Galactic Archeology through advanced statistical methods and 3D non-local thermodynamic equilibrium radiative transfer

Ella Xi Wang

Supervisory panel

Prof. Helmut Jerjen, Dr. Thomas Nordlander
Dr. Yuan-Sen Ting, Prof. Karin Lind

PhD Thesis Proposal

The Australian National University
Research School of Astronomy & Astrophysics

March 22, 2021

Contents

Chapter 1

Background

1.1 Introduction

Galactic archaeology is the study of elemental abundances and stellar dynamics of stars within the Milky Way to learn about the formation and evolutionary history of our Galaxy (Freeman & Bland-Hawthorn, 2002). The commonly accepted paradigm of galaxy formation is the Λ Cold Dark Matter (ΛCDM) model, which hypothesises that larger, younger galaxies form through accretion of smaller, older galaxies. The disk stars accumulated from these accretion events should leave signatures in their stellar dynamics. However, it is unlikely that stellar dynamics remain unchanged from accretion till present day, due to e.g radial migration (Francois & Matteucci, 1993; Sellwood & Binney, 2002). Therefore, we also need to look at other signatures of stellar history still present in the star today. Stars with similar elemental abundances and ages likely co-evolved (Bland-Hawthorn et al., 2010), though even stars formed from the same protocloud have some inherent distribution in their elemental abundance due to inhomogeneities within the cloud (Armillotta et al., 2018). Through measuring both the stellar dynamics and elemental abundances of a large sample of stars, we expect to learn more about the formation history of the Galaxy.

Detailed analysis of stellar spectra offers insights in the properties of stars. In particular, spectra can tell us the stellar dynamics, stellar parameters (effective temperature, $T_{\rm eff}$, surface gravity, $\log g$, metallicity, [Fe/H]), and elemental abundances within a star. All together, the combination of stellar parameters and abundances is referred to in the context of machine learning as stellar labels. Knowing the stellar labels for a representative sample of stars can give us insights into Galactic archaeology. For example, the elemental abundances in old metal-poor stars tells us of the production of elements from first stars, the stellar labels of stars also allow us to identify stars that formed together but drifted apart over time. Ideally, we'd like to derive the stellar labels of a large sample of stars based on analysing a large set of spectral data employing accurate and fast techniques. However, in reality, techniques used to analyse stellar spectra are often inaccurate and/or

slow. The analysis of observed spectra usually relies on the generation and fitting of synthetic spectra. To be able to generate synthetic spectra for a multi-dimensional grid of stellar labels, whilst quantifying the systematic errors in interpolation will be a huge improvement to the precision and accuracy of the measured stellar labels. In addition to large spectral data sets, we also have complementary photometric data sets of the same target stars. Exploiting the information present in both spectroscopic and photometric data will further improve accuracy of derived stellar labels and help break partial degeneracies between individual stellar labels particularly for spectra with limited coverage or low signal-to-noise (SNR).

1.2 Stellar Spectroscopy

Stellar parameters, elemental abundances, and stellar dynamics are properties of stars which offer insight to Galactic archaeology. For example, stars with similar stellar dynamics and comparable elemental abundances likely formed in the same protocloud. The process of determining the stellar labels is typically done through computing a reference database of synthetic stellar spectra and comparing with observed spectra.

Spectral synthesis is done through solving the equation of radiative transfer in a model stellar atmosphere with atomic line parameters (Amarsi et al., 2016a). Atomic line parameters specify the line profile and include: ionisation state, excitation energy, $\log(gf)$, and broadening parameters. The ionisation state and excitation energy gives the populations, whilst the $\log(gf)$ gives the likelihood that a photon is absorbed.

