








# Coupling 1D stellar evolution with 3D-hydrodynamical simulations on-the-fly – III: stellar evolution at different metallicities

Yixiao Zhou <sup>1,2★</sup>, Jakob L. Rørsted <sup>1,2,3†</sup>, Achim Weiss <sup>4</sup>, Andreas C. S. Jørgensen,<sup>5,6</sup> Cis Lagae,<sup>7,8,9</sup> Luisa F. Rodríguez Díaz,<sup>2</sup> Yaguang Li <sup>10</sup>, Mark L. Winther <sup>2</sup>, Jens R. Larsen,<sup>2</sup> Jørgen Christensen-Dalsgaard,<sup>2</sup> Oleg Kochukhov <sup>11</sup>, Karen R. Pollard<sup>1</sup> and Tanda Li <sup>12,13</sup>

<sup>1</sup>*School of Physical and Chemical Sciences – Te Kura Matū, University of Canterbury, Private Bag 4800, Christchurch 8140, Aotearoa, New Zealand*

<sup>2</sup>*Stellar Astrophysics Centre, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark*

<sup>3</sup>*Aarhus Astronomy Data Centre (AADC), Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark*

<sup>4</sup>*Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany*

<sup>5</sup>*Department of Mathematics, Imperial College London, Exhibition Rd, South Kensington, London SW7 2BX, UK*

<sup>6</sup>*I-X Centre for AI In Science, Imperial College London, White City Campus, 84 Wood Lane, London W12 0BZ, UK*

<sup>7</sup>*Department of Astronomy, Stockholm University Albanova University Center, 106 91 Stockholm, Sweden*

<sup>8</sup>*Department of Physics, University of Warwick, Coventry CV4 7AL, UK*

<sup>9</sup>*Center for Exoplanets and Habitability, University of Warwick, Coventry CV4 7AL, UK*

<sup>10</sup>*Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA*

<sup>11</sup>*Department of Physics and Astronomy, Uppsala University, Box 516, 751 20 Uppsala, Sweden*

<sup>12</sup>*School of Physics and Astronomy, Beijing Normal University, Beijing 100875, China*

<sup>13</sup>*Institute for Frontiers in Astronomy and Astrophysics, Beijing Normal University, Beijing 102206, China*

Accepted 2025 June 4. Received 2025 June 3; in original form 2025 February 14

## ABSTRACT

A major weakness in one-dimensional (1D) stellar structure and evolution modelling is the simplified treatment of convection, which leads to erroneous near-surface stratification and considerable uncertainties in predicted effective temperatures and luminosities of low-mass stars. In a series of preceding works, a novel method for coupling 1D stellar structural models with a grid of 3D surface convection simulations during stellar evolution was developed, at solar metallicity. This 1D–3D coupling method slightly shifts evolutionary tracks relative to standard calculations, meanwhile providing oscillation frequencies that agree more closely with asteroseismic observations. Here, we extend this method to model metal-poor and metal-rich F, G, K-type stars, by implementing interpolations on-the-fly across metallicity ( $-3 < [\text{Fe}/\text{H}] < 0.5$ ) for mean 3D models during stellar evolution. We demonstrate quantitatively that the fundamental stellar parameters modelled within our framework are insensitive to the mixing-length parameter. A 20 per cent change in the mixing-length parameter results in evolutionary tracks with a temperature shift of less than 30 K, compared to a difference of over 200 K in standard evolution calculations. Our extension is validated against eclipsing binary systems with extremely precise observational constraints as well as stars in binaries with asteroseismic data. Using a fixed mixing-length parameter that merely governs convective heat transport in the near-adiabatic layers, the 1D–3D coupling method successfully reproduces most observational constraints for all target stars. Coupling 1D stellar evolution models with 3D simulations greatly reduces uncertainties associated with the choice of atmosphere boundary conditions and mixing-length parameters, hence offering a powerful tool for characterizing stars with seismic measurements and determining ages for globular clusters.

**Key words:** convection – stars: atmospheres – methods: numerical – stars: evolution – stars: interiors – binaries: eclipsing.

## 1 INTRODUCTION

1D models of stellar structure and evolution are among the most important instruments in astrophysics. They are capable of predicting the production and transport of chemical elements (Pinsonneault

1997; Herwig 2005), pulsation frequencies (Aerts, Christensen-Dalsgaard & Kurtz 2010), solar neutrino fluxes (Bahcall & Ulrich 1988), among others. Moreover, they are a major method for determining the age of stars, which is a crucial quantity for revealing the evolution history of our Galaxy (Silva Aguirre et al. 2018; Chaplin et al. 2020) as well as population analysis for star clusters (Martocchia et al. 2018). Nevertheless, the predicting power of stellar evolution calculations is limited by several defects, with the modelling of the near-surface convective region being the most

\* E-mail: [yixiao.zhou@qq.com](mailto:yixiao.zhou@qq.com)

† Formerly Jakob Rørsted Mosumgaard.

notable one for low-mass stars. Most stellar evolution codes adopt the mixing-length theory (MLT; Böhm-Vitense 1958) to describe stellar convection. As a phenomenological theory, the MLT pictures heat transfer in the convective region as discrete ‘fluid parcels’ that travel vertically with a distance determined locally, called the mixing length, before merging into their surroundings. The MLT performs well in the deep convection region where the temperature gradient is very close to adiabatic, but it is well-known to fail near the stellar surface where heat transport is far from adiabatic and asymmetries between upflows and downflows are prominent. For example, stellar structure models of solar-type and giant stars computed with MLT predict systematically higher pressure-mode oscillation frequencies than observations at high frequencies, referred to as the asteroseismic surface effect (Christensen-Dalsgaard et al. 1996). This mismatch partly originates from incorrect pressure and density stratification near the stellar surface. Moreover, the convective velocity given by MLT abruptly drops to zero at boundaries between convective and radiative zones, leading to unphysical hard boundaries without convective overshoot. Additionally, MLT contains several free parameters including the mixing-length parameter  $\alpha_{\text{MLT}}$ , which is one of the most uncertain factors in stellar evolution calculations (Joyce & Tayar 2023). Another associated uncertainty in stellar evolution calculations is the choice of surface boundary condition. Physical quantities at the outermost point of the stellar structural model are determined by the stellar model atmosphere. Choi et al. (2018) found that different atmosphere boundary conditions could lead to  $\sim 100$  K effective temperature offset at the red-giant branch (RGB) of stellar evolution. This adds further complications to the interpretation of observational data for red-giant stars.

Considerable efforts have been made to obtain a more realistic description of convection, by numerically solving the equations of hydrodynamics and radiative transfer in 3D space and time (see Nordlund, Stein & Asplund 2009 and references therein). 3D numerical simulations reveal that convection operates in a fundamentally different way than what is suggested by MLT – the stellar envelope convection is driven by cooling at the surface, which generates finger-like downflows that merge as they descend into deeper convective regions (e.g. Käpylä et al. 2017). These downward plumes move faster but occupy less space relative to upflows, thus resulting in a downward net kinetic energy flux. Solar surface convection simulations have been tested extensively against observations and proven to be superior to their 1D counterpart in all aspects. Without the need for tunable parameters necessary in 1D line formation calculations, Asplund et al. (2000) reported an excellent match when comparing the detailed spectral lines predicted from 3D solar simulations with observations. Pereira et al. (2013) carried out a detailed validation of 3D hydrodynamical models of the solar lower atmosphere and demonstrated that the centre-to-limb variations predicted by the 3D solar model agree remarkably well with corresponding observations, indicating the modelled temperature stratification near the optical depth unity is highly realistic. Motivated by the success of the solar case, grids of 3D surface convection simulations such as the CIFIIST grid (Ludwig et al. 2009), the Trampedach et al. (2013) grid, and the STAGGER grid (Magic et al. 2013a; Rodríguez Díaz et al. 2024) have been computed, covering the main-sequence and giant phase of low-mass stars. These 3D grids have been applied to diverse topics in the physics of low-mass stars across the Hertzsprung–Russell (HR) diagram, including the analysis of chemical composition (Amarsi, Nissen & Skúladóttir 2019; Wang et al. 2024), granulation and oscillation properties (Zhou et al. 2020; Rodríguez Díaz et al. 2022).

Despite the weakness in describing turbulent convection near the stellar surface, 1D calculations of stellar evolution will not be

replaced by full 3D hydrodynamics simulations in the near future because the latter are computationally not affordable in evolutionary time-scales (Kupka & Muthsam 2017). On the other hand, although realistic theories of turbulent convection have been developed by different groups, such as the non-local mixing-length model of Gough (1977), the non-local time-dependent theory of Xiong (2021) and Li (2012) derived from hydrodynamic equations and turbulence theory, those theories are not implemented in most stellar evolution codes due to their complexity. Moreover, advanced theories of turbulent convection involve more equations for various interconnected physical quantities, making it much more difficult to obtain a converged solution.

To this end, utilizing pre-computed 3D surface convection simulations to improve 1D evolution models is arguably a promising way forward. A pioneering effort in this direction was made by Ludwig, Freytag & Steffen (1999), who calibrated mixing-length parameters at various effective temperatures  $T_{\text{eff}}$ , surface gravity  $\log g$  and chemical compositions based on the entropy profile of their 2D convection simulations. Similar  $\alpha_{\text{MLT}}$  calibrations were later presented by Trampedach et al. (2014b) and Magic, Weiss & Asplund (2015), based on the Trampedach et al. (2013) and STAGGER grid of 3D simulations respectively. It is worth noting that mixing-length calibrations based on 3D surface convection simulations all predict that  $\alpha_{\text{MLT}}$  of cool dwarfs are larger than the solar value, whereas  $\alpha_{\text{MLT}}$  of warm main-sequence-turn-off stars and red giants is smaller (Trampedach et al. 2014b; Magic et al. 2015), indicating less efficient convection. A different calibration of  $\alpha_{\text{MLT}}$  based on stars in binary systems with asteroseismic data by Li et al. (2018), however, concluded that  $\alpha_{\text{MLT}}$  is more than 10 percent larger for red giants than for the Sun. Furthermore, when comparing effective temperatures from the APOGEE-Kepler catalogue (Pinsonneault et al. 2018) and theoretical predictions, Tayar et al. (2017) discovered a trend between temperature mismatch and [Fe/H] that can only be explained by a positive correlation between  $\alpha_{\text{MLT}}$  and [Fe/H].<sup>1</sup> This positive correlation was supported by Viani et al. (2018) and Li et al. (2024b) who performed  $\alpha_{\text{MLT}}$  calibrations across different metallicities based on observational data, but it was not seen in  $\alpha_{\text{MLT}}$  calibrated from grids of 3D simulations (cf. fig. 6 of Li et al. 2024b). The underlying reason for tensions between mixing-length parameters calibrated from 3D simulations and observations is not understood.

Instead of applying pre-calibrated  $\alpha_{\text{MLT}}$  and  $T(\tau)$  relations, properties of 3D simulation can be used to inform stellar evolution calculations on-the-fly. Jørgensen et al. (2017) developed a robust method to interpolate the horizontal- and time-averaged 3D model atmospheres over the ( $T_{\text{eff}}$ ,  $\log g$ ) plane. The interpolation scheme therefore enables the coupling of the (interpolated) mean 3D stratification with the 1D interior model at every time step of the stellar evolution. In practice, Jørgensen et al. (2018) replaced the near-surface regime of the 1D evolutionary model with the mean 3D model and set the outer boundary condition of the stellar evolution calculation from the mean 3D model at every time step of the stellar evolution. Based on this novel approach of coupling 1D stellar evolution and 3D simulation on-the-fly, Jørgensen et al. (2018, hereafter *paper I*) obtained a new solar structural model and demonstrated that the outermost layers are consistent with the mean 3D solar atmosphere and the coupled model better reproduces observed solar p-mode frequencies. In the

<sup>1</sup>Unless otherwise specified,  $[A/B] = \log(n_A/n_B) - \log(n_A/n_B)_\odot$  where  $n_A/n_B$  and  $(n_A/n_B)_\odot$  represent number density ratio between element A and B in the star and the Sun, respectively.

following paper, Mosumgaard et al. (2020, paper II) further explored how the 1D–3D coupling method would shift evolutionary tracks and reduce the asteroseismic surface effects with respect to the traditional method. However, both previous works were confined to solar metallicity.

In order to extend the 1D–3D coupling method to the modelling of stars with different metallicities with the aim of determining the age of metal-poor stars which are of particular interest in galactic archaeology, in the present work we implement the interpolation of mean 3D models over chemical composition in our stellar evolution code (Section 2). We compare evolutionary tracks computed based on the 1D–3D coupling method with results of standard evolution calculations in Section 3. The 1D–3D coupling method and newly implemented metallicity interpolation are validated against stars with different compositions in eclipsing binary systems (Section 4).

## 2 MODELLING

### 2.1 1D stellar structure and evolution models

Stellar structure models are computed using the Garching Stellar Evolution Code (GARSTEC; Weiss & Schlattl 2008). GARSTEC is among the first generation of programs for numerical simulation of stellar evolution. It solves the equations of mass, energy conservation, hydrostatic equilibrium, and heat transport together at each (spherically symmetric) shell with the Henyey, Vardya & Bodenheimer (1965) method (see also Kippenhahn, Weigert & Weiss 2012 chapter 12.2). Structural models are evolved by evaluating the variation of element abundances throughout the star with time. The relationship between thermodynamic quantities, such as density  $\rho$ , temperature  $T$ , and pressure  $P$ , is described by the equation of state (EOS). GARSTEC is equipped with realistic EOSs for stellar interior calculations, including the OPAL EOS (Rogers & Nayfonov 2002) and `FreeEOS`<sup>2</sup> (Irwin 2004, 2012), which are provided in the form of pre-computed tables for different hydrogen, helium, and heavy element mass fractions. Radiative energy transport is modelled with the diffusion approximation, where the Rosseland mean opacities are supplied from opacity tables tabulated according to their underlying chemical composition. In particular, GARSTEC includes opacity tables with varying  $\alpha$  element abundance, with  $[\alpha/\text{Fe}]$  ranging from  $-0.2$  to  $0.6$  dex. The code has been applied to various topics in stellar physics, such as constructing standard solar models (Vinyoles et al. 2017), detailed modelling of solar-type stars with precise asteroseismic measurements (Silva Aguirre et al. 2015), investigating the helium flash and associated element mixing (Cassisi et al. 2003), and nucleosynthesis in asymptotic giant branch stars (Remple, Battich & Weiss 2024). Meanwhile, the code undergoes continuous development. Recent updates include the implementation of a non-local theory of convection (Ahlborn et al. 2022; Kupka, Ahlborn & Weiss 2022; Braun, Ahlborn & Weiss 2024) and alternative surface boundary conditions obtained from 3D simulations (Mosumgaard et al. 2018; paper I; paper II).

Throughout this study we use the `FreeEOS` equation-of-state. Above  $\log T = 4.1$  (in K), we use OPAL opacity tables (Iglesias & Rogers 1996). At lower temperatures, we gradually switch to the opacity tables of Ferguson et al. (2005). However, as will be detailed below, low-temperature opacities rarely enter into stellar evolution calculations due to the location of the outer boundary in our 1D–3D coupling method. All evolution calculations are based on the

Asplund et al. (2009) metal mixture. For nuclear reaction rates and reaction networks, we adopt default settings in GARSTEC described in Weiss & Schlattl (2008) section 3.2.1.

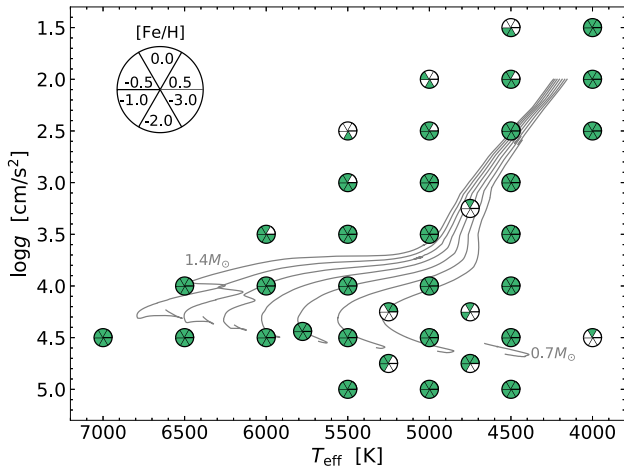
Apart from nuclear reactions, the distribution of chemical elements over time is also changed by element diffusion, which consists of gravitational settling (also called pressure diffusion) caused by the pressure gradient; thermal diffusion caused by the temperature gradient; concentration diffusion sources from chemical inhomogeneity; and radiative levitation associated with radiation pressure. In this work, diffusion from the first three mechanisms is considered for H, He, C, N, O, Ne, Mg, Si, and Fe. Radiative levitation is not implemented in the code. GARSTEC calculates the diffusion coefficient following the method of Thoul, Bahcall & Loeb (1994), and evaluates the change of chemical composition due to nuclear burning and element diffusion simultaneously.

