w} \theta) - \frac{2}{\rho} \frac{\partial \bar{w}^2 \theta}{\partial z} (\rho \bar{w}^2 \theta) - \frac{2}{c_p} (A + \sigma \bar{w}^2) (w^3 \theta + 2 w u^2 \theta - \bar{w} u^2) = 0.
\]
\[ \frac{D\bar{\theta}^3}{Dt} + \frac{1}{\rho} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^3 \bar{\theta}^2) - 3 \Delta \nabla T \bar{w} \bar{\theta}^2 - \frac{3}{c_r} \frac{\partial}{\partial z} (A + B_o) \bar{w}^2 \bar{\theta}^2 + 3 \left( \frac{\alpha^2}{\rho c_T} \frac{DP}{Dt} + \frac{K_T}{\rho c_r} \frac{\partial^2 T}{\partial z^2} + \frac{Q_I}{\rho c_r} \right) \bar{\theta}^2 \frac{3 \alpha^2}{\rho c_T} \frac{DP}{Dt} (w \theta - w \theta^2) + (3D + 2Ba) \bar{\theta}^3 = 0. \]  
(\text{B20})

\[ \frac{D \bar{w} u^2}{Dt} + \frac{1}{\rho} \frac{\partial}{\partial z} (\bar{\rho} \bar{w} u^2 \theta) + \frac{\alpha}{\rho c_T} \frac{\partial P}{\partial T} \bar{w} \theta + \left( 3A + \frac{\partial^2 \bar{u}_T}{\partial z^2} + 2Ba \right) \bar{w} u^2 - \frac{u^2}{\rho} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^2) = 0. \]  
(B21)

\[ \frac{D \bar{u}^2 \bar{\theta}}{Dt} + \frac{1}{\rho} \frac{\partial}{\partial z} (\bar{\rho} \bar{w} u^2 \theta) + (2A + D + 2Ba) \bar{u} \bar{\theta} - \Delta \nabla T \bar{w} u^2 - \frac{\alpha^2}{\rho c_T} \frac{DP}{Dt} + \frac{K_T}{\rho c_r} \frac{\partial^2 T}{\partial z^2} + \frac{Q_I}{\rho c_r} \frac{\bar{u}^2 \theta}{\rho T} \]  
\[ = -\frac{\alpha^2}{\rho c_T} \frac{DP}{Dt} (\bar{w} u^2 \theta - w \theta u^2) - \frac{u^2}{\rho} \frac{\partial}{\partial z} (\bar{\rho} \bar{w}^2) - \frac{1}{c_r} (A + B_o) \bar{w}^2 u^2 + 2u^4 - u^2 \sigma^2) = 0. \]  
(B22)

The remaining third moments, \( \bar{u}', \bar{w}'u \) and \( \bar{u}^2 \), are zero in steady state, so we have not bothered to write their moment equations. These third-moment equations are nearly identical to the original ones, except for the modifications for three-dimensional motions and one other significant difference. Because of the form of the scattering term, turbulent losses enter all these equations (and indeed would enter all higher-moment equations) as a term proportional to \( 2Ba \). For example, the damping term of the \( \bar{w}^2 \)-equation is \( (3A + 2Ba) \bar{w}^2 \). In the original formulation, turbulent damping enters all third moments as \( 3Ba \), and, in general, the coefficient corresponds to the order of the equations. Although the two formulations describe the same local behaviour in the limit \( \xi = \sigma = 0 \), the two formulations do not describe the same non-local behaviour in any limit. The differences between these and the original equations seem minor, but the quantitative effects are significant.

We have written the moment equations in complete detail for comparison with those in Paper I. The equations that we actually solve in this paper have had the same simplifications applied as those that yield the equations in Appendix A. Namely, we assume \( \bar{v}_T = 0 \), since this is true in the steady state, and we neglect \( K_T \) and \( Q_T \) terms, and terms with \( DP/Dt \).

### B3 The local limit of the scattering theory

In Paper I we showed that, if \( \xi = 0, \sigma = 0 \), in the local limit these equations are identical to those of the original theory. Here we present results for the more general case of non-zero \( \xi \) and \( \sigma \). Recalling that the local limit is obtained when all third and higher order moments are set equal to zero, we can solve the four second-moment equations (B13)-(B16) to yield a relation between the velocity dispersion and the unstable superadiabatic gradient,

\[ \frac{(A + D + 2Ba)(A + Ba)(1 - \sigma^2)(A + Ba(1 - \xi^2))}{[A + Ba + D + Ba(1 - \xi^2)](A + Ba(1 - \xi^2))} \frac{ga}{T} \Delta \nabla T. \]  
(B23)

If \( \Delta \nabla T \) is stable (i.e. \( \Delta \nabla T < TD/ga \)), the solution is \( \sigma = 0 \). In astrophysical convection, turbulent viscosity is always dominant over microscopic viscosity, so that \( Ba/A \gg 1 \). In this limit, the ratio of horizontal to vertical velocity dispersion is

\[ \bar{u}^2 = \frac{\bar{w}^2}{1 - 2\xi^2}. \]  
(B24)

Thus, if there is no scattering into the horizontal direction (i.e. \( \xi = 0 \)) in equation (B4a), there is no horizontal velocity dispersion. Of course, even if there is no vertical scattering, the vertical velocity dispersion remains finite since buoyancy also generates vertical velocities. If scattering is maximal (\( \xi = 1 \)), \( \bar{u}^2 = \bar{w}^2 \) and the turbulence is isotropic.

The ratio of turbulent to radiative diffusion may be more or less than unity in stars. Convection is efficient if \( Ba/D \gg 1 \), and is inefficient if \( Ba/D \ll 1 \). In the limit of efficient convection, the velocity–temperature correlation is

\[ \bar{w} \theta = \frac{2(1 - \sigma^2)(1 - \xi^2) + \xi^2(1 - \sigma^2)}{2(1 - 2\xi^2)} \]  
(B25)
If $\xi = a^\nu = 0$, this correlation is unity, but approaches zero as $\xi$ and $a^\nu$ go to unity. In the inefficient limit, the correlation is
\[
\frac{\bar{w}\theta}{\sigma_w \sigma_\theta} = 1,
\]
(B26)
independent of $\xi$ and $a^\nu$. In this limit, the Peclet number, the efficiency measure, is

\[
Bo = \frac{gaT\Delta T}{D D^2(1 - \xi^2)},
\]
(B27)
Thus, for a given set of parameters in the inefficient regime, as $\xi$ approaches unity, convection becomes increasingly efficient, and, for large $\xi$, equation (B26) becomes a poor approximation.