Most of the radiation from stars originate in the photosphere, with only a small fraction originating in the hot but tenuous chromosphere and corona, both of which are located above the photosphere. Therefore, stellar models used for detailed radiative transfer only model the temperature, gas pressure, electron pressure, and density of the photosphere; and are often referred to as stellar atmosphere models. Traditionally, these stellar atmospheres have been 1D hydrostatic model atmospheres, with mixing-length theory used to estimate convection (Böhm-Vitense, 1958). 3D hydrodynamic atmospheres modelling convection from first principles without empirical calibrations are available for decades (Nordlund, 1982; Nordlund & Dravins, 1990), but only recently has it become possible to compute large grids (Magic et al., 2013). This physical description of convection results in not only horizontal inhomogeneities, but also a temperature structure that on average differs from the 1D models, and velocity fields that introduce local Doppler shifts. These shifts are approximated through microturbulence in 1D models, but are actually due to macroscopic convective motions rather than subscale turbulent flows (Asplund et al., 2000; Pereira et al., 2013).

The equation of radiative transfer in 1D is given by:

$$\mu \frac{\mathrm{dI}_\nu}{\mathrm{d}\tau_\nu} = I_\nu - S_\nu, \tag{1.1}$$

where I_ν is the intensity, S_ν is the source function, μ is the viewing angle (note $\mu = \cos(\theta)$, where θ is the angle of the line of sight relative to the vertical), and τ_ν is the optical depth given by:

$$\tau_\nu = \int \frac{\alpha_\nu}{|\mu|} \mathrm{dz}, \tag{1.2}$$

where α_ν is the volume extinction coefficient, and z is the vertical depth in the stellar atmosphere.

Solving the radiative transfer equation to find the emergent flux can be done under the assumption of local thermodynamic equilibrium (LTE) or relaxing that assumption, non-LTE (NLTE). Assuming LTE, meaning that collisional transitions dominate over radiation transitions, the Boltzmann and Saha distributions hold locally and can be used to compute the level populations. In addition, in LTE, the source function is assumed to be described by the Planck function, and the opacity is then given by the number of particles times the cross-section per particle. In NLTE, the level populations are given by the statistical equilibrium:

$$\sum_{l \neq i} (n_l P_{li} - n_i P_{il}) = 0, \tag{1.3}$$

where n_i is the number of particles in the ith state, P_{ij} is the probability of a transition from the ith state to the jth state, and is defined as $P_{ij} = R_{ij} + C_{ij}$, where C is the collisional probability and R is the radiative probability, further broken down into $R_{ij} = A_{ij} + B_{ij}\bar{J}_\nu$, with Einstein coefficients A for radiative emission and B for radiative absorption, and mean intensity J_ν. In other words, the level populations depend on all possible collisional and radiative transitions. In NLTE, S_ν^l, the source function for a line, l, is approximated as:

$$S_\nu^l \approx \frac{b_u}{b_l} B_\nu, \tag{1.4}$$

where B_ν is the Planck function, and b_l is the departure coefficient for the lower state, given by:

$$b_i = \frac{n_i^{\mathrm{NLTE}}}{n_i^{\mathrm{LTE}}}. \tag{1.5}$$

S_ν^l depends on the intensity field, so to model the NLTE line strength of any transition, we need the correct lower and upper level populations. When radiation deviates from the (local) Planck function at wavelengths where strong radiative transitions are located, these radiative transitions will break the collisional equilibrium, which results in the level populations deviating from the LTE state ($b_i \neq 1$). Therefore, NLTE better models reality when there is an intense radiation field.

NLTE radiative transfer solved in 3D stellar atmospheres is more realistic, because in 3D atmospheres there are horizontal inhomogeneities present, such that a given point in the photosphere is illuminated by non-local radiation from many different directions. However, 3D NLTE radiative transfer is also more computationally intensive due to the

larger number of points in the model and additional horizontal and diagonal rays which need to be propagated.