Convective energy transport in the stellar interior is treated by means of the MLT in the formulation of Kippenhahn et al. (2012). It is worth noting that the strong superadiabatic region near the stellar surface, where MLT is well known to be problematic, is not modelled by GARSTEC in our approach but supplied from the 3D model. The outer boundary of the stellar interior model is located below the photosphere, at the near-adiabatic convective layers. Pressure and luminosity at the outermost point, as well as their partial derivatives with respect to temperature and radius, are supplied by interpolating mean 3D model atmospheres at every time step of stellar evolution (cf. Sections 2.2 and 2.3). The choice of outer boundary conditions is the core of our methodology. At other convective boundary layers such as the tachocline located at the bottom of the convective envelope, GARSTEC has the option to account for element mixing caused by convective overshoot (or undershoot) that extends beyond the Schwarzschild boundary of convection. The code employs an exponential decay of overshooting diffusion coefficient as suggested in Freytag, Ludwig & Steffen (1996). The overshoot parameter  $f_{ov}$  controls the  $e$ -folding distance of the diffusion coefficient.

### 2.2 3D model atmospheres and the interpolation technique

The outer boundary conditions of the stellar evolution calculations are provided by the STAGGER-grid – a grid of 3D model atmospheres, also called surface convection simulations, of F, G, K-type dwarf and giant stars originally presented by Magic et al. (2013a) and subsequently refined and extended by Rodríguez Díaz et al. (2024). The grid was generated from the STAGGER code (Nordlund & Galsgaard 1995; Collet et al. 2018; Stein et al. 2024), a versatile magnetohydrodynamics code that solves the equations of mass, momentum, energy conservation, and the induction equation together with the equation of radiative transfer in three spatial dimensions and time. Magnetic fields were ignored in the calculation of STAGGER-grid, whereas radiative energy transport was modelled in detail for every mesh and time step. The spatial dependence of the radiative transfer equation was taken into account by solving the equation at different polar and azimuthal angles, while the frequency dependence of the problem was approximated by the opacity binning method (also called multigroup method, see Magic et al. 2013a section 2 and Collet et al. 2018 section 2.4). The STAGGER-grid was generated based on the Trampedach et al. (2013) version of the Mihalas, Dappen & Hummer (1988) EOS. A comprehensive collection of continuous opacity sources, summarized in Hayek et al. (2010), and line opacities from the MARCS opacity sampling data (Gustafsson et al. 2008; Plez 2008) were used in the radiative transfer calculations. All models adopt the Asplund et al. (2009) metal mixture scaled with

<sup>2</sup>Available at <http://freeeos.sourceforge.net/documentation.html>



**Figure 1.** Kiel diagram showing the global parameters of the STAGGER-grid models, with shaded portions indicating the 3D model atmospheres used in this work (see also fig. 5 of Rodríguez Díaz et al. 2024). Note that the mean effective temperatures of the 3D models are close to, but generally not exact multiples of 500 K. Evolutionary tracks of 0.7–1.4  $M_{\odot}$  stars, whose initial chemical compositions are determined from solar calibration (Table 1), computed based on the 1D–3D coupling method are depicted in the background. The 0.7  $M_{\odot}$  track stops at 20 Gyr.

metallicity. A 0.4 dex enhancement of  $\alpha$ -element abundances was applied to metal-poor models with  $[\text{Fe}/\text{H}] \leq -1$ .

At the current time, the STAGGER-grid consists of more than 220 3D models with  $T_{\text{eff}}$  ranges from 3500 to 7000 K,  $\log(g)/[\text{cm}/\text{s}^2]$  from 1.5 to 5, and  $[\text{Fe}/\text{H}]$  from  $-4$  to 0.5. Unlike  $\log g$  and  $[\text{Fe}/\text{H}]$ , which are input parameters of the simulation, the  $T_{\text{eff}}$  of each simulation snapshot is derived from the radiative flux at the top of the simulation domain. The effective temperature is therefore an emergent property of the simulation fluctuating with time. In this work, we utilize the whole grid except for the  $[\text{Fe}/\text{H}] = -4$  ones and relatively shallow simulations whose lower boundary is located at or near the strong superadiabatic layers (Fig. 1). The simulation domain is a cuboid box with  $240 \times 240$  mesh points in the horizontal plane and 230 meshes along the vertical. Although the characteristic length could differ by several orders of magnitude across the HR diagram, the horizontal area of the simulation was optimally determined such that at least ten granules are enclosed at any time of each simulation (Magic et al. 2013a, section 2.1). Vertically, 3D model atmospheres cover a small but nevertheless important region of the star. All our selected models span from the near-adiabatic convective layers (typically, Rosseland optical depth  $\tau_{\text{Ross}} \sim 10^6$ – $10^7$  at the bottom boundary) to the base of the chromosphere, with vertical size being at least 15 times the photospheric pressure scale height. The fact that STAGGER-grid models extend relatively deep below the surface (compared to most 1D model atmospheres) enables us to match 1D and 3D models at the near-adiabatic region, which makes the stellar evolution calculation insensitive to the value of mixing-length parameter (cf. Section 3.1). Each 3D model contains at least 100 simulation snapshots evenly sampled in time. The time sequence of these simulations covers at least six periods of the main pressure mode (p mode) in the simulation box, corresponding to at least  $\sim 25$  convective turnover time (defined as pressure scale height divided by convective velocity) at the photosphere.

We average 3D model atmospheres in space and time in order to apply them to stellar evolution calculations. Here, the horizontally averaged model is obtained by taking the mean value at constant

geometric depth, i.e. the simple horizontal average. Compared to other averaging methods, such as averaging over constant optical depth, pressures and densities obtained through the simple horizontal and time averaging method very closely obey hydrostatic equilibrium (cf. Magic et al. 2013b fig. A.2 and Zhou et al. 2023 fig. 9). We refer to the simple horizontal- and time-averaged 3D model as mean 3D or  $\langle 3D \rangle$  model hereinafter. The effective temperature of the mean 3D model is the time-averaged value from all simulation snapshots.

Using mean 3D models at every time step of the stellar evolution calculation requires interpolating the model grid at different  $T_{\text{eff}}$ ,  $\log g$ , and metallicities. Jørgensen et al. (2017, 2019) conceived and developed a robust interpolation method to obtain a mean 3D model for any given fundamental stellar parameter covered by the grid. The key component of their interpolation scheme is the ‘density inflection region’ (called density jump in Jørgensen et al. 2017), where the density gradient is relatively small, or, for some stars, a density inversion occurs (e.g. Collet, Asplund & Trampedach 2007, fig. 2). For F, G, and K-type stars, the density inflection is associated with the large superadiabatic temperature gradient at the top of the hydrogen ionization zone just below the stellar surface. Quantitatively, the location of the density inflection corresponds to the local minimum of the partial derivative of density with respect to pressure,  $\partial \ln \rho / \partial \ln P$ . For each mean 3D model, all quantities are scaled based on their value at the density inflection. The scaled structure  $f' = f / f_{\text{di}}$ , where  $f_{\text{di}}$  could be density, pressure, temperature etc. at the location of density inflection, is highly similar across the whole STAGGER-grid (see fig. 2 of Jørgensen et al. 2017), thus ensuring a stable and robust interpolation.

In practice, we first trim all  $\langle 3D \rangle$  models by locating the minimum superadiabatic temperature gradient in the optically thick region then discarding all layers beneath it. The reasoning is that the superadiabatic temperature gradient  $\nabla_{\text{T}} - \nabla_{\text{ad}}$ , where  $\nabla_{\text{T}}$  and  $\nabla_{\text{ad}}$  is the actual and adiabatic temperature gradient respectively, should decrease monotonically when moving deeper into the convection zone, hence an increase of  $\nabla_{\text{T}} - \nabla_{\text{ad}}$  with depth reflects numerical effects from the bottom boundary conditions. At each metallicity, the trimmed  $\langle 3D \rangle$  models define the reference scaled pressures  $P' = P / P_{\text{di}}$  used in the interpolation. To avoid extrapolations, the maximum  $P'$  is dictated by the trimmed model which has the smallest innermost scaled pressure. For given effective temperature and surface gravity, scaled physical quantities  $f'$  are interpolated at each value of scaled pressures. The same procedure is repeated for each metallicity value, followed by a 1D interpolation at the target metallicity to obtain the desired scaled stratification. Finally, the interpolated stratification  $f'(P')$  is multiplied by its value at the density inflection  $f_{\text{di}}$ , which is likewise obtained via interpolation, to get the interpolated  $\langle 3D \rangle$  model.

2D linear interpolation in  $(T_{\text{eff}}, \log g)$  following the technique described above has been implemented in GARSTEC in paper I and paper II. Nevertheless, as the STAGGER-grid has a relatively large interval in effective temperature and surface gravity, the stratification of the interpolated mean 3D model is likely affected by the interpolation method. An optimal interpolation method should give mean 3D structures in best agreement with the ones computed from 3D surface convection simulations. Here, we carry out multiple tests at different  $(T_{\text{eff}}, \log g)$  pairs by taking one  $\langle 3D \rangle$  model out from the grid, constructing an interpolated model at corresponding surface temperature and gravity based on the remaining models, then comparing the interpolated  $\langle 3D \rangle$  model with the actual profile predicted by the 3D simulation. The thus evaluated interpolation errors from the linear and cubic method are shown in Figs A1–A6 in Appendix A for atmosphere parameters corresponding to the Sun,

dwarf and giant stars. Errors below the location of density inflection are the main focus here, since the matching between (3D) and the stellar interior model takes place in the deeper layers. We find that the mean structure of cool dwarfs is better reproduced by linear interpolation. Conversely, cubic interpolation gives smaller errors for warm dwarfs, whose effective temperature is around and greater than the solar value, as well as for RGB stars. In this study, we opt to employ 2D linear interpolation when  $\log g \geq 4$  and 2D cubic interpolation when  $\log g \leq 3.5$ . A smooth transition<sup>3</sup> between results from the two interpolation methods is ensured when  $3.5 < \log g < 4$  to prevent abrupt changes in the interpolated mean 3D profile that may cause convergence problems in GARSTEC. Therefore, our choice of interpolation method is suitable for cool dwarfs and giants. Meanwhile, we caution that systematic errors due to interpolation are likely larger in the high  $T_{\text{eff}}$  and  $\log g$  regions.

The interpolation in ( $T_{\text{eff}}$ ,  $\log g$ ) is performed at each metallicity of the grid, which gives an univariate function between metallicity and physical quantities. We base the metallicity interpolation on the metal-to-hydrogen ratio  $[M/H] = \log_{10}(Z/X) - \log_{10}(Z/X)_{\odot}$  rather than the iron-to-hydrogen ratio  $[Fe/H]$ , where  $X$ ,  $Y$ ,  $Z$  are the mass fraction of hydrogen, helium and heavier elements, respectively. Interpolation in  $[M/H]$  is achieved using the cubic monotonic method of Steffen (1990, see also Paxton et al. 2011). The thus obtained interpolation results always lie within the range of the input data. The metallicity interpolation is implemented in GARSTEC, making the 1D–3D coupling method applicable to the evolution of stars with different  $[M/H]$ , and in addition leads to more consistent atmosphere boundary conditions when considering element diffusion in stellar evolution calculations.

### 2.3 Coupling 1D stellar evolution with 3D model atmospheres

In this section, we summarize how 1D stellar interior models are coupled with (3D) model atmospheres at every evolution time step. Our adopted approach was originally developed in paper I and paper II. We refer the readers to these previous works for more details about the methodology and workflow.

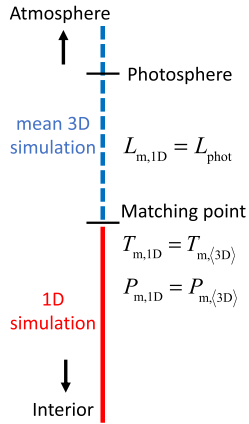
First, we define the location where the stellar interior and (3D) atmosphere models are connected, that is, the matching point. As the matching point serves as the outermost shell of the stellar interior model, ideally it should be placed as deep as the 3D model could cover where the temperature profile is closer to adiabatic and the phenomenological treatment of convective heat transport in 1D models is of less concern. In reality, given that (1) numerical effects from the bottom boundary condition have an impact on the deepest layers in 3D simulations, and (2) the interpolation error increases with depth in deep layers of mean 3D models (Figs A1–A6), an optimal matching point location is therefore determined based on a compromise between these considerations. Here, we follow paper I and paper II to locate the matching point according to pressure at the density inflection. The pressure at the matching point,  $P_{m,(3D)}$ , is always approximately 15.8 times greater than the pressure at the density inflection,  $P_{m,(3D)} = 10^{1.2} P_{di}$ . This criterion ensures the matching between 1D and (3D) models takes place well beneath the optical surface (at  $\tau_{\text{Ross}} \gtrsim 10^4$ ) where the superadiabatic temperature gradient  $\nabla_T - \nabla_{ad}$  is an order of magnitude smaller than the peak value just below stellar photosphere (Fig. 5). The effect of the matching point location on stellar evolution from the 1D–3D coupling method was thoroughly investigated in paper II.

<sup>3</sup>Using a harmonic function with continuous first derivatives.

It is worth noting that the turbulent pressure, which originates from velocity fluctuations, contributes to the total pressure. The effects of turbulent pressure on the evolution of a one solar-mass star and solar oscillation frequencies were studied by Jørgensen & Weiss (2019), who concluded that including turbulent pressure has negligible impact on the evolutionary track. Indeed, for G- and K-type stars, turbulent pressure consists less than 5 per cent of the total pressure below the matching point as predicted by STAGGER-grid models (e.g. Magic et al. 2013a, fig. 21), due to smaller velocity fluctuations in deeper layers. The fraction of turbulent pressure increases in warmer stars, but the ratio between turbulent and total pressure is less than 10 per cent even for the warmest STAGGER-grid models ( $T_{\text{eff}} \approx 7000$  K). Although it could be beneficial to include turbulent pressure in the modelling of early F-type stars and their oscillation frequencies, it is not taken into account in stellar interior modelling in the present work. For consistent couplings between 1D and (3D) models, the turbulent pressure component in (3D) models is also excluded entirely during the interpolation. We therefore do not distinguish between thermal and total pressure and simply symbolize them as  $P$ .

Global stellar parameters  $T_{\text{eff}}$ ,  $\log g$ , and  $[M/H]$  play an important role in the interpolation of mean STAGGER-grid models and specifying outer boundary conditions for GARSTEC. In each iteration, surface gravity is evaluated at the matching point. We note that the distance between the matching point and photosphere is less than 2.5 per cent of the stellar radii (less than 0.5 per cent of the total radii for dwarf stars) for all models except for the  $\log g = 1.5$  ones. Equalizing gravity at the matching point to the surface value is therefore a reasonable approximation and meanwhile consistent with the fact that all atmosphere models were constructed assuming constant gravitational acceleration in the simulation domain. Surface metallicity  $[M/H]$  is also determined at the matching point, which is justified due to the surface convection zone being well mixed and hence having a uniform chemical composition.