1.3 Neural Networks

Computing the synthetic spectra for one set of stellar labels can be expensive, especially under 3D NLTE. With millions of observable stars in the Milky Way, computing the synthetic spectra for each star becomes an impossible task with the current computational power available. Therefore, to analyse spectral data sets, we often compute a grid of synthetic spectra for some stellar labels then interpolate to arbitrary stellar labels. This interpolation can be done through multiple methods, traditionally polynomials are used (Ness et al., 2015; Mott et al., 2020), but with advancements in computational techniques, neural networks have been used to interpolate grids of synthetic spectra (Ting et al., 2019).

The first neural networks were known as perceptrons and were inspired by information storage in the brain (Rosenblatt, 1958). The natural extension to perceptrons are fully connected feedforward neural networks (FFNN), consisting of one input layer, one or more hidden layers, and one output layer, which are all fully connected, see Fig. 1.1 for a schematic. FFNNs have seen a wide variety of use on data sets, ranging from imaging, computer vision, predictions on time series data. Let x_i be the ith input, then to propagate this value through a hidden layer, we use:

$$h_j^m = f_{act}\left(\sum_i^N w_{ij}x_i + b_i\right) \tag{1.6}$$

where h_j^m is the value of the propagated input in the mth hidden layer, f_{act} is the activation function needed for non-linearity. Common activation functions include: TanH (hyperbolic tangent function), ReLU, sigmoid, amongst others; these activation functions chosen based upon their properties, e.g: gradient, bounded output, etc. w_{ij} is the weight associated with the ith value in the input layer and the jth value in the hidden layer, and b_i is the bias value which acts as a constant offset. This propagation is repeated for all inputs and through all hidden layers to get the final output (Goodfellow et al., 2016). Many extensions for FFNNs have been developed, e.g: convolutional neural networks, recurrent neural networks, and bayesian neural networks.

FFNNs have been used to interpolate synthetic spectra with higher accuracy than traditional methods (Ting et al., 2019), however, are still not perfect yet. During the training phase, the weights are often initialised from a normal distribution centered at 0 to help with gradient decent (Goodfellow et al., 2016). Due to the randomness in initiating weights for FFNNs and that there are no optimal convergence guarantees, the interpolated spectrum has a noise-like signal in it, see the purple fit in Fig. 1.2. Such noise is problematic for synthetic spectra used to fit observed spectra. This fitting process often involves interpolating the synthetic spectra to the same wavelength points as the observed spectra,

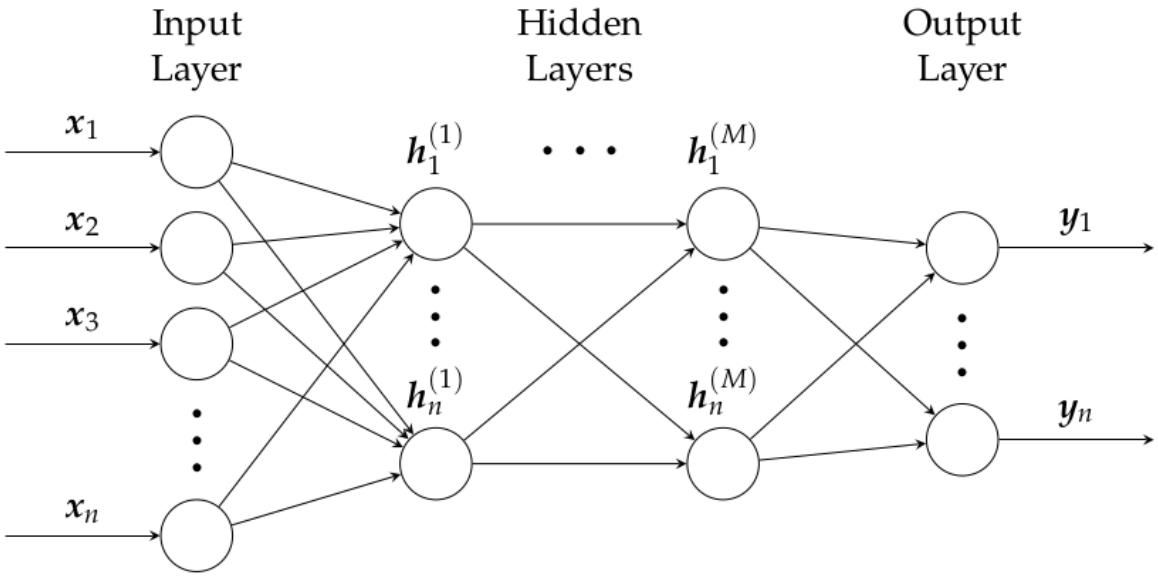

Figure 1.1: Schematic for fully connected feedforward neural network, with input layers represented by x_i, hidden layers represented by h_i^m, weights and biases represented by the lines drawn between the neurons, and output represented by y_i.

where this noise will cause oscillations in the interpolation. In addition, the noise also affects low resolution synthetic spectra, where line features are weaker due to the lower resolution. In these cases, noise will result in inaccurate feature detection. When interpolating synthetic spectra, the inputs are usually the stellar labels and the output is the spectrum; this forward modelling approach results in an output with higher dimensionality than the input, which is part of the reason for the noise. Therefore, it is possible to reduce the signal through reduction of dimensionality techniques such as principal component analysis (PCA), or smoothing techniques e.g: a convolutional layer in the neural network, or the noise can potentially be "averaged out" through the use of an ensamble of FFNNs. In addition to this noise issue, vanilla FFNN do not offer a fundamental way of quantifying systematic errors within the network. Quantifying systematic interpolation errors is crucial as it will allow a more accurate determination of stellar labels, especially on low SNR data where there are partial degeneracies between labels. In addition, areas with high systematic errors could indicate where additional training models need to be located. Bayesian neural networks (BNN) are a variation on FFNN that are potentially able to improve accuracy (Wenzel et al., 2020), reduce the noise-like signal in interpolating spectra, and quantify systematic interpolation errors.

BNN differ from FFNN in that their weights are posterior probability distributions instead of point estimates. This change allows us to capture the aleatoric (random) uncertainty, i.e. the uncertainty in the data set, and the epistemic uncertainty, i.e. the uncertainty in the model (Charnock et al., 2020). The change from point estimates will also alleviate