Based on the temperature at the outer boundary of the interior model  $T_{m,1D}$ ,  $\log g$ , and  $[M/H]$ , we interpolate the bivariate function  $T_{\text{eff}}(T_{m,1D}, \log g)$ , which is obtained from the relationship between effective temperature and temperature of the matching location for STAGGER models, followed by the metallicity interpolation to obtain the effective temperature. The thus obtained  $T_{\text{eff}}$  is then used to derive the mean 3D model as outlined in Section 2.2. Nevertheless, the interpolated temperature at the matching point,  $T_{m,(3D)}$ , might differ from  $T_{m,1D}$  because of interpolation errors. That is, the interpolation is not reversible. In view of this, we adjust the effective temperature iteratively until the interpolated temperature at the matching point agrees with  $T_{m,1D}$  within 0.01 per cent. The interpolated temperature and (thermal) pressure profile, which extends to the lower stellar atmosphere, is subsequently appended on top of the 1D model. The density stratification above the matching point is evaluated from FreeEOS in GARSTEC based on temperature and pressure of the mean 3D model, rather than directly taken from the mean 3D model, to avoid discontinuities. Having established the density and pressure structure, we integrate the equations of mass conservation and hydrostatic equilibrium outwards to calculate stellar radius and enclosed mass above the matching point. The complete stellar structure model yields total luminosity,  $L_{\text{phot}}$ , from the Stefan–Boltzmann law, where the photospheric radius  $R$  is the location where  $T = T_{\text{eff}}$ . The aforementioned procedure provides two constraints on the stellar evolution calculation. First, pressure at the outermost shell of the interior model must equal to the (3D) value at the matching point,  $P_{m,1D} = P_{m,(3D)}$ . Second, since no nuclear reaction takes place at the near-surface region, luminosity at the top of the 1D model



**Figure 2.** Schematic overview of the 1D–3D coupling approach. The mean 3D model is employed as the outer boundary condition for the stellar evolution calculation. The matching point is located well below the photosphere. The subscript ‘m’ denotes quantity at the matching point. See also Jørgensen & Weiss (2019) fig. 1 for a detailed flowchart of the method.

must equal to the photospheric value,  $L_{m,1D} = L_{\text{phot}}$ . The numerical solver in GARSTEC will iteratively alter the interior structure to fulfill the two outer boundary conditions before advancing to the next evolutionary time step. Therefore, our 1D–3D coupling method ensures continuous pressure and temperature at the matching point, as well as a consistent luminosity between the 1D interior model and the full (patched) model throughout the evolution calculation.

In short, as outlined in Fig. 2, the 1D–3D coupling method replaces the near-surface regime of the 1D evolutionary model with the (interpolated) mean 3D model, and sets the outer boundary condition of the stellar evolution calculation from the mean 3D model in every time step of stellar evolution.

### 3 STELLAR EVOLUTION AND STRUCTURE

#### 3.1 Solar calibration and evolution

A basic requirement of all stellar evolution codes is that given appropriate input parameters, evolution calculation must be able to reproduce the properties of the Sun at solar age, 4.57 Gyr. These strict constraints allow us to calibrate free parameters in the evolution code such as  $\alpha_{\text{MLT}}$ , which is referred to as solar calibration. In GARSTEC, solar calibration is carried out as an iterative process. In each evolution iteration, a  $1 M_{\odot}$  model is evolved from pre-main-sequence to the solar age with a new combination of mixing-length parameter, initial helium, and metal mass fraction. This process is repeated until the modelled luminosity, total radius, and mass fraction ratio  $Z/X$  agree with the corresponding solar values at solar age within a given error range.

Results of the solar calibration based on the 1D–3D coupling approach are summarized in Table 1. The adopted input physics and settings of our evolution calculations are explained in Section 2. We note that considering element diffusion in the construction of solar models improves the agreement between the model-predicted interior structure and that determined from helioseismic data (Christensen-Dalsgaard, Proffitt & Thompson 1993); therefore diffusion is taken into account in our solar calibrations and all subsequent evolution calculations, as outlined in Section 2.1. The change of surface composition over time is treated more consistently between 1D and (3D) model due to the newly implemented interpolation in  $[M/H]$ .

Convective overshoot is not included in the solar calibration, and mass-loss is not considered in this work. Meanwhile, as control groups, we carry out solar calibrations for standard evolution calculations adopting the Eddington grey atmosphere, the Krishna Swamy (1966, KS) analytical  $T(\tau)$  relation fitted to the temperature profile of the solar atmosphere, and the  $T(\tau)$  relation fitted to the Vernazza, Avrett & Loeser (1981, their model C, abbreviated to VAL-C) solar atmosphere (Salaris et al. 2018) as outer boundary conditions. Additionally, GARSTEC is able to carry out evolution calculations using  $T(\tau)$  relations and  $\alpha_{\text{MLT}}$  calibrated from 3D simulations, as implemented in Mosumgaard et al. (2018). We extract the radiative  $T(\tau)$  relations from all solar-metallicity STAGGER grid models according to the formulation of Trampedach et al. (2014a), and adopt mixing-length parameters calibrated by Magic et al. (2015). In this method, the solar calibration sets the scaling factor for the mixing-length parameter, whereas the relative values of  $\alpha_{\text{MLT}}$  remain unchanged compared to Magic et al. (2015). The utilization of 3D-calibrated  $T(\tau)$  relations and  $\alpha_{\text{MLT}}$  improves the near-surface temperature stratification throughout the evolution, but we emphasize that it is fundamentally different from the 1D–3D coupling method focused on in this work (see Section 5). Input physics of the standard evolution calculations as well as the 3D-calibrated  $T(\tau)$  method are identical to the 1D–3D coupling case except for the employed atmospheric boundary condition.

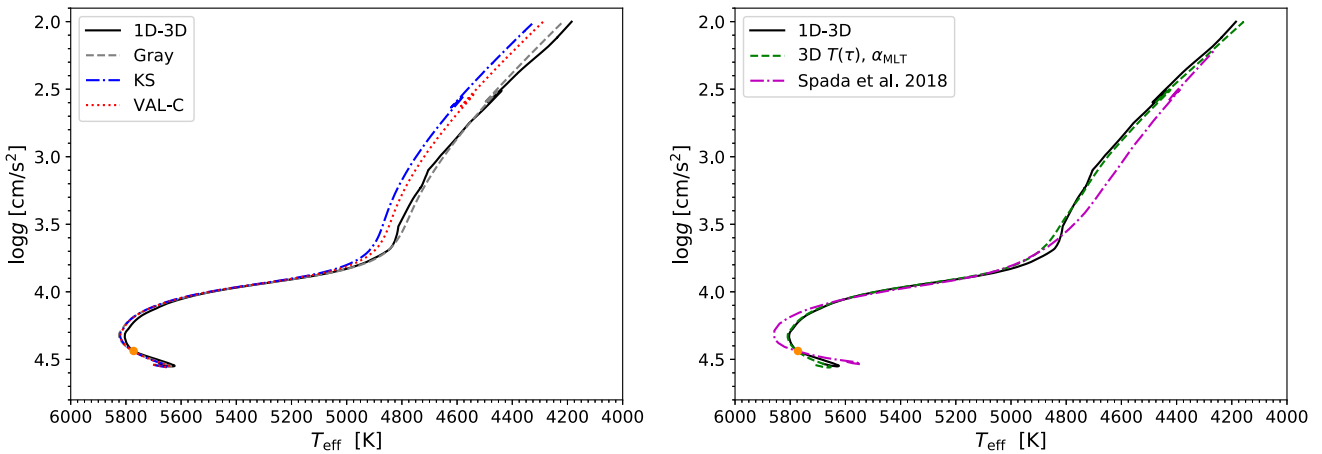
Evolutionary tracks of  $1 M_{\odot}$  stars from the zero-age-main-sequence until  $\log(g/[cm/s^2]) = 2$  with solar-calibrated input parameters are shown in the Kiel diagram (Fig. 3) for these five outer boundary conditions. An additional track adopted from Spada et al. (2018) is shown in the *right panel* of Fig. 3, which was obtained from the entropy calibration method developed in the same study. Spada et al. (2018) adjusted  $\alpha_{\text{MLT}}$  at every time step of stellar evolution so that the adiabatic entropy of the structure model matches that given by the STAGGER grid models (see also Section 5). The effect of the atmospheric conditions is most pronounced on the RGB, where surface temperatures predicted by the KS atmosphere are about 100 K hotter than the case of the grey atmosphere (see also Choi et al. 2018). The evolutionary track given by the 1D–3D coupling method agrees reasonably well with that computed using the grey atmosphere and the 3D-calibrated  $T(\tau)$  method with less than 25 K discrepancy in  $T_{\text{eff}}$  throughout the evolution, albeit a few slight bends are seen in the RGB of the 1D–3D track which will be discussed in Section 3.2. However, the track from the entropy calibration method differs from the 1D–3D coupling result at the main-sequence-turn-off and by up to about 80 K along the RGB, which deserves further investigation as both methods rely on the same 3D simulations. Implementing both methods to the same evolution code is necessary for a convincing identification of the source of the this difference.

A notable result from Table 1 is that the mixing-length parameter calibrated with the 1D–3D coupling method is significantly larger than values from standard evolution calculations. According to MLT,  $\alpha_{\text{MLT}}$  directly impacts the temperature gradient in the convective region. In standard evolution calculations where the model reaches the photosphere, the calibrated  $\alpha_{\text{MLT}}$  governs the difference between  $T_{\text{eff}}$  and temperature at the base of the surface convection zone. However, in our 1D–3D coupling approach,  $\alpha_{\text{MLT}}$  is calibrated to reproduce temperature at the matching point of the mean 3D simulation in which convection is modelled in a fundamentally different approach than MLT. Given that (1) the matching temperature  $T_{m,(3D)}$  differs from temperature predicted by standard evolution calculations at the corresponding location (see Fig. 6), and (2) the superadiabatic temperature gradient at and below the surface and the matching point behaves differently (Fig. 5), mixing-length parameters calibrated

**Table 1.** Solar calibration results based on different descriptions of the outer boundary, including the Eddington grey atmosphere, the empirical  $T(\tau)$  relation by Krishna Swamy (1966) and the  $T(\tau)$  relation fitted to the model C of Vernazza et al. (1981, VAL-C). Key input parameters of evolution calculations obtained from solar calibrations are listed in the lower part.

	Reference	1D–3D coupling	Grey atmosphere	Krishna Swamy (KS)	VAL-C
$T_{\text{eff}}$ (K)	5772 (a)	5772	5772	5772	5772
$R$ ( $10^{10}$ cm)	6.957 (a)	6.957	6.957	6.957	6.957
$X_{\text{surf}}$	0.7381 (b)	0.7490	0.7489	0.7489	0.7489
$Y_{\text{surf}}$	0.2485 (c)	0.2374	0.2375	0.2375	0.2375
$Z_{\text{surf}}$	0.0134 (b)	0.0136	0.0136	0.0136	0.0136
$(Z/X)_{\text{surf}}$	0.0181 (b)	0.01815	0.01815	0.01815	0.01815
$\alpha_{\text{MLT},\odot}$	–	2.77	1.77	2.07	1.98
$X_{\text{init}}$	–	0.7232	0.7232	0.7232	0.7232
$Y_{\text{init}}$	–	0.2622	0.2622	0.2622	0.2622
$Z_{\text{init}}$	–	0.0146	0.0146	0.0146	0.0146

Note. (a) Prša et al. (2016); (b) Asplund et al. (2009); (c) Basu & Antia (2004).



**Figure 3.** Evolution of  $1 M_{\odot}$  stars computed with different treatments of the outer boundary and otherwise identical input physics (see Section 3.1 for the abbreviations used for the outer boundary conditions). The dash-dotted magenta line in the *right panel* is an exception, representing a model computed with the entropy calibration method adopted from Spada et al. (2018, model ESM1 in their fig. 10). For all four tracks in the *left panel*, the mixing-length parameter and initial mass fractions are determined via solar calibration. In the case where the evolution is carried out with 3D-calibrated  $T(\tau)$  relations and mixing-length parameters (green dashed line in the *right panel*), the solar calibration sets the scaling factor (cf. Mosumgaard et al. 2018, equation 2) while relative values of  $\alpha_{\text{MLT}}$  follow those determined by Magic et al. (2015). All tracks pass through the solar parameters (marked with an orange circle) at the present solar age.

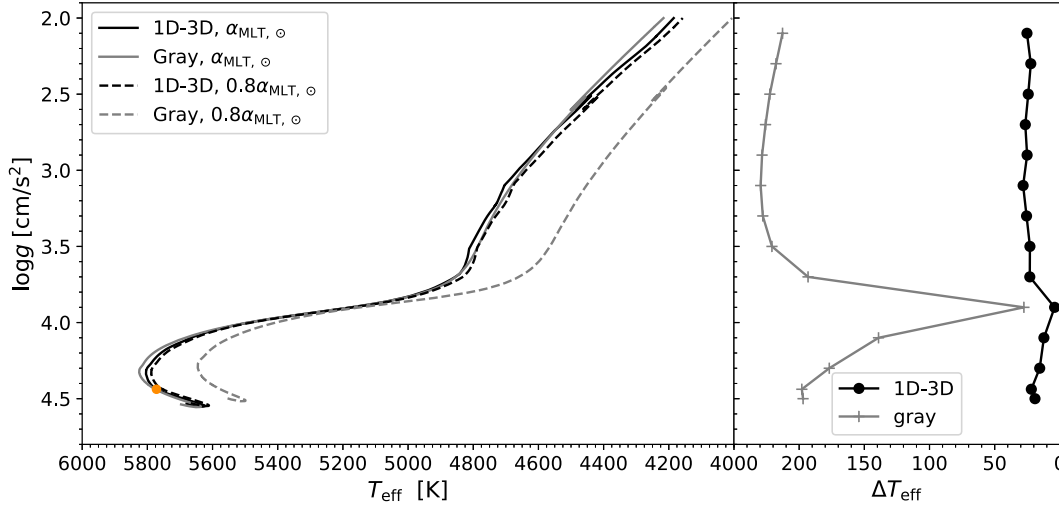
from the two approaches are not comparable.<sup>4</sup> Similar discussions can be found in section 3 of [paper II](#).

One of the major advantages of the 1D–3D coupling method is that it makes evolution calculations insensitive to  $\alpha_{\text{MLT}}$ . As shown in Fig. 4, a 20 percent relative change on  $\alpha_{\text{MLT}}$  shifts  $T_{\text{eff}}$  by less than 30 K at all evolutionary stages. For standard evolutionary calculations, the same relative change alters  $T_{\text{eff}}$  by more than 150 K on the main-sequence and more than 200 K along the RGB. As discussed above and seen in Fig. 5, convection is efficient underneath the matching point and convective fluxes are transported via a small

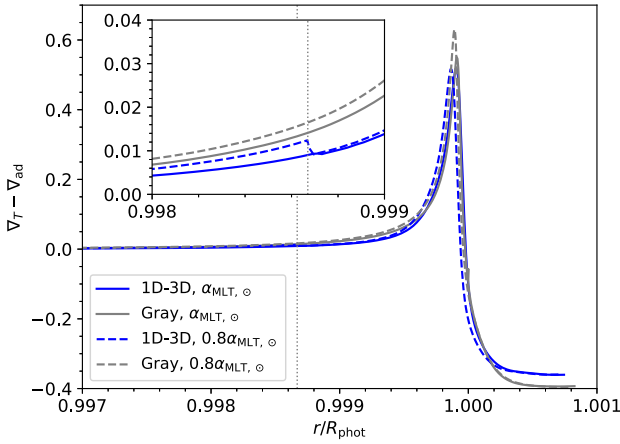
<sup>4</sup>Trampedach & Stein (2011) defined the mass mixing length in the context of hydrodynamics as the turnover distance of convective upflow (or equivalently, downflow). They quantified the ratio between the mass mixing length, being the ‘scale height’ of the mass flux of upflow, and the pressure scale height across their simulation domain and found that the mass mixing-length parameter in the near-adiabatic convective region is greater than its surface value. A higher  $\alpha_{\text{MLT}}$  calibrated from the 1D–3D coupling method appears qualitatively in line with results from 3D hydrodynamical simulations, although the two mixing-length parameters arise from distinct theories of convection and hence are not directly comparable.

$\nabla_T - \nabla_{\text{ad}}$ . Adjusting  $\alpha_{\text{MLT}}$  therefore has a marginal effect on the temperature at the outer boundary of stellar interior models. In the usual scenario where the stellar structure model extends to the photosphere, varying  $\alpha_{\text{MLT}}$  by the same magnitude leaves a larger imprint on the effective temperature and near-surface layers, due to the strong superadiabaticity there (Fig. 5). Although the mixing-length parameter is still necessary for modelling convection in stellar structure calculations, it is regarded as constant and fixed to the solar-calibrated value in our approach. We note that observational uncertainties of most stars surpass variations of stellar parameters caused by changing  $\alpha_{\text{MLT}}$ , therefore the 1D–3D coupling method practically makes stellar evolution independent of the mixing-length parameter.