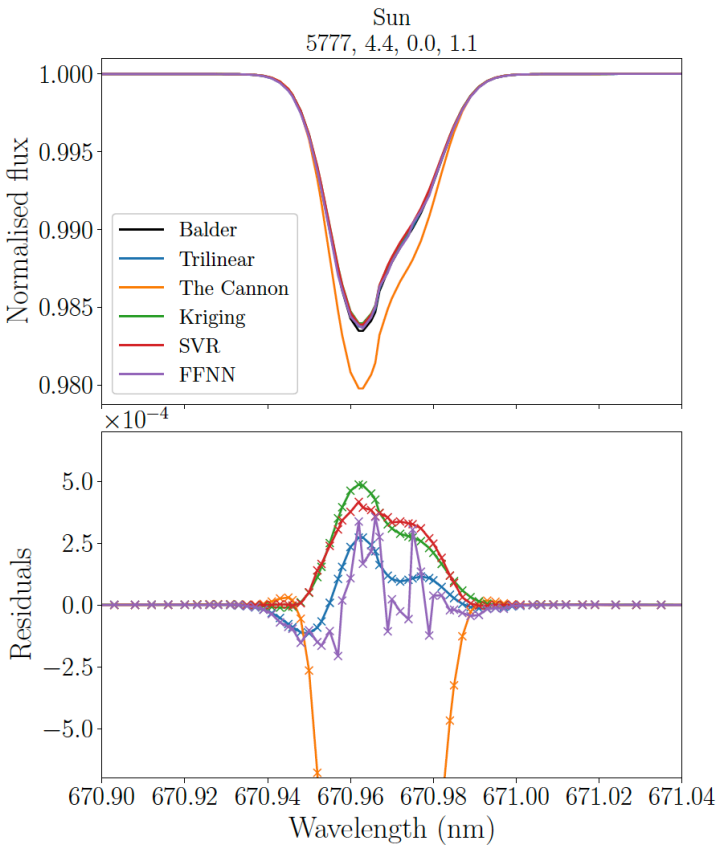

Figure 1.2: A test of how well five different interpolation methods perform on the Sun (Balder - shown in black), with residual computed by comparing the interpolated spectra to a directly synthesised one. The FFNN result is shown in purple, and the noise-like signal can be clearly seen in the residual.

the noise-like signal seen in interpolating spectra as we no longer have point estimates, instead a distribution. In practice, training a BNN with many posteriors is very computationally intensive, so approximations are necessary, for example, Gaussian approximations, Monte Carlo Markov chain (MCMC) methods, and variational inference are often used to approximate posteriors (Neal, 1996; Charnock et al., 2020).

1.4 Probability Density Estimators

In addition to quantifying the systematic errors in interpolating synthetic spectra, propagating the observational errors into the stellar labels would also improve accuracy. This can be done through sampling the posterior of the stellar labels given some observed spectra, with the added benefit of finding correlations between stellar labels.

MCMC methods comprise algorithms which approximate a posterior, this is done by sampling from the posterior to create a sample which has approximately the same distribution. Most MCMC methods rely on the use of walkers, initiated at random points in the probability distribution, that walk around the distribution based on the probability of taking that step, see Fig. 1.3 for an illustration. Each walker generates a chain based on the steps that it has taken, and this chain is what forms the approximated probability distribution. The downsides to MCMC is that the next step a walker takes is dependent on the current location of the walker, so each chain is dependent on the previous steps in the chain and is influenced slightly by its starting position. In order to avoid this influence, usually a large number of walkers are employed to produce independent chains. Furthermore, each walker is initiated randomly, which means that the first few hundred steps it takes is usually dependent on its initial position, so MCMC algorithms usually throw away the first few hundred steps the walker takes. This process is known as the "burn-in" period. In addition, MCMC walkers can easily get stuck in local minima, resulting in a wrong approximation of the distribution. The burn-in period, getting trapped in local minima, combined with the dependence on the previous step means that the use of more walkers results in a better approximation and a more independent sample. However, this is also much more computationally intensive due to the fact that every walker requires a separate burn-in. Despite these overheads, MCMC methods are used quite a lot in astrophysics, see Foreman-Mackey et al. (2013) and the applications within. Measuring stellar labels requires comparing the synthetic spectra to the observed spectra, the posterior of the stellar labels can be approximated through MCMC methods, which usually take $\sim$1-2 hours per star (Cargile et al., 2019).

Taking 1-2 hours to analyse each star becomes cumbersome when the sample size becomes large. Projects nowadays may have samples that contain millions of stars, so faster sampling is required. Normalising flows are probability density estimators like MCMC methods. Normalising flows transform a simple probability distribution to a more complex one through a series of normalised invertible mappings (Rezende & Mohamed, 2016). Normalising flows avoid all the aforementioned issues with MCMC methods due to the use of

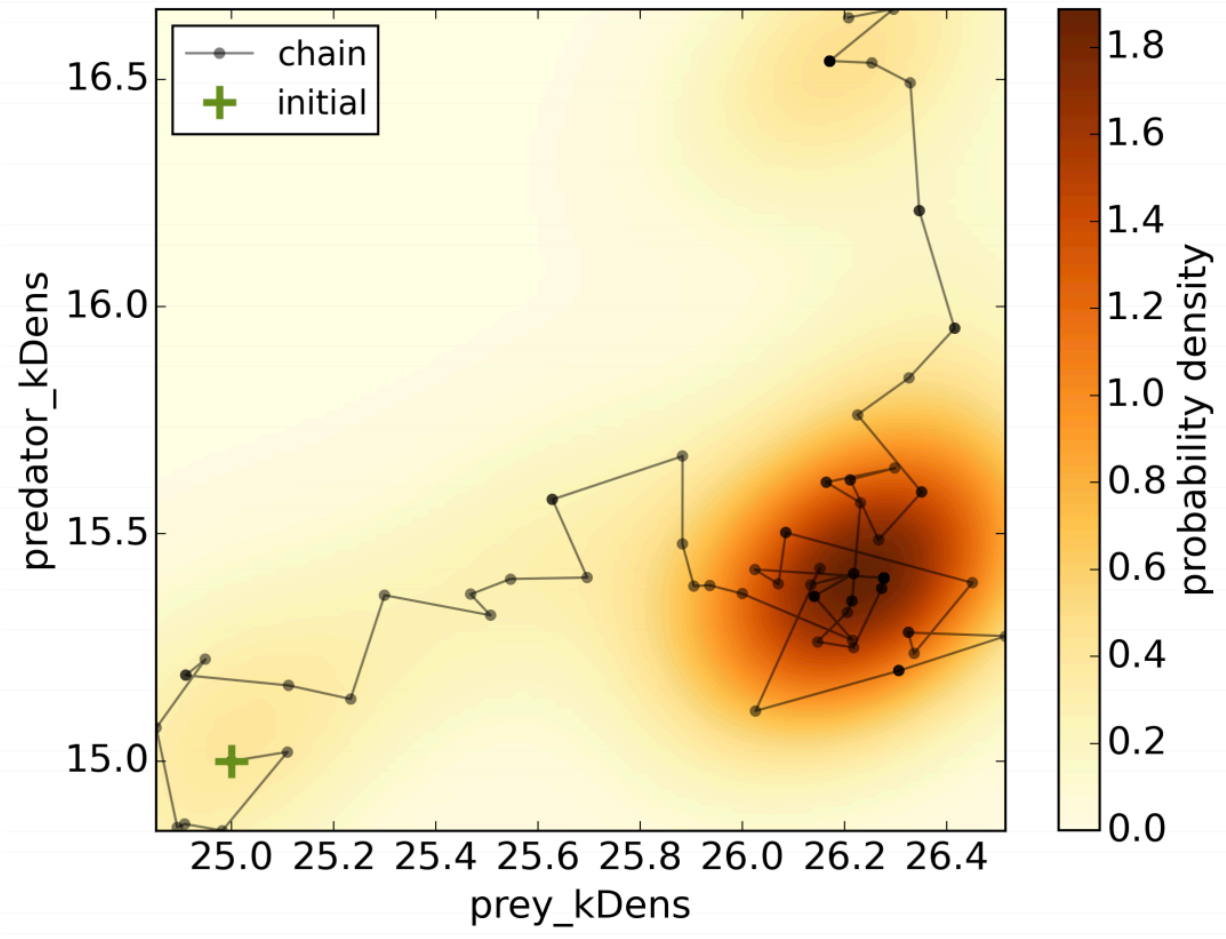

Figure 1.3: Illustrative MCMC walker figure from Šukys & Kattwinkel (2017), where the walker is initialised at the green plus (located at 25,15), and walks around the parameter space to form a chain (black dots). More steps in the chain come from higher probability density regions.

invertible mappings as opposed to walkers to approximate the probability distribution, and are faster than traditional MCMC methods. Within astrophysics, normalising flows have been used to approximate the posterior of physical parameters of stars from photometric observations (Ksoll et al., 2020).

Another probability density estimator is variational inference which turns a sampling of the posterior into an optimization problem, assuming some functional form for the posterior. In addition to MCMC and normalising flows, the proposed PhD project will also consider variational inferences as a possible probability density estimator that is faster than MCMC.