The near-surface structure of our GARSTEC solar models is compared with the mean 3D STAGGER-grid solar model in Fig. 6 (see also [paper I](#) figs 2–4 and [paper II](#) fig. 5). It is clear that temperature of the 1D–3D coupled model is continuous at the matching point, and the coupled model nearly reproduces the temperature profile of the mean 3D model with less than 0.15 percent disagreement in the optically thick regime, which validates the technical aspect of the 1D–3D coupling scheme. The remaining discrepancy in  $T$  is

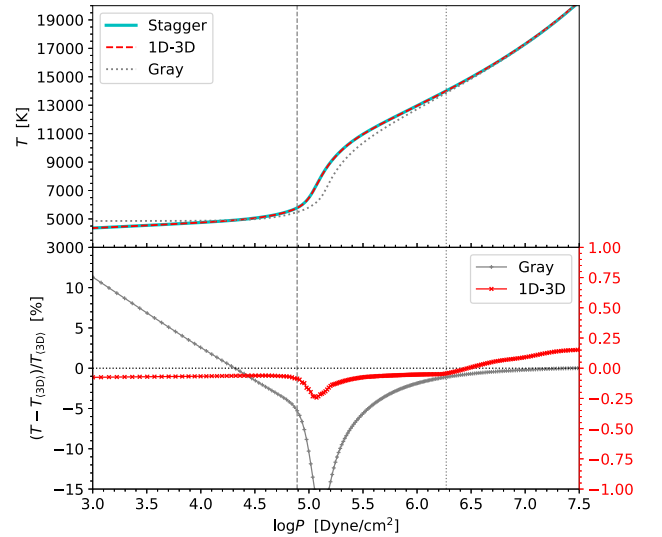


**Figure 4.** The effect of changing  $\alpha_{\text{MLT}}$  on surface properties of stars for the 1D–3D coupling approach and standard stellar evolution with grey atmosphere. Black and grey solid lines in the *left panel* are evolutionary tracks of solar mass star calculated with solar-calibrated  $\alpha_{\text{MLT}}$ , which are also shown in Fig. 3. The sole difference between black (also grey) dashed and solid line is the adopted  $\alpha_{\text{MLT}}$  value. The absolute difference in  $T_{\text{eff}}$  at given surface gravity caused by varying  $\alpha_{\text{MLT}}$  is demonstrated in the *right panel*. In standard evolution calculations, a 20 per cent relative difference in  $\alpha_{\text{MLT}}$  shifts  $T_{\text{eff}}$  by about 200 K on the main sequence and around 225 K along the giant branch. However, changing  $\alpha_{\text{MLT}}$  by the same relative amount results in less than 30 K difference in  $T_{\text{eff}}$  when the 1D–3D coupling method is used.



**Figure 5.** The distribution of superadiabatic temperature gradient in the near-surface region, as predicted by GARSTEC solar models constructed with different outer boundary conditions (blue and grey solid line). Corresponding results from models at solar  $\log g$  (hence solar radius) but calculated with a smaller  $\alpha_{\text{MLT}}$  (see dashed lines in Fig. 4) are shown in dashed lines. The vertical dotted line marks the location of the matching point within the 1D–3D coupling approach, above which the superadiabatic temperature gradient is obtained from the interpolated ⟨3D⟩ models. The superadiabatic temperature gradient peaks just below the photosphere.

due to slight offsets in global parameters between the 1D and 3D model. More realistic near-surface structures lead to more accurate theoretical oscillation frequencies. Indeed, as detailed in [paper I](#) and [paper II](#), a major advantage of the 1D–3D coupling method is that the resulting stellar models predict oscillation frequencies in better agreement with measured values, making this novel method suitable for modelling stars with asteroseismic data (cf. Section 4, [paper II](#), and [Jørgensen et al. 2021](#)).



**Figure 6.** *Upper panel:* temperature profile relative to pressure near the photosphere from the mean STAGGER-grid solar model (cyan solid line), as well as from GARSTEC solar models calculated with 1D–3D coupling method (red dashed line) and the grey atmosphere (grey dotted line). The vertical dashed line indicates the surface location of the solar model from the 1D–3D coupling method, where temperature equals the solar effective temperature. The matching point (cf. Fig. 2) is marked by the vertical dotted line. *Lower panel:* relative difference (in percentage) between GARSTEC and the mean STAGGER solar model. Left vertical axis corresponds to temperature differences when the grey atmosphere is employed in the calculation of stellar structure, while a different vertical axis in red is used for the 1D–3D case to make the difference visible.

### 3.2 Stellar evolution at different metallicities

The wide metallicity range covered by the STAGGER-grid and the interpolation of mean 3D models across  $[M/H]$  implemented in this

work allow us to extend the 1D–3D coupling method to the evolution of stars with different metallicities. Particularly interesting targets are metal-poor stars which are important in Galactic archaeology, but their fundamental properties predicted by standard evolution calculations sometimes contradict values measured from observations (e.g. Creevey et al. 2015, 2024; Joyce & Chaboyer 2018, but see Huber et al. 2024 for evidence supporting the reliability of stellar models at low metallicity).

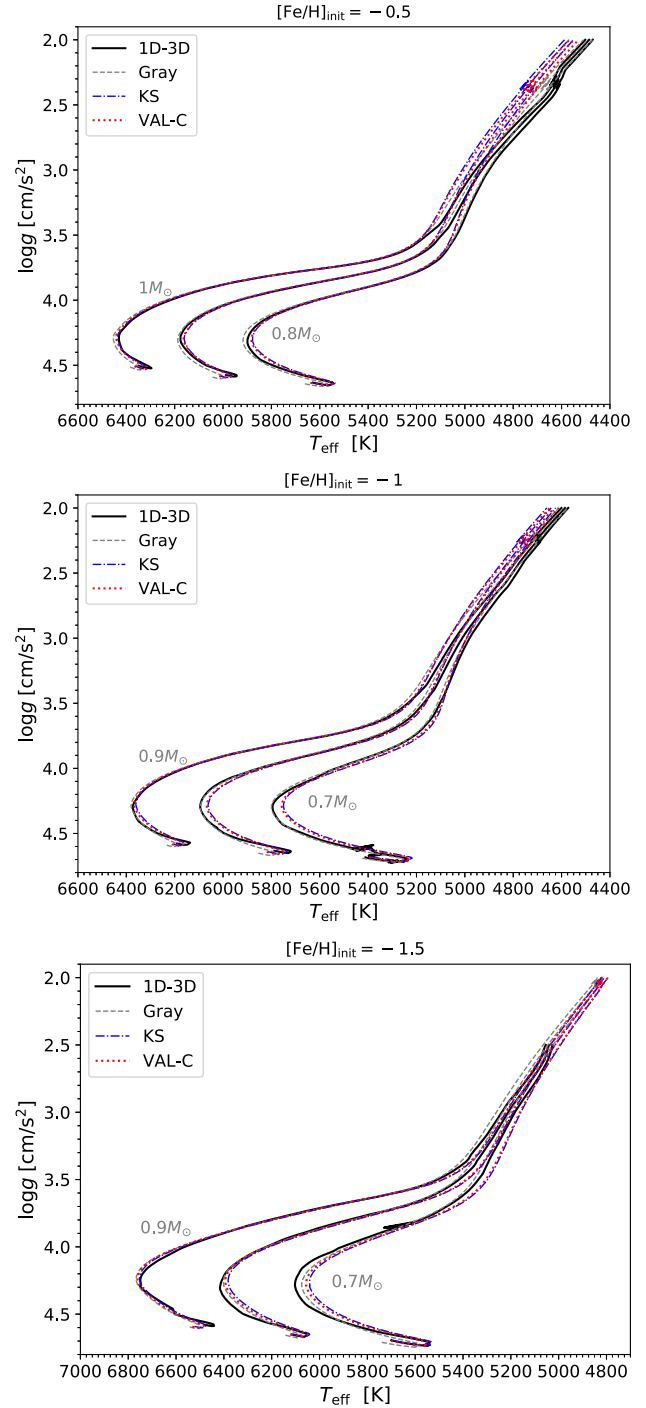
Evolutionary tracks for various masses and initial chemical compositions are shown in Fig. 7. All calculations employ their respective solar-calibrated  $\alpha_{\text{MLT}}$ , and other input physics are identical to the solar case. For a given initial  $[\text{Fe}/\text{H}]$ , the helium mass fraction of the starting model is determined via  $Y_{\text{init}} = Y_0 + (\Delta Y/\Delta Z)Z_{\text{init}}$ , where  $Y_0$  is set to 0.248. The helium enrichment coefficient  $\Delta Y/\Delta Z = 0.97$  is fixed by solving this equation with  $Y_{\text{init}}$  and  $Z_{\text{init}}$  obtained from solar calibration (cf. Table 1). An investigation of the helium enrichment law is beyond the scope of this paper.

It is worth noting that for warmer stars whose  $T_{\text{eff}} \gtrsim 6300$  K, including the standard treatment of element diffusion results in significant or complete depletion of helium and metal by gravitational settling (see fig. 1 of Verma & Silva Aguirre 2019). Heavier elements are easily sedimented in A and F-type stars due to their thin surface convection zone and large pressure gradient near the surface. Such surface abundance patterns predicted by standard diffusion calculations are not in line with helium and metal abundances measured in F-type stars (Varenne & Monier 1999; Verma et al. 2019). Since the metallicity interpolation in our 1D–3D coupling method is based on  $Z/X$ , the strong depletion of metals when  $T_{\text{eff}} \gtrsim 6300$  K is also a practical problem that results in the interpolation scheme picking up unrealistically metal-poor (3D) models or even going out of the metallicity range.

To this end, we implemented the so-called turbulent diffusion in our calculations to counter the effect of gravitational settling. Turbulent diffusion was put forward as an extra mixing term to reproduce the surface abundance of lithium in warm population II stars (Proffitt & Michaud 1991; Richard, Michaud & Richer 2005) and the abundance pattern of turn-off stars in metal-poor globular clusters (Nordlander et al. 2012). It aims to represent the effect of turbulence that might be generated by mild convective overshooting in the radiative regions below the surface convection zone, as stated in Proffitt & Michaud (1991). Here, the turbulent diffusion coefficient is evaluated following equation (1) of Dotter et al. (2017):

$$D_{\text{T}} = D_0 \left( \frac{\rho_{\text{CZ}}}{\rho} \right)^3 \left( \frac{M_{\text{CZ}}}{M} \right)^{-3/2}, \quad (1)$$

where  $\rho_{\text{CZ}}$  and  $M_{\text{CZ}}$  are density at the base of the surface convection zone and the total mass of the surface convection zone, respectively. The free parameter  $D_0$  is set to  $1 \text{ cm}^2 \text{ s}^{-1}$  (Dotter et al. 2017). We apply turbulent diffusion at and under the base of the surface convection zone. If convective overshoot is also included in stellar structure modelling,  $D_{\text{T}}$  is added on top of the diffusion coefficient from overshooting. We emphasize that although turbulent diffusion is conceived based on reasonable assumptions and motivations, it does not have a solid physical basis. Radiative acceleration and rotationally induced mixing are physically well-justified processes that counteract gravitational settling; however it is challenging to incorporate them into stellar evolution calculations (see Richer et al. 1998 and Deal et al. 2020 for efforts in this topic).



**Figure 7.** Evolutionary tracks at different stellar masses and metallicities computed based on the 1D–3D coupling method, as well as three atmospheric boundary conditions frequently used in stellar evolution calculations. For models with initial metallicity  $[\text{Fe}/\text{H}]_{\text{init}} = -1$  and  $-1.5$ , a 0.4 dex enhancement of  $\alpha$ -element abundance is applied in GARSTEC calculations, consistent with the  $\alpha$ -enhanced STAGGER models when  $[\text{Fe}/\text{H}] \leq -1$ . Owing to missing 3D models at  $T_{\text{eff}} = 5000$  K,  $\log g = 2$  (cf. Fig. 1),  $[\text{Fe}/\text{H}]_{\text{init}} = -1.5$  tracks are stopped at  $\log g = 2.5$  to avoid extrapolations.

Fig. 7 shows how surface boundary conditions affect the evolution of metal-poor stars. At all metallicities,<sup>5</sup> evolutionary tracks computed with the 1D–3D coupling approach are similar to that given by the grey atmosphere. Less than 25 K differences in  $T_{\text{eff}}$  between the two types of evolution calculations are found at the main-sequence turn-off, while on the RGB the discrepancy depends on metallicity and the evolutionary stage but does not exceed 40 K in all cases investigated. Overall, the 1D–3D coupling method gives cooler tracks along the RGB than the grey atmosphere scenario. On the other hand, metal-poor tracks computed from the KS and VAL-C atmosphere resemble each other closely, but they are clearly distinguishable from the 1D–3D coupling result. It is worth noting that we find no physical explanation for the similarity in evolutionary tracks between the 1D–3D coupling approach and the grey atmosphere scenario. The latter adopted unrealistic assumptions in stellar atmospheres and predicted significantly different  $T(\tau)$  relations than 3D model atmospheres (see fig. 2 of Zhou et al. 2024). Moreover, the treatment of the outer boundary and surface convection is fundamentally different between the standard evolution calculation and the 1D–3D coupling method. Therefore, we speculate the observed reasonable agreement is a pure coincidence—the results seen in Fig. 7 do not imply that the grey atmosphere is more realistic than KS or VAL-C, nor can they be interpreted as evidence to support the use of the grey atmosphere boundary conditions in standard evolution calculations.

Another notable aspect of the 1D–3D coupling method is that the computed tracks demonstrate discontinuities and bends around certain stellar parameters. This can be seen from zigzags around  $T_{\text{eff}} = 5400$  K,  $\log(g/[\text{cm/s}^2]) = 4.65$  in the middle panel and around  $T_{\text{eff}} = 5700$  K,  $\log(g/[\text{cm/s}^2]) = 3.85$  in the bottom panel of Fig. 7. Bends along the RGB, which were referred to as ‘kinks’ and discussed in detail in paper II, are visible in Figs 1, 3, and 7. Discontinuities in the tracks are certainly unphysical and stem from numerical issues. They are diagnosed to be mainly associated with problems in partial derivatives at the outermost shell of the GARSTEC model. Mean 3D models not only supply pressure and luminosity at the outer boundary but also their partial derivatives with respect to temperature and radius, which are evaluated numerically as well. As derivatives from 2D linear interpolation are not guaranteed to be continuous, large jumps in pressure and/or luminosity derivatives may occur when global parameters of the star vary across the boundary of two triangulations defined in the interpolation scheme. Large variations in partial derivatives will result in the Heney solver converging to a new solution that is far away from the previous time step. Bends along RGB tracks appear from similar underlying reasons. Although the 2D cubic interpolation of Renka & Cline (1984) employed during the red-giant evolution is  $C^1$  continuous, i.e. continuous first derivatives, we observe rapid change of derivatives at the parameter space where the bend (or change of direction) takes place. Apart from making the 3D grid denser as advocated in paper I and paper II, which will likely reduce sudden variations during interpolation, a suitable interpolation method and carefully evaluated partial derivatives are also necessary to make evolutionary tracks smoother. We will focus on these two aspects for further improvements on the 1D–3D coupling method in the future.

<sup>5</sup>We have checked that different treatments of the surface layer hardly affect element diffusion and the evolution of surface abundance. This is also evidenced from the solar calibration that initial and modelled present-day  $X$ ,  $Y$ ,  $Z$  values are insensitive to the outer boundary condition used (Table 1).

## 4 VALIDATION AGAINST BINARY SYSTEMS

Detached binary systems are ideal candidates to test stellar evolution calculations. It is commonly believed that two components are born in the same epoch from a common molecular cloud (see Saffe et al. 2024; however, for evidence pointing to a different scenario) thus should have identical age and initial chemical composition. This characteristic, combined with accurate stellar mass and radius determined from orbits, chemical abundance from spectroscopy, and luminosity or effective temperature from photometry, put strong constraints on stellar evolution modelling. Moreover, oscillating stars in binary systems, which stand on the intersection of two effective approaches for accurate determination of stellar parameters, are valuable targets for calibrating adjustable parameters that govern convection (e.g. Li et al. 2018, 2024b) and internal mixing (Schmid & Aerts 2016; Johnston et al. 2019) in 1D stellar modelling, and for testing the asteroseismic scaling relation (Brogaard et al. 2016; Gaulme et al. 2016). Here, we validate the 1D–3D coupling method against both components in the eclipsing binary AI Phe and two oscillating giants in binary systems observed by *Kepler*, KIC 9970396 and KIC 10001167. Given that stellar evolution using the 1D–3D coupling method is insensitive to the mixing-length parameter (Fig. 4), we fix  $\alpha_{\text{MLT}}$  to the solar-calibrated value throughout. Having invariant  $\alpha_{\text{MLT}}$  and tight constraint on stellar mass, the only major tunable input parameter is the initial chemical composition  $X$ ,  $Y$ , and  $Z$ . With very few ‘degrees of freedom’ in the evolution calculation, we demonstrate below that our models reproduce various observables reasonably well.

### 4.1 AI Phe

AI Phe (HD 6980) is an eclipsing binary system composed of a subgiant and a more evolved star at the base of RGB. As a bright target, it has been studied in detail observationally by, among others, Reipurth (1978) and Maxted et al. (2020), making it an important benchmark for testing stellar evolution models (Spada et al. 2013; Higl & Weiss 2017; Valle et al. 2023). We adopt stellar masses and radii measured by Maxted et al. (2020), who carefully determined geometric and orbital parameters of AI Phe from *TESS* (Transiting Exoplanet Survey Satellite) light curves as well as spectroscopic orbits. Fundamental stellar parameters of each component and their sources are tabulated in Table 2.

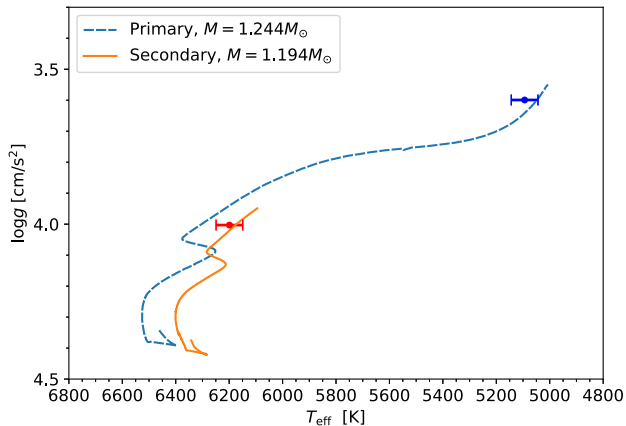
In our modelling, we require the two components to have identical initial chemical composition, and the difference in their age must be less than 1 per cent. For each star, the model-predicted global properties, that is, radius,  $T_{\text{eff}}$  and surface  $[\text{Fe}/\text{H}]$ , should match the measured values within their uncertainties. A grid of evolutionary tracks is computed for each component with fixed stellar mass due to the strong observational constraint. We vary initial  $[\text{Fe}/\text{H}]$  from  $-0.24$  to  $-0.04$  dex, initial helium mass fraction from 0.255 to 0.275, and the overshoot parameter in the stellar structure calculation,  $f_{\text{ov}}$  (cf. Section 2.1), from 0 (no overshoot) to 0.03. Overshooting above the surface convection zone is not modelled by GARSTEC, so  $f_{\text{ov}}$  has no effect in that region.

Given these requirements and relatively few adjustable parameters, the 1D–3D coupling method is able to produce acceptable models for AI Phe. A representative track for both components is shown in the Kiel diagram in Fig. 8, and the corresponding modelled stellar parameters are listed in Table 2. For both components, the modelled  $T_{\text{eff}}$  is in reasonable agreement with measured values. Considering the statistical uncertainties given by Miller et al. (2020) and Valle et al. (2023) and in view of additional systematic errors in the

**Table 2.** Fundamental parameters of the binary system AI Phe from observational data and evolution modelling using the 1D–3D coupling method.

Stellar parameter		$M/M_{\odot}$	$R/R_{\odot}$	$T_{\text{eff}}$ [K]	[Fe/H]	Age ( $10^9$ yr)
Primary	Observation	$1.2438 \pm 0.0008$ (a)	$2.9303 \pm 0.0023$ (a)	$5094 \pm 50$ (b)	$-0.13 \pm 0.1$ (c)	–
	Modelling	1.244	2.934	5047	–0.10	4.73
Secondary	Observation	$1.1938 \pm 0.0008$ (a)	$1.8036 \pm 0.0022$ (a)	$6199 \pm 50$ (b)	$-0.17 \pm 0.1$ (c)	–
	Modelling	1.194	1.803	6177	–0.22	4.69

Note. Reference: (a) Maxted et al. (2020), (b) Miller, Maxted & Smalley (2020), uncertainties quoted from Valle et al. (2023), (c) Andersen et al. (1988).



**Figure 8.** Evolutionary tracks computed using the 1D–3D coupling method, which are good representations of the primary and secondary component of AI Phe, are compared with the corresponding observational data (blue and red dots with error bars). Stellar parameters of the best-matching model are listed in Table 2.

determination of effective temperature, the less than 50 K discrepancy between evolutionary models and observations is acceptable. Another useful perspective is to examine the  $T_{\text{eff}}$  ratio of the two components, a quantity directly determined from the observed light curves (Higl & Weiss 2017). The modelled temperature ratio between the primary and secondary component is 0.817, consistent with the measured value  $0.821 \pm 0.004$  (Miller et al. 2020, their table 5) within uncertainties. The predicted age of the binary system is about 4.7 Gyr, which agrees well with detailed modelling results of Higl & Weiss (2017) and Valle et al. (2023).

Additionally, as the metallicity of the primary and secondary component of AI Phe are determined with the same method and observational data (Andersen et al. 1988), the measured metallicity difference,  $\Delta[\text{Fe}/\text{H}]_{\text{obs}} = 0.04$ , should be realistic. This is analogous to the differential abundance analysis for solar twins that leads to extremely precise element abundance for stars similar to the Sun (e.g. Meléndez et al. 2014). However, our modelling yields  $\Delta[\text{Fe}/\text{H}]_{\text{mod}} = 0.12$ . The significant discrepancy in  $\Delta[\text{Fe}/\text{H}]$  indicates room for improvement in the treatment of element diffusion, including the turbulent diffusion, in our modelling.

#### 4.2 KIC 9970396

KIC 9970396 is a detached eclipsing binary system detected by *Kepler* (Borucki et al. 2010; Prša et al. 2011). It consists of an oscillating red giant and a G or late F companion. Global parameters of the binary and key asteroseismic parameters were first provided in Gaulme et al. (2013, 2016) and subsequently investigated in detail by Brogaard et al. (2018). Furthermore, Li et al. (2018) extracted 40 oscillation frequencies of low-degree modes for KIC 9970396

from *Kepler* photometric data. With well-determined stellar mass and radius from the binary orbit and a wealth of mode frequencies that provide strong constraints on the interior structure, KIC 9970396 is an ideal test bed for stellar structure modelling. The red giant was used as an important target in asteroseismic modelling by Li et al. (2018) and Zhang et al. (2022) for calibrating the mixing-length parameter and stellar age determination, respectively.

In this work, we aim to construct evolutionary models for KIC 9970396 that reasonably reproduce the observed stellar parameters and, more importantly, oscillation frequencies. We consider not only measured radial modes but also the  $l = 1$  and  $l = 2$  mixed modes where  $l$  is the mode degree. Mixed modes, as explained in Hekker & Christensen-Dalsgaard (2017), are found in subgiant and red-giant stars in which propagation cavities of pressure modes and gravity modes (g mode) are coupled (or not well separated). Mixed modes have substantial amplitudes in both the outer region and near the stellar core, demonstrating distinctive characteristics different from p or g modes. Incorporating mixed modes in our analysis provides unique constraints on the core properties and evolutionary stage of our model.

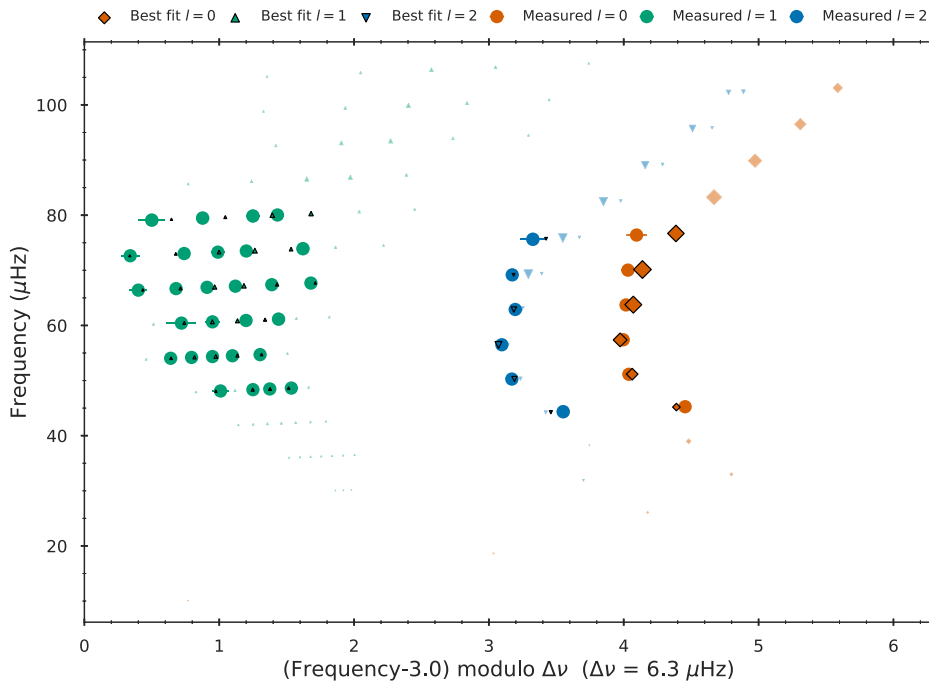
Similar to the case of AI Phe, we computed a grid of evolutionary tracks with different stellar masses, initial metallicities and amount of convective overshoot in stellar interior models. Stellar structural models are selected once the modelled stellar radius approaches the measured value (cf. Table 3). We subsequently compute theoretical frequencies for all selected models using the Aarhus adiabatic oscillation package (ADIPLS; Christensen-Dalsgaard 2008). It is worth noting that for red giants, numerical calculations predict numerous theoretical mixed modes that could potentially be excited. This makes the calculation of the frequencies for the stellar models computationally demanding especially in the context of grid-based modelling. Nevertheless, only mixed modes with low inertia, that is, oscillations that involve relatively small fraction of the stellar mass (often referred to as p-dominated mixed modes), are observable from power spectra (e.g. fig. 5 of Li et al. 2018) since they have a higher amplitude at a given excitation and damping rate. To facilitate straightforward comparison between theoretical and measured mixed mode frequencies, we employ the *truncated scanning method* recently developed by Larsen et al. (2024). As evident from its name, the method utilizes a truncated stellar model with the innermost region removed to estimate the acoustic resonances (the most p-mode-like mixed modes whose frequencies approximately satisfy the acoustic asymptotic relation) within each mode order. They are then used as centrepoints of the observable frequency ranges, for which the complete structure model is subsequently scanned to obtain the observable p-dominated mixed modes. The truncated scanning method outputs much fewer mixed mode solutions, hence significantly reducing the computation cost.

All selected stellar structure models and their corresponding oscillation frequencies computed with the truncated scanning method are subsequently fed into the BAYesian STellar Algorithm (BASTA; Silva Aguirre et al. 2015; Aguirre Børsen-Koch et al. 2022) to find

**Table 3.** Fundamental parameters of KIC 9970396 from observations, the best-fitting model, and median values together with the 16 per cent and 84 per cent limit of the inferred posterior distribution (the 68 per cent Bayesian credibility interval) from BASTA. The frequency of maximum oscillation power,  $\nu_{\max}$ , of the models is estimated from the asteroseismic scaling relation (Brown et al. 1991; Kjeldsen & Bedding 1995).

Stellar parameter	$M/M_{\odot}$	$R/R_{\odot}$	$L/L_{\odot}$	$T_{\text{eff}}$ [K]	[Fe/H]	$\nu_{\max}$ [ $\mu\text{Hz}$ ]	Age ( $10^9$ yr)
Observation	$1.178 \pm 0.015$ (a)	$8.035 \pm 0.074$ (a)	$30.53 \pm 1.85$ (b)	$4825 \pm 85$ (c)	$-0.27 \pm 0.08$ (c)	$63.8 \pm 0.5$ (d)	–
Best model	1.229	8.052	30.83	4793	–0.36	64.5	4.05
Inferred values	$1.214^{+0.020}_{-0.030}$	$8.023^{+0.043}_{-0.068}$	$31.89^{+0.97}_{-1.23}$	$4852^{+39}_{-57}$	$-0.46^{+0.07}_{-0.05}$	$63.6^{+0.6}_{-0.4}$	$4.04^{+0.30}_{-0.22}$

*Note.* (a) Brogaard et al. (2018); (b) Determined based on  $K_s$ -band magnitude (Cutri et al. 2003), *Gaia* DR3 distance (Bailer-Jones et al. 2021), extinction correction (Green et al. 2019), and bolometric correction (Choi et al. 2016), using the procedure implemented in ISOCCLASSY (Huber et al. 2017; Berger et al. 2020). We have accounted for the parallax zero-point offset in the *Gaia* data (Groenewegen 2021) and have inflated uncertainties following the methods proposed by El-Badry, Rix & Heintz (2021); (c) The parameter and its error are averaged from Gaulme et al. (2016, based on spectra from APOGEE) and Brogaard et al. (2018); (d) Li et al. (2018).



**Figure 9.** Échelle diagram of the best-fitting model of KIC 9970396 (Table 3). Here,  $\Delta\nu$  is the large frequency separation. The orange, green, and blue filled circles indicate frequencies of  $l = 0, 1, 2$  modes measured by Li et al. (2018). Theoretical frequencies calculated from ADIPLS using the truncated scanning method are orange rhombuses, green triangles, and blue inverted triangles whose sizes are inversely proportional to their mode inertias. No frequency correction is applied to the theoretical frequencies. Transparent symbols correspond to theoretical frequencies that are not compared with observations.

the model that best represents KIC 9970396. Based on Bayesian statistics, BASTA<sup>6</sup> is a versatile fitting algorithm that infers stellar properties using a given set of observational data which could stem from astrometry, spectroscopy, asteroseismology or combinations thereof. Meanwhile, it is highly flexible in terms of stellar tracks or isochrones adopted in the fitting. In the case of KIC 9970396, the observed radius, effective temperature, as well as all individual frequencies extracted in Li et al. (2018, their table 3) are fitted to the models. For matching the observed individual frequencies with corresponding frequencies in the model, we use the default algorithm in BASTA described in Stokholm et al. (2019) and Aguirre Børsen-Koch et al. (2022).<sup>7</sup>

<sup>6</sup><https://github.com/BASTAcodes/BASTA>

<sup>7</sup>A detailed description of the frequency-fitting algorithm used in BASTA will be presented in Stokholm et al. (in preparation).

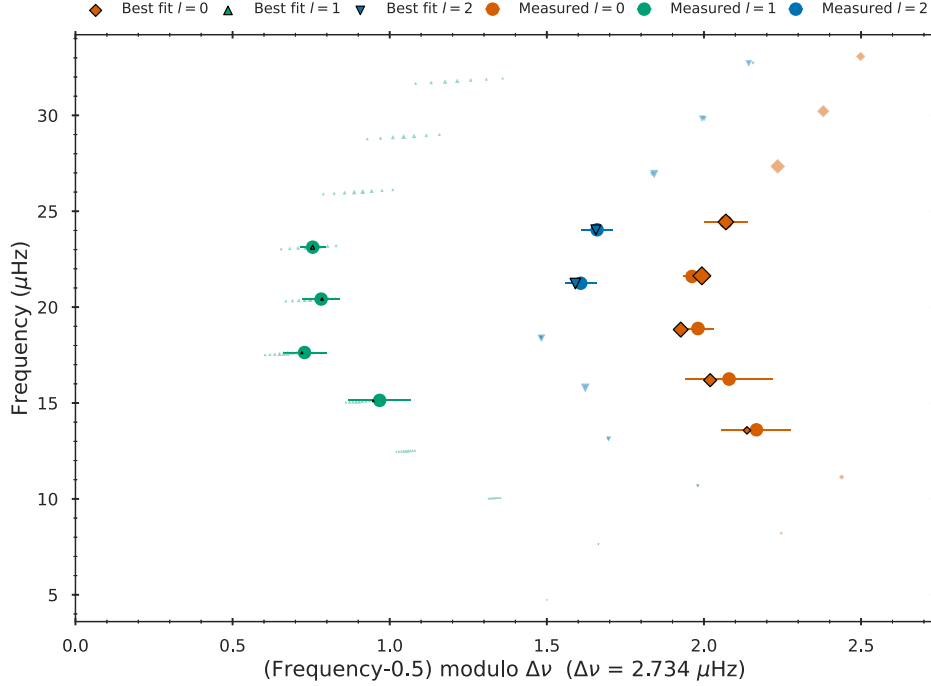
The best-fitting model is defined as the one with maximum likelihood, whose fundamental parameters are listed in Table 3 and theoretical frequencies are compared against observations through the échelle diagram shown in Fig. 9. To avoid seismic measurements dominating the likelihood,  $\chi^2$  computed from seismic data is divided by the total number of frequencies (see Aguirre Børsen-Koch et al. 2022, equation 4). We note that no frequency correction for the asteroseismic surface effect is applied in the fitting process, as (1) the 1D–3D coupling approach greatly reduces the discrepancy between the theoretical and measured frequencies (cf. paper I fig. 5 and paper II fig. 10) and (2) the remaining surface terms have a different shape than those from traditional models. The median, 16th, and 84th quantiles of posterior distributions of stellar parameters inferred based on our model grid for KIC 9970396 are also listed in Table 3.

The échelle diagram depicts the relationship between mode frequency and  $\nu \bmod \Delta\nu$ , where  $\Delta\nu$  is the large frequency separation defined as the frequency difference between two modes with the

**Table 4.** Fundamental parameters of KIC 10001167 from observations, the best-fitting model and median values together with the 16 per cent and 84 per cent limit of the inferred posterior distributions.

Stellar parameter	$M/M_{\odot}$	$R/R_{\odot}$	$L/L_{\odot}$	$T_{\text{eff}}$ [K]	[Fe/H]	$[\alpha/\text{Fe}]$	$\nu_{\text{max}}$ [ $\mu\text{Hz}$ ]	Age ( $10^9$ yr)
Observation	$0.934 \pm 0.008$ (a)	$13.03 \pm 0.21$ (a)	$71.7 \pm 3.8$ (b)	$4625 \pm 59$ (a)	$-0.73 \pm 0.10$ (a)	$0.37 \pm 0.10$ (a)	$19.9 \pm 1.0$ (c)	–
Best model	0.998	13.03	69.2	4618	–0.82	0.4	20.4	7.57
Inferred values	$0.969^{+0.025}_{-0.030}$	$12.90^{+0.10}_{-0.15}$	$66.5^{+3.9}_{-3.6}$	$4603^{+54}_{-62}$	$-0.75^{+0.10}_{-0.10}$	–	$20.2^{+0.2}_{-0.2}$	$8.60^{+1.24}_{-0.79}$

Note. (a) Thomsen et al. (2025); (b) cf. note (b) in Table 3; (c) Li et al. (2024b).

**Figure 10.** Échelle diagram of the best-fitting model of KIC 10001167 (Table 4). Orange, green, and blue filled circles indicate measured frequencies of  $l = 0, 1, 2$  modes. No frequency correction is applied to the theoretical frequencies.

same degree and consecutive radial order. The diagram divides the frequency spectrum into sections so that p modes with the same  $l$  line up nearly vertically in the diagram whereas modes with different degrees are well separated, thus offering a clear manner to compare theoretical and observed frequencies. As seen in Fig. 9, theoretical frequencies predicted from our best model agree reasonably well with observation for both radial p modes and non-radial mixed modes. Systematic offsets are still present at high frequencies, which is expected as the 1D–3D coupling method is known to reduce but not eliminate the asteroseismic surface effect. Although the mixing-length parameter is not adjusted within our approach, we are able to model oscillation frequencies of KIC 9970396 at a similar level of fidelity as the detailed investigation by Li et al. (2018, see their fig. A6; after applying surface corrections they achieve better agreement at high frequencies). The stellar mass inferred from our modelling falls between the dynamical mass and the higher values determined from the (corrected) asteroseismic scaling relation (see Brogaard et al. 2018 fig. 7). Given that the global properties of our model are also compatible with measured values, we are confident that the 1D–3D coupling method produces reliable structural models for KIC 9970396 with realistic near-surface stratification, which is an encouraging validation of our method at non-solar metallicity.

### 4.3 KIC 10001167

KIC 10001167 is another eclipsing binary system observed by *Kepler*. Similar to KIC 9970396, it is composed of a G or F-type dwarf and a more evolved oscillating red giant (Gaulme et al. 2013). The system was thoroughly analysed in a recent study by Thomsen et al. (2025), who derived precise dynamical mass and radius for the oscillating giant. Its low metallicity,  $[\text{Fe}/\text{H}] = -0.73$  and a 0.37 dex enhancement of  $\alpha$ -element (Thomsen et al. 2025), makes KIC 10001167 an ideal target for validating the 1D–3D coupling method. The red giant was modelled in detail in Li et al. (2024b) as one of the binary samples to identify the relationship between metallicity and the mixing-length parameter.

We measure the oscillation frequencies of  $l = 0, 1, 2$  modes following the method described in Li et al. (2024a). In brief, we identify frequency bins with signal-to-noise ratios greater than five as prominent spikes, and then extract a mode where a number of spikes appear within the typical width of an oscillation mode. By visual inspection, we identify the  $l = 0, 2$  modes based on their characteristic pair-peak features, and treat the rest as the  $l = 1$  modes. We fit a mode with the Lorentz function to measure the frequency and estimate the uncertainty using the method described in Li et al. (2018), who applied the Monte Carlo approach and produced 1000 simulated power spectra by multiplying the power spectrum by a random noise following a  $\chi^2$  distribution with 2 degrees of freedom.

The standard deviation of the 1000 measured frequencies is used as the uncertainty of a mode.

A grid of evolutionary tracks with stellar masses ranging from 0.86 to  $1 M_{\odot}$  and initial [Fe/H] from  $-0.85$  to  $-0.6$  is computed to model KIC 10001167. The  $\alpha$ -enhancement used in our evolutionary grid calculation is 0.3 or 0.4 dex. The frequency calculation and fitting procedure are described in Section 4.2. Theoretical models are fitted to the dynamical radius, effective temperature (cf. Table 4), as well as measured frequencies of  $l = 0, 1, 2$  modes. Classical parameters inferred from our grid-based modelling and parameters of the model with the maximum likelihood are tabulated in Table 4, and Fig. 10 is the échelle diagram of KIC 10001167.

The 1D–3D coupling method is able to produce acceptable models for KIC 10001167, whose radius, effective temperature, and individual frequencies agree reasonably well with corresponding observations. The inferred stellar mass and mass of the best-fitting model are greater than that measured from the binary orbit, similar to the case of KIC 9970396. If we instead fit the coupled models to the measured mass, effective temperature and mode frequencies, the derived stellar radius turns out to be  $12.73 \pm 0.05 R_{\odot}$ , being smaller than the dynamical radius. Results from these two scenarios indicate that when both fundamental stellar parameters and oscillation frequencies are used as constraints, the 1D–3D coupling method tends to slightly overestimate surface gravity and mean density ( $\bar{\rho}$ ) along the RGB. This assertion is further supported by the fact that our best model for KIC 10001167 predicts slightly larger  $\Delta\nu$  ( $2.767$  versus  $2.734 \pm 0.033 \mu\text{Hz}$  as measured by Li et al. 2024b) and  $v_{\text{max}}$ . These two key asteroseismic observables are proportional to  $\sqrt{\bar{\rho}}$  and  $g/\sqrt{T_{\text{eff}}}$ , respectively, according to the asteroseismic scaling relations. The underlying reason for this offset is not understood.<sup>8</sup> One possible cause from the modelling aspect could be the remaining small surface effects from coupled models (e.g. Fig. 10 and fig. 5 of paper I) that were not accounted for in our fitting process, which might lead to systematics in derived stellar parameters. A detailed study of asteroseismic surface effects in red giants using coupled or patched models whose near-surface stratification is supplied from 3D simulations is needed to shed light on this issue.

## 5 DISCUSSION: DIFFERENT APPROACHES OF IMPROVING STELLAR EVOLUTIONARY MODELS WITH 3D SURFACE CONVECTION SIMULATIONS

As outlined in the introduction, considerable effort has been devoted to overcoming the outstanding problem in stellar interior modelling, where phenomenological theories of convection such as the MLT fail in the near-surface convective layers. To our knowledge, there are three well-established methods that utilize 3D model atmosphere grids to improve the treatment of near-surface layers in stellar evolution calculations. Apart from the 1D–3D coupling method developed and tested in paper I, paper II, and this work, Trampedach et al. (2014a, b) carefully extracted the  $T(\tau)$  relation from the Trampedach et al. (2013) grid and performed a calibration of  $\alpha_{\text{MLT}}$  based on the same 3D grid; Spada et al. (2018), Spada, Demarque & Kupka (2021), and Manchon et al. (2024) used the adiabatic entropy  $s_{\text{ad}}$  derived from 3D simulations to calibrate  $\alpha_{\text{MLT}}$  throughout the evolution calculation. In this section, we compare the three methods

by summarizing their reasoning and methodology, and pointing out their respective advantages and limitations.

The 1D–3D coupling method has two outstanding advantages. By replacing near-surface layers of the 1D stellar structure model with the mean 3D model atmosphere, the model-predicted theoretical frequencies agree better with asteroseismic observations. Placing the matching point in the near-adiabatic region also makes the evolution tracks insensitive to the choice of  $\alpha_{\text{MLT}}$ , effectively eliminating this important adjustable parameter. Another accompanying benefit is that consistent opacity between 1D and 3D models is not essential for this method, as the flux of radiative diffusion is negligible compared to the total energy flux at the matching location.

On the other hand, since the matching point is located in the hydrogen ionization zone ( $T_{\text{m}} > 10\,000\text{ K}$ ), where thermodynamics quantities are relatively more sensitive to the adopted EOS and the hydrogen mass fraction, using identical EOS between the 1D and 3D models is advantageous. It is worth noting that element diffusion alters hydrogen, helium, and metal mass fractions at the outermost shell of the 1D model, whereas the relationship between  $X$  and  $Z$  is fixed by the chemical composition employed in the 3D simulation. Although our metallicity interpolation ensures that  $Z/X$  is identical between the atmosphere and interior model, individual H, He, and metal mass fractions are not necessarily consistent between the two parts. This is a common problem when using tabulated model atmospheres as outer boundary conditions for stellar evolution. Nevertheless, we note that Spada et al. (2021) and Manchon et al. (2024) have taken into account the change of mean molecular weight in their entropy calibration, as discussed below. Finally, the 1D–3D coupling method requires not only the mean temperature and pressure but also partial derivatives at the matching point, making this method more susceptible to the interpolation scheme and the density of the 3D grid compared to other approaches.

Trampedach et al. (2014a, b) conceived and developed a different approach to improve stellar evolution calculation using information from 3D simulations. Being aware that convection is described in a fundamentally different way in 1D models and 3D hydrodynamics simulations, Trampedach et al. (2014a) extracted the radiative  $T(\tau)$  relation that excludes the contribution from convection from their 3D grid. The  $\alpha_{\text{MLT}}$  calibration is carried out using a stellar envelope code with identical microphysics. At each  $T_{\text{eff}}$  and  $\log g$ , the appropriate  $\alpha_{\text{MLT}}$  is determined by requiring the temperature of the envelope and the 3D model to agree at the matching point. The radiative  $T(\tau)$  relation provides atmosphere boundary conditions for stellar evolution calculations and meanwhile affects the temperature stratification near the photosphere, and should be applied in conjunction with the calibrated  $\alpha_{\text{MLT}}$ . Similar to standard stellar evolution, the outer boundary is located at the photosphere (but not  $\tau = 2/3$  as in the case of the grey atmosphere), above which atmosphere integration is performed based on the  $T(\tau)$  relation. This method relies on tabulated  $T(\tau)$  relations and  $\alpha_{\text{MLT}}$  values, making it less affected by the choice of interpolation method and less prone to numerical issues compared to the 1D–3D coupling approach. However, when utilizing the 3D-calibrated  $T(\tau)$  relation and  $\alpha_{\text{MLT}}$  in stellar evolution calculation, it is preferred to use identical EOS and opacity as the 3D grid. If the stellar evolution code adopts different microphysics, an additional scaling factor to the calibrated mixing-length parameter is necessary to compensate for this inconsistency (see Mosumgaard et al. 2018, equation 2). As investigated in Mosumgaard et al. (2018), implementing the 3D-calibrated  $T(\tau)$  relation and  $\alpha_{\text{MLT}}$  shifts the evolutionary track to slightly hotter  $T_{\text{eff}}$  in the RGB relative to the grey atmosphere. The improved structural model, however, has minimal effect on the predicted p-mode frequencies (Mosumgaard et al. 2018,

<sup>8</sup>We note that Thomsen et al. (2025) modelled KIC 10001167 with traditional evolutionary calculations and also predicted a larger surface gravity and mean density compared with observations (see their table G.2).

fig. 7), because the near-surface layers are still described by the MLT.

The entropy calibration method proposed and implemented by Tanner, Basu & Demarque (2016), Spada et al. (2018, 2021), and Manchon et al. (2024) is essentially an on-the-fly calibration of  $\alpha_{\text{MLT}}$  during stellar evolution. At every evolution time step, the mixing-length parameter is adjusted such that entropy at the bottom convective boundary equals the entropy of the inflow at the bottom boundary of the simulation (referred to as asymptotic entropy in Magic et al. 2015) at that  $T_{\text{eff}}$ ,  $\log g$ , and metallicity. The asymptotic entropy was expressed as a function of global stellar parameters for easy implementation. As  $\alpha_{\text{MLT}}$  is fixed by the entropy calibration, it is no longer a free parameter within this approach. Possible inconsistencies in chemical composition between the 1D and 3D models are mended by introducing an extra factor based on the mean molecular weight (Spada et al. 2021; Manchon et al. 2024). Moreover, Spada et al. (2018) and Manchon et al. (2024) included a constant correction factor to account for different entropy ‘zero points’ defined in EOSs. Manchon et al. (2024) showed that the entropy correction factor does not vary significantly across the parameter space, suggesting different EOSs between 1D and 3D models are acceptable within this approach provided that the correction factor is carefully determined. Finally, like the calibrated  $T(\tau)$  and  $\alpha_{\text{MLT}}$  method, the entropy calibration is anchored to surface parameters such as  $T_{\text{eff}}$  and the adiabatic entropy, hence will not lead to notable improvement on theoretical oscillation frequencies.

To sum up, the major advantage of the 1D–3D coupling approach compared to other methods is that near-surface layers are represented by the mean 3D model which is more realistic, making our method suitable for modelling stars with asteroseismic data. On the other hand, the calibrated  $T(\tau)$  relation and  $\alpha_{\text{MLT}}$  as well as the entropy calibration method, especially the latter, is less influenced by numerical challenges. It will be worthwhile to implement all three approaches to the same stellar evolution code for quantitative comparisons of evolutionary tracks and interior structures predicted by each method.

## 6 CONCLUSIONS

As a continuation of [paper I](#) and [paper II](#), we refined and further developed the method of coupling 1D stellar evolution with 3D model atmosphere on-the-fly by implementing the interpolation of mean 3D models across metallicities in GARSTEC. This effort allows us to utilize the majority of STAGGER-grid models and extend the capability of the 1D–3D coupling method to stars with non-solar metallicity. We demonstrate quantitatively that by placing the outer boundary of the stellar interior model well beneath the stellar surface in the near-adiabatic layer, stellar evolution tracks computed with the 1D–3D coupling method become far less sensitive to  $\alpha_{\text{MLT}}$  compared to the standard approach. The consequence is that  $\alpha_{\text{MLT}}$  can be regarded as a constant determined by solar calibration.

The 1D–3D coupling method was validated against stars in detached eclipsing binary systems, where accurate stellar mass and radius could be measured from binary orbits. It is clear from Section 4 that the method can be applied for realistic evolutionary modelling of stars with different metallicities. With  $\alpha_{\text{MLT}}$  fixed and few input parameters to adjust, our method is capable of generating stellar models that reproduce the basic properties for both components of AI Phe, whose fundamental parameters are extremely well constrained. Tests on oscillating *Kepler* giants in binary systems highlight the strength of the 1D–3D coupling method. For both giants, their radial and mixed mode frequencies are fitted by our coupled models at a satisfactory level, without the need for artificial surface corrections.

Although the surface term is still not fully eliminated, our method of coupling 1D and 3D models during stellar evolution demonstrates great potential for modelling stars with precise asteroseismic data. The best models of two *Kepler* giants have slightly higher masses than inferred from observations. As an overestimated mass leads to a younger stellar age, this systematic offset deserves further investigation in future studies.

Meanwhile, bends (or kinks) observed in the evolutionary tracks indicate room for improvement in the 1D–3D coupling method. An optimum numerical approach for interpolating mean 3D models and evaluating partial derivatives at the outer boundary is necessary. Of course, the reliability of the method and accuracy of the interpolation would benefit from increasing the density of the grid in the  $(T_{\text{eff}}, \log g)$  plane, and supplying missing 3D models at the edges of the grid (e.g.  $T_{\text{eff}} = 5000$  K,  $\log g = 2$ ) would allow metal-poor tracks to reach higher up the RGB. Nevertheless, preliminary investigations with a denser grid indicate that this would typically result in changes less than about 20 K in the effective temperature. Another open problem in metallicity interpolation is that the method is based on the mass fraction ratio  $Z/X$ . As a result, individual hydrogen, helium, and metal mass fractions may differ between the 1D and 3D models when element diffusion is taken into account. We will address these issues in future work to fully exploit the potential of the 1D–3D coupling method for modelling oscillating stars and determining the ages of star clusters.

The fact that our 1D–3D coupling models remove the need for  $\alpha_{\text{MLT}}$  in calculating thermal structure near the surface introduces new opportunities for stellar modelling. This is particularly relevant for red giants, where  $\alpha_{\text{MLT}}$  exhibits strong degeneracies with  $Y_{\text{init}}$ , and as a result affecting stellar age estimates by up to 30 per cent. This new approach could lead to more accurate stellar age estimates as well as potentially tighter constraints on helium abundance – both are crucial for understanding the history of Galactic chemical enrichment. Additionally, since stellar fundamental parameters such as radius and  $T_{\text{eff}}$  are insensitive to  $\alpha_{\text{MLT}}$ , the frequency of maximum oscillation power estimated from the  $\nu_{\text{max}} \propto g/\sqrt{T_{\text{eff}}}$  scaling relation is less affected by modelling uncertainties, permitting more meaningful comparisons against observed  $\nu_{\text{max}}$ . This provides a means to test the validity of the  $\nu_{\text{max}}$  scaling relation in red giants. Such a test would be very valuable for metal-poor stars, where the  $\nu_{\text{max}}$  scaling relation remains questionable to use but is highly needed.

## ACKNOWLEDGEMENTS

The authors are grateful to Jeppe Thomsen for providing observational data of KIC 10001167 and to Anish Amarsi and Mia Lundkvist for reading and commenting on this manuscript. We thank also Meridith Joyce, Victor Aguirre Børsen-Koch, Regner Trampedach, and Teresa Braun for valuable comments and fruitful discussions. YZ thanks the hospitality of Max Planck Institute for Astrophysics during his visit. YZ gratefully acknowledges support from the Elaine P. Snowden Fellowship. ACSJ is supported by the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship, a Schmidt Sciences programme. CL acknowledges funds from the UKRI Future Leaders Fellowship (MR/S035214/1). CL also acknowledged funds from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 852977). OK acknowledges support by the Swedish Research Council (grant agreement 2023–03667) and the Swedish National Space Agency. This work was supported by a research grant (42101) from VILLUM FONDEN. Funding for the Stellar Astrophysics Centre was provided by The Danish National Research Foundation

(grant DNR106). The numerical results presented in this work were partly obtained at the Centre for Scientific Computing, Aarhus <https://phys.au.dk/forskning/faciliteter/cscaa/>. This research was supported by computational resources provided by the Australian Government through the National Computational Infrastructure (NCI) under the National Computational Merit Allocation Scheme and the ANU Merit Allocation Scheme (project y89).

## DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

- Aerts C., Christensen-Dalsgaard J., Kurtz D. W., 2010, *Asteroseismology*. Springer, Dordrecht
- Aguirre Børsen-Koch V. et al., 2022, *MNRAS*, 509, 4344
- Ahlborn F., Kupka F., Weiss A., Flaskamp M., 2022, *A&A*, 667, A97
- Amarsi A. M., Nissen P. E., Skúladóttir Á., 2019, *A&A*, 630, A104
- Andersen J., Clausen J. V., Nordstrom B., Gustafsson B., Vandenberg D. A., 1988, *A&A*, 196, 128
- Asplund M., Nordlund Å., Trampedach R., Allende Prieto C., Stein R. F., 2000, *A&A*, 359, 729
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
- Bahcall J. N., Ulrich R. K., 1988, *Rev. Mod. Phys.*, 60, 297
- Bailer-Jones C. A. L., Rybizki J., Fousneau M., Demleitner M., Andrae R., 2021, *AJ*, 161, 147
- Basu S., Antia H. M., 2004, *ApJ*, 606, L85
- Berger T. A., Huber D., van Saders J. L., Gaidos E., Tayar J., Kraus A. L., 2020, *AJ*, 159, 280
- Böhm-Vitense E., 1958, *Z. Astrophys.*, 46, 108
- Borucki W. J. et al., 2010, *Science*, 327, 977
- Braun T. A. M., Ahlborn F., Weiss A., 2024, *A&A*, 689, A292
- Brogaard K. et al., 2016, *Astron. Nachr.*, 337, 793
- Brogaard K. et al., 2018, *MNRAS*, 476, 3729
- Brown T. M., Gilliland R. L., Noyes R. W., Ramsey L. W., 1991, *ApJ*, 368, 599
- Cassisi S., Schlattl H., Salaris M., Weiss A., 2003, *ApJ*, 582, L43
- Chaplin W. J. et al., 2020, *Nat. Astron.*, 4, 382
- Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, *ApJ*, 823, 102
- Choi J., Dotter A., Conroy C., Ting Y.-S., 2018, *ApJ*, 860, 131
- Christensen-Dalsgaard J., 2008, *Ap&SS*, 316, 113
- Christensen-Dalsgaard J., Proffitt C. R., Thompson M. J., 1993, *ApJ*, 403, L75
- Christensen-Dalsgaard J. et al., 1996, *Science*, 272, 1286
- Collet R., Asplund M., Trampedach R., 2007, *A&A*, 469, 687
- Collet R., Nordlund Å., Asplund M., Hayek W., Trampedach R., 2018, *MNRAS*, 475, 3369
- Creevey O. L. et al., 2015, *A&A*, 575, A26
- Creevey O. L., Cassisi S., Thévenin F., Salaris M., Pietrinferni A., 2024, *A&A*, 689, A243
- Cutri R. M. et al., 2003, 2MASS All Sky Catalog of point sources. [Webpage](http://www.ipac.caltech.edu/2mass/)
- Deal M., Goupil M. J., Marques J. P., Reese D. R., Lebreton Y., 2020, *A&A*, 633, A23
- Dotter A., Conroy C., Cargile P., Asplund M., 2017, *ApJ*, 840, 99
- El-Badry K., Rix H.-W., Heintz T. M., 2021, *MNRAS*, 506, 2269
- Ferguson J. W., Alexander D. R., Allard F., Barman T., Bodnarik J. G., Hauschildt P. H., Heffner-Wong A., Tamanai A., 2005, *ApJ*, 623, 585
- Freytag B., Ludwig H. G., Steffen M., 1996, *A&A*, 313, 497
- Gaulme P., McKeever J., Rawls M. L., Jackiewicz J., Mosser B., Guzik J. A., 2013, *ApJ*, 767, 82
- Gaulme P. et al., 2016, *ApJ*, 832, 121
- Gough D. O., 1977, *ApJ*, 214, 196
- Green G. M., Schlafly E., Zucker C., Speagle J. S., Finkbeiner D., 2019, *ApJ*, 887, 93
- Groenewegen M. A. T., 2021, *A&A*, 654, A20
- Gustafsson B., Edvardsson B., Eriksson K., Jørgensen U. G., Nordlund Å., Plez B., 2008, *A&A*, 486, 951
- Hayek W., Asplund M., Carlsson M., Trampedach R., Collet R., Gudiksen B. V., Hansteen V. H., Leenaarts J., 2010, *A&A*, 517, A49
- Hekker S., Christensen-Dalsgaard J., 2017, *A&AR*, 25, 1
- Heney L., Vardya M. S., Bodenheimer P., 1965, *ApJ*, 142, 841
- Herwig F., 2005, *ARA&A*, 43, 435
- Higl J., Weiss A., 2017, *A&A*, 608, A62
- Huber D. et al., 2017, *ApJ*, 844, 102
- Huber D. et al., 2024, *ApJ*, 975, 19
- Iglesias C. A., Rogers F. J., 1996, *ApJ*, 464, 943
- Irwin A. W., 2004, unpublished, available at [http://freeeos.sourceforge.net/ef\\_fit.pdf](http://freeeos.sourceforge.net/ef_fit.pdf)
- Irwin A. W., 2012, Astrophysics Source Code Library, record ascl:1211.002
- Johnston C., Tkachenko A., Aerts C., Molenberghs G., Bowman D. M., Pedersen M. G., Buysschaert B., Pápics P. I., 2019, *MNRAS*, 482, 1231
- Jørgensen A. C. S., Weiss A., 2019, *MNRAS*, 488, 3463
- Jørgensen A. C. S., Weiss A., Mosumgaard J. R., Silva Aguirre V., Sahlholdt C. L., 2017, *MNRAS*, 472, 3264
- Jørgensen A. C. S., Mosumgaard J. R., Weiss A., Silva Aguirre V., Christensen-Dalsgaard J., 2018, *MNRAS*, 481, L35 (paper I)
- Jørgensen A. C. S., Weiss A., Angelou G., Silva Aguirre V., 2019, *MNRAS*, 484, 5551
- Jørgensen A. C. S. et al., 2021, *MNRAS*, 500, 4277
- Joyce M., Chaboyer B., 2018, *ApJ*, 856, 10
- Joyce M., Tayar J., 2023, *Galaxies*, 11, 75
- Käpylä P. J., Rheinhardt M., Brandenburg A., Arlt R., Käpylä M. J., Lagg A., Olsperg N., Warnecke J., 2017, *ApJ*, 845, L23
- Kippenhahn R., Weigert A., Weiss A., 2012, *Stellar Structure and Evolution*, Astronomy and Astrophysics Library, Springer, Berlin, Heidelberg
- Kjeldsen H., Bedding T. R., 1995, *A&A*, 293, 87
- Krishna Swamy K. S., 1966, *ApJ*, 145, 174 (KS)
- Kupka F., Muthsam H. J., 2017, *Living Rev. Comput. Astrophys.*, 3, 1
- Kupka F., Ahlborn F., Weiss A., 2022, *A&A*, 667, A96
- Larsen J. R., Christensen-Dalsgaard J., Lundkvist M. S., Rørsted J. L., Winther M. L., Kjeldsen H., 2024, *A&A*, 690, A394
- Li Y., 2012, *ApJ*, 756, 37
- Li T., Bedding T. R., Huber D., Ball W. H., Stello D., Murphy S. J., Bland-Hawthorn J., 2018, *MNRAS*, 475, 981
- Li T., Bi S., Davies G. R., Bedding T. R., Li Y., Stello D., Reyes C., 2024a, *MNRAS*, 530, 2810
- Li Y. et al., 2024b, *ApJ*, 974, 77
- Ludwig H.-G., Freytag B., Steffen M., 1999, *A&A*, 346, 111
- Ludwig H. G., Caffau E., Steffen M., Freytag B., Bonifacio P., Kučinskas A., 2009, *Mem. Soc. Astron. Italiana*, 80, 711
- Magic Z., Collet R., Asplund M., Trampedach R., Hayek W., Chiavassa A., Stein R. F., Nordlund Å., 2013a, *A&A*, 557, A26
- Magic Z., Collet R., Hayek W., Asplund M., 2013b, *A&A*, 560, A8
- Magic Z., Weiss A., Asplund M., 2015, *A&A*, 573, A89
- Manchon L. et al., 2024, *A&A*, 687, A146
- Martocchia S. et al., 2018, *MNRAS*, 473, 2688
- Maxted P. F. L. et al., 2020, *MNRAS*, 498, 332
- Meléndez J. et al., 2014, *ApJ*, 791, 14
- Mihalas D., Dappen W., Hummer D. G., 1988, *ApJ*, 331, 815
- Miller N. J., Maxted P. F. L., Smalley B., 2020, *MNRAS*, 497, 2899
- Mosumgaard J. R., Ball W. H., Silva Aguirre V., Weiss A., Christensen-Dalsgaard J., 2018, *MNRAS*, 478, 5650
- Mosumgaard J. R., Jørgensen A. C. S., Weiss A., Silva Aguirre V., Christensen-Dalsgaard J., 2020, *MNRAS*, 491, 1160 (paper II)
- Nordlander T., Korn A. J., Richard O., Lind K., 2012, *ApJ*, 753, 48
- Nordlund Å., Galsgaard K., 1995, Tech. rep., A 3D MHD code for Parallel Computers, Astronomical Observatory, Copenhagen University
- Nordlund Å., Stein R. F., Asplund M., 2009, *Living Rev. Sol. Phys.*, 6, 2
- Paxton B., Bildsten L., Dotter A., Herwig F., Lesaffre P., Timmes F., 2011, *ApJS*, 192, 3

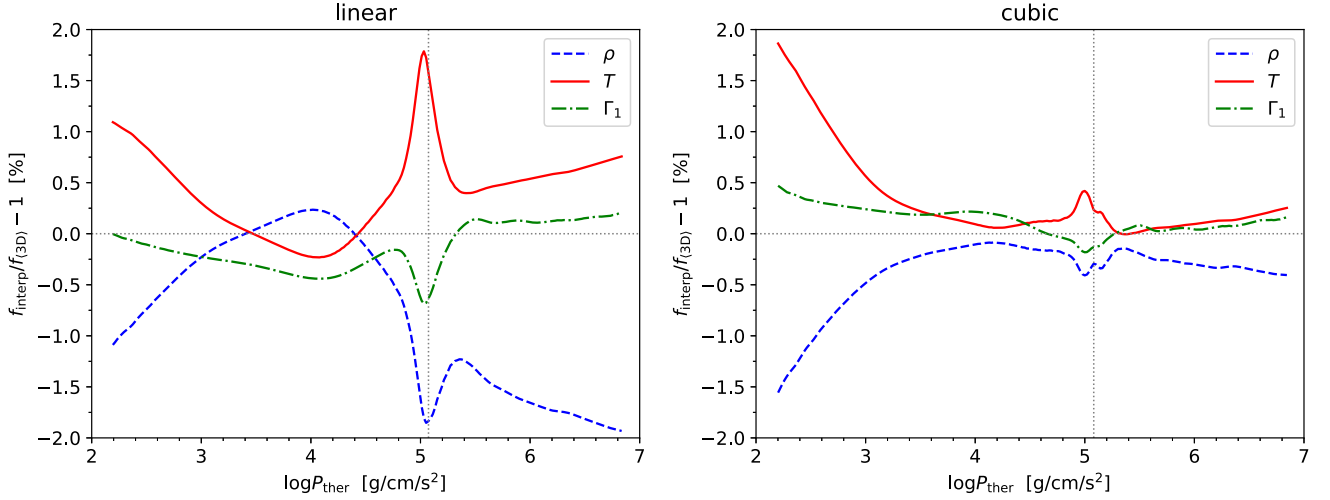
- Pereira T. M. D., Asplund M., Collet R., Thaler I., Trampedach R., Leenaarts J., 2013, *A&A*, 554, A118
- Pinsonneault M., 1997, *ARA&A*, 35, 557
- Pinsonneault M. H. et al., 2018, *ApJS*, 239, 32
- Plez B., 2008, *Phys. Scr. T*, 133, 014003
- Proffitt C. R., Michaud G., 1991, *ApJ*, 371, 584
- Prša A. et al., 2016, *AJ*, 152, 41
- Prša A. et al., 2011, *AJ*, 141, 83
- Reipurth B., 1978, *Inf. Bull. Var. Stars*, 1419, 1
- Remple B. A., Battich T., Weiss A., 2024, *A&A*, 687, A260
- Renka R. J., Cline A. K., 1984, *The Rocky Mountain J. Math.*, 14, 223
- Richard O., Michaud G., Richer J., 2005, *ApJ*, 619, 538
- Richer J., Michaud G., Rogers F., Iglesias C., Turcotte S., LeBlanc F., 1998, *ApJ*, 492, 833
- Rodríguez Díaz L. F., Bigot L., Aguirre Børsen-Koch V., Lund M. N., Rørsted J. L., Kallinger T., Sulis S., Mary D., 2022, *MNRAS*, 514, 1741
- Rodríguez Díaz L. F. et al., 2024, *A&A*, 688, A212
- Rogers F. J., Nayfonov A., 2002, *ApJ*, 576, 1064
- Saffe C. et al., 2024, *A&A*, 682, L23
- Salaris M., Cassisi S., Schiavon R. P., Pietrinferni A., 2018, *A&A*, 612, A68
- Schmid V. S., Aerts C., 2016, *A&A*, 592, A116
- Silva Aguirre V. et al., 2015, *MNRAS*, 452, 2127
- Silva Aguirre V. et al., 2018, *MNRAS*, 475, 5487
- Spada F., Demarque P., Kim Y. C., Sills A., 2013, *ApJ*, 776, 87
- Spada F., Demarque P., Basu S., Tanner J. D., 2018, *ApJ*, 869, 135
- Spada F., Demarque P., Kupka F., 2021, *MNRAS*, 504, 3128
- Steffen M., 1990, *A&A*, 239, 443
- Stein R. F., Nordlund Å., Collet R., Trampedach R., 2024, *ApJ*, 970, 24
- Stokholm A., Nissen P. E., Silva Aguirre V., White T. R., Lund M. N., Mosumgaard J. R., Huber D., Jessen-Hansen J., 2019, *MNRAS*, 489, 928
- Tanner J. D., Basu S., Demarque P., 2016, *ApJ*, 822, L17
- Tayar J. et al., 2017, *ApJ*, 840, 17
- Thomsen J. S. et al., 2025, preprint ([arXiv:2504.17853](https://arxiv.org/abs/2504.17853)),
- Thoul A. A., Bahcall J. N., Loeb,
- Trampedach R., Stein R. F., 2011, *ApJ*, 731, 78
- Trampedach R., Asplund M., Collet R., Nordlund Å., Stein R. F., 2013, *ApJ*, 769, 18
- Trampedach R., Stein R. F., Christensen-Dalsgaard J., Nordlund Å., Asplund M., 2014a, *MNRAS*, 442, 805
- Trampedach R., Stein R. F., Christensen-Dalsgaard J., Nordlund Å., Asplund M., 2014b, *MNRAS*, 445, 4366
- Valle G., Dell’Omodarme M., Prada Moroni P. G., Degl’Innocenti S., 2023, *A&A*, 678, A203
- Varenne O., Monier R., 1999, *A&A*, 351, 247
- Verma K., Silva Aguirre V., 2019, *MNRAS*, 489, 1850
- Verma K., Raodeo K., Basu S., Silva Aguirre V., Mazumdar A., Mosumgaard J. R., Lund M. N., Ranadive P., 2019, *MNRAS*, 483, 4678
- Vernazza J. E., Avrett E. H., Loeser R., 1981, *ApJS*, 45, 635
- Viani L. S., Basu S., Ong J. M. J., Bonaca A., Chaplin W. J., 2018, *ApJ*, 858, 28
- Vinyoles N. et al., 2017, *ApJ*, 835, 202
- Wang E. X. et al., 2024, *MNRAS*, 528, 5394
- Weiss A., Schlattl H., 2008, *Ap&SS*, 316, 99
- Xiong D.-r., 2021, *Front. Astron. Space Sci.*, 7, 95
- Zhang X., Cai T., Li Y., Wu T., 2022, *ApJ*, 931, 64
- Zhou Y., Asplund M., Collet R., Joyce M., 2020, *MNRAS*, 495, 4904
- Zhou Y., Amarsi A. M., Aguirre Børsen-Koch V., Karlsmosse K. G., Collet R., Nordlander T., 2023, *A&A*, 677, A98
- Zhou Y., Christensen-Dalsgaard J., Asplund M., Li Y., Trampedach R., Ting Y.-S., Rørsted J. L., 2024, *ApJ*, 962, 118

## APPENDIX A: INTERPOLATION ERRORS FOR MEAN 3D MODELS

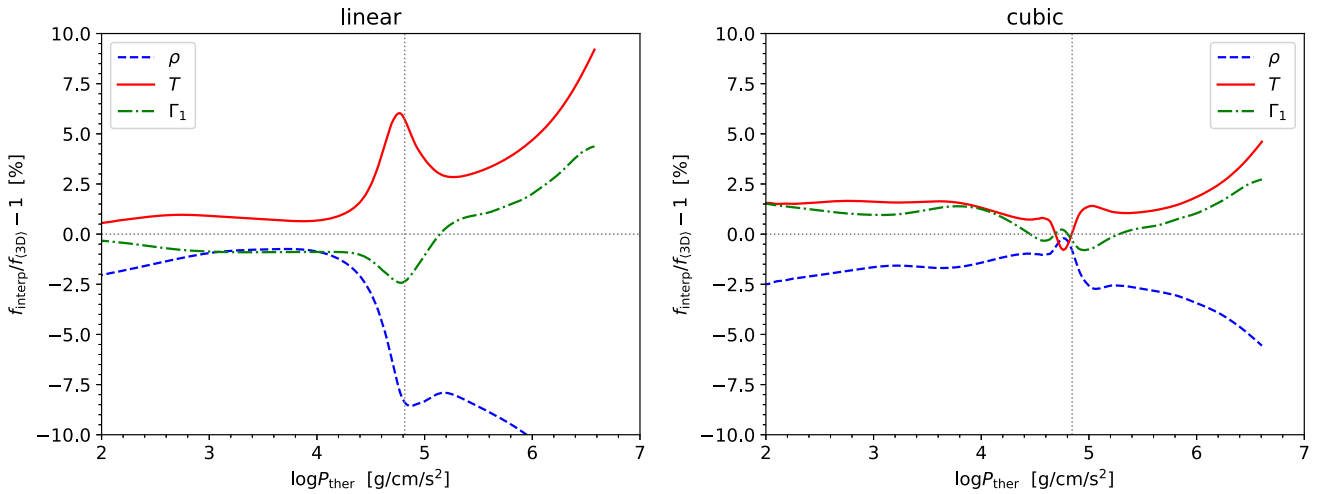
We quantify the errors of our interpolation scheme introduced in Section 2.2 across the HR diagram by taking one (3D) model out from the grid before constructing an interpolated mean stratification at the corresponding surface temperature and gravity based on the remaining models, or by comparing our interpolation results with an additional 3D model not included in the grid. As shown in Table A1, multiple tests are conducted for the solar parameter, models correspond to warm (F-type) and cool (K-type) dwarfs as well as red giants. Given the 3D grid adopted in this work, the optimal interpolation method (linear or cubic) depends on the stellar parameters, see Fig. A1–A6.

**Table A1.** Effective temperature and surface gravity of (3D) models used in the interpolation test. Except for the solar model, model names are constructed from the effective temperature (labelled as ‘ $t$ ’), surface gravity (‘ $g$ ’), and ‘m00’ stands for solar metallicity. Effective temperatures listed here are the mean value averaged from all simulation snapshots. We note that models  $t62g43m00$  and  $t47g27m00$  are not included in the STAGGER-grid (Fig. 1).

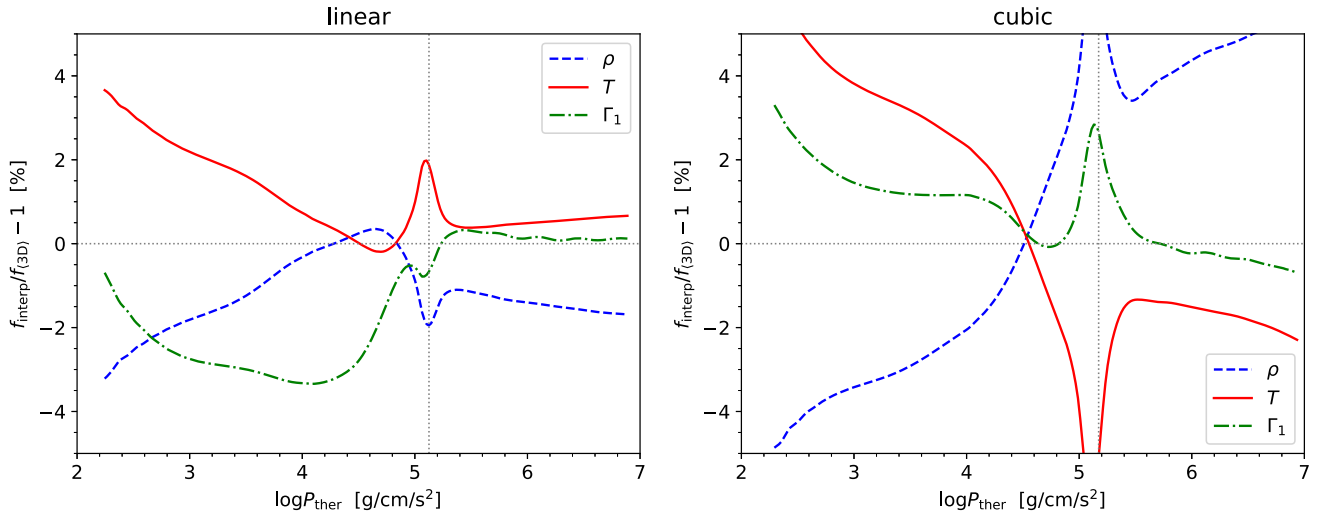
Model name	solar	$t62g43m00$	$t52g42m00$	$t47g47m00$	$t47g32m00$	$t47g27m00$
$T_{\text{eff}}$ (K)	5776	6231	5242	4743	4766	4695
$\log g$ (cgs)	4.44	4.319	4.25	4.75	3.25	2.67



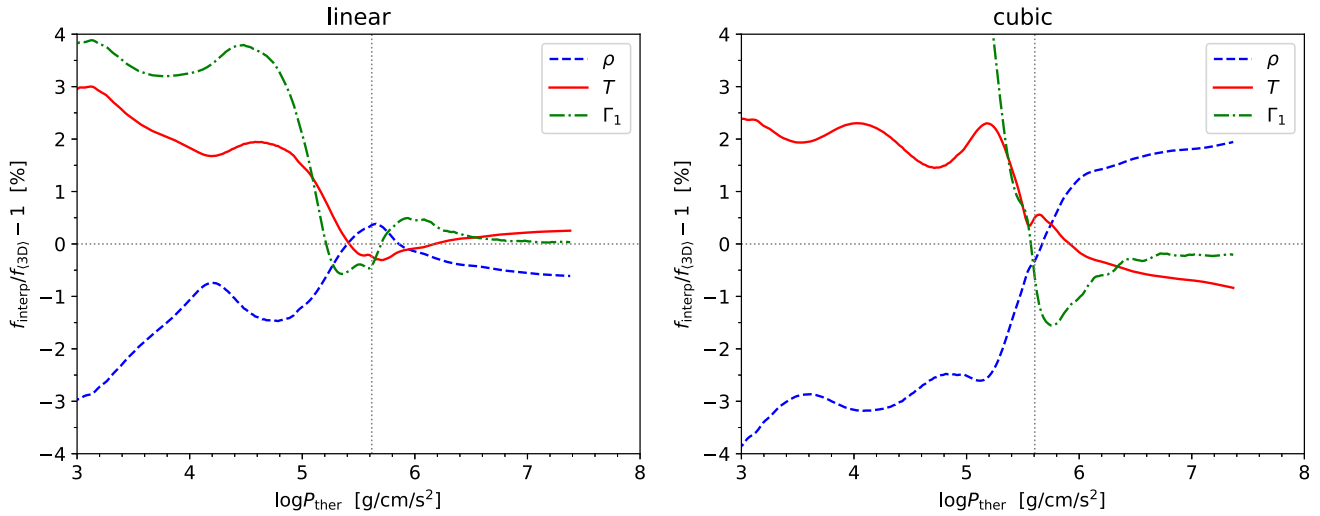
**Figure A1.** Interpolation error of density, temperature, and the first adiabatic exponent  $\Gamma_1$  for the mean 3D solar model. Error from linear (cubic) interpolation at different thermal pressure is shown in the left (right) panel. The vertical dotted line marks the location of density inflection defined as the local minimum of  $\partial \ln \rho / \partial \ln P_{\text{ther}}$ . Below the density inflection, the error in temperature ranges from about 0.5 per cent to over 0.7 per cent for linear interpolation and less than 0.25 per cent for cubic interpolation.



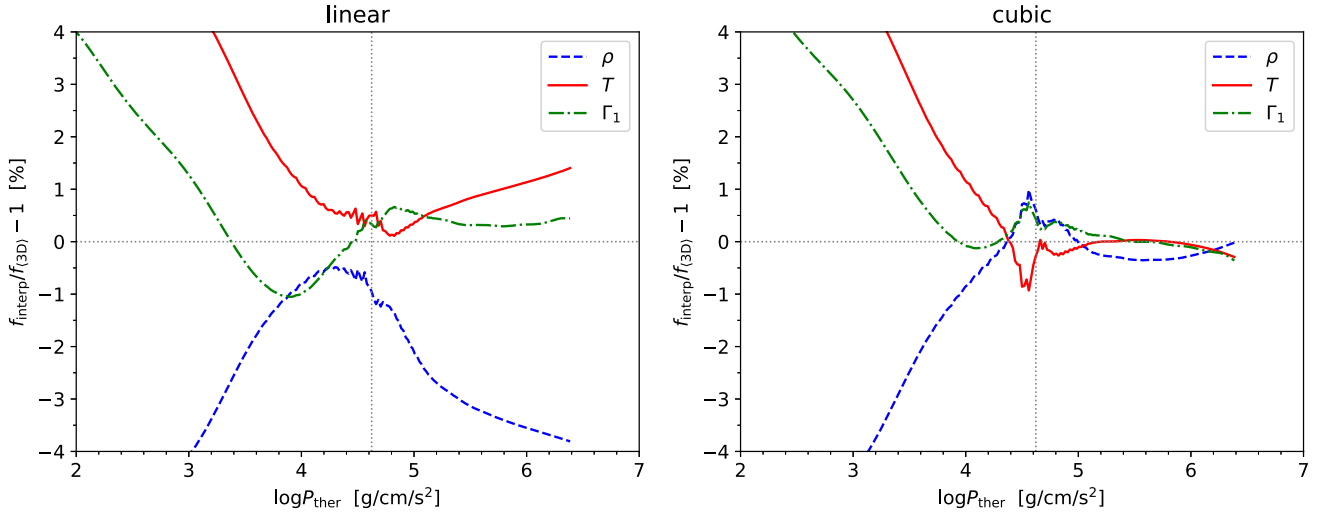
**Figure A2.** Interpolation error for the mean structure of model  $t62g43m00$  (see Table A1). The error of linear interpolation is larger than the cubic method around and below the density inflection. In deeper layers, relative error from linear interpolation reaches up to 9 per cent for temperature, while the cubic method results in less than 5 per cent error.



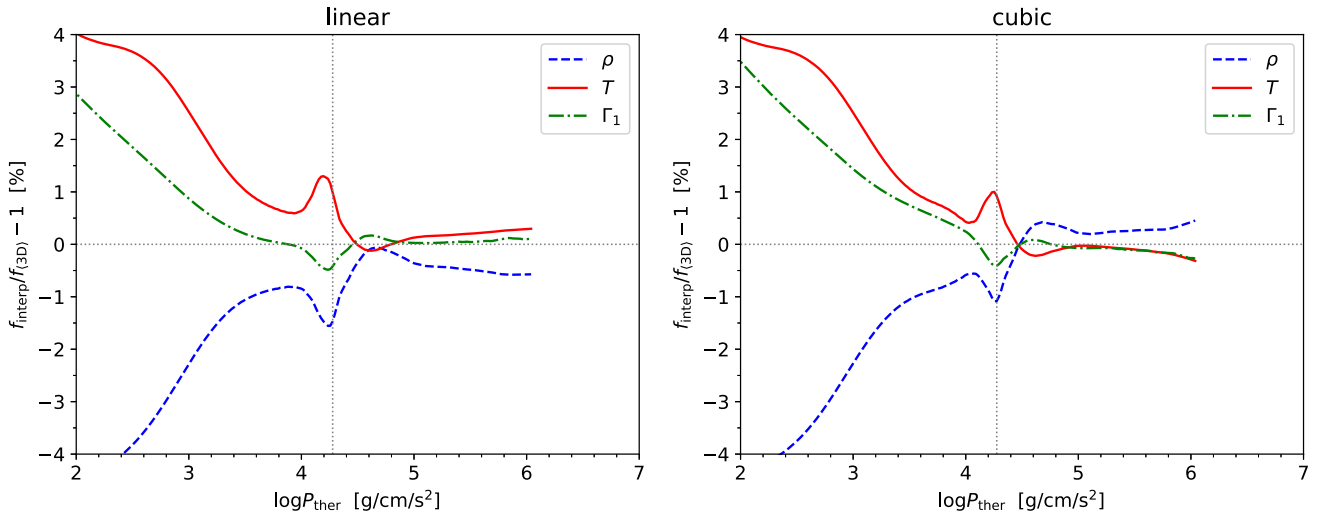
**Figure A3.** Interpolation error for the mean structure of model  $t52g42m00$ . Below the density inflection, the error of cubic interpolation is larger than the linear method: temperature profile of the (3D) model differs from the linear interpolated values by less than 0.75 per cent. The relative differences increase to nearly 2 per cent for the cubic method.



**Figure A4.** Interpolation error for the mean structure of model  $t47g47m00$ .



**Figure A5.** Interpolation error for the mean structure of model  $t47g32m00$ . Below the density inflection, the error in temperature ranges from less than 0.25 per cent to about 1.5 per cent for linear interpolation and less than 0.25 per cent for the cubic method.



**Figure A6.** Interpolation error for the mean structure of model  $t47g27m00$ .

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.