1.5 Thesis Overview and Outline

In this PhD thesis, I propose to use synthetic spectra computed from advanced spectrum synthesis methods to measure stellar parameters and elemental abundances (referred to as stellar labels) whilst quantifying systematic errors through BNN, improving accuracy through incorporating different observations, and improving sampling speed through normalising flows. Traditional stellar spectroscopic analyses focussed only on individual line features that are determined to be isolated (unblended), but the vast majority of spectral lines overlap to some extent. More information can be derived from spectra if all pixels are taken into account as opposed to only some pixels. However, using all the information contained in spectra is currently bottlenecked by two problems: incomplete and incorrect line parameters, slow and/or inaccurate stellar labels from advanced spectral synthesis methods.

Project 1 of this thesis will tackle a better simulation of line profiles for calcium (Ca) through use of 3D NLTE calculations. The line profile of calcium is dominated by metallicity, with some influence by Ca abundance (Da Costa, 2016). In particular the calcium infrared triplet (8498 Å, 8542 Å, and 8662 Å) is often used to estimate stellar radial velocities and metallicities, but also exhibits a smaller dependence on the abundance of calcium. Better line profiles from models will help improve radial velocity measurements, metallicity measurements as well as Ca abundance measurements.

Project 2 of this thesis will create a rapid but accurate pipeline that measures stellar labels whilst propagating systematic and observational errors. This will require improving synthetic line profiles through advanced spectrum synthesis methods, faster posterior sampling for accurate and reliable error estimates, using information in more spectral pixels, and including photometric data and parallax data in the analysis. This will be done as part of the analysis pipeline in GALAH, which is a high-resolution spectroscopic survey of hundreds of thousands of stars in the Galaxy (Buder et al., 2018).

Project 3 of this thesis will push the limit on extracting information from spectra at lower resolution and smaller wavelength range from the Gaia Radial Velocity Spectrometer (RVS) data set (Gaia Collaboration et al., 2016) by incorporating elements from both projects 1 and 2. Gaia RVS captures the Ca triplet, which we will have 3D NLTE synthetic profiles

of to determine the radial velocity, [Fe/H], and Ca abundance to a better accuracy. The lower resolution and smaller wavelength range of Gaia RVS compared to GALAH will likely result in a less precise and partially degenerate constraint on stellar labels. However, Gaia captures a larger data set than GALAH. The ability to quantify systematic errors, fast sampling of posterior, and incorporate additional observations (i.e. information) from project 2 will help produce robust estimates of stellar labels from the Gaia RVS data set.

Chapter 2

Proposal

2.1 Supervisory panel

Chair of supervisory panel: Prof. Helmut Jerjen (ANU).
Panel members: Dr. Thomas Nordlander, Dr. Yuan-Sen Ting (ANU, starting Dec 2020), Prof. Karin Lind (Stockholm University), and Dr. Michael Hayden (University of Sydney).
Future panel member: Dr. Ioana Ciuca (ANU, starting Nov 2020)

2.2 Project 1: 3D NLTE calcium line formation in late-type stars

Aim

The aim of this project is to produce a grid of 3D NLTE Ca synthetic spectra for FGK-type stars. The calcium H and K (3968Å and 3934Å) lines and the near-infrared triplet (Ca IRT) are very strong and form partially in the chromosphere of the star, where the high velocity fields are what will effect line shape the most. Simulating the chromosphere is hard, so instead I will test the change in line shape by adding velocity fields to the topmost layers of the photospheric atmosphere simulations, and produce a grid of 3D NLTE Ca synthetic spectra incorporating these effects.

Background

Ca lines indicate the Ca abundance. However, the Ca triplet (CaT) strength is dominated by metallicity, due to its sensitivity to the background opacity, dominated by H^-, which in turn depends on the electron density and temperature structure, both of which is determined by metallicity. As a result, CaT can be used to measure [Fe/H] in stars. This can be done in many different ways, one of which is by fitting synthetic line profiles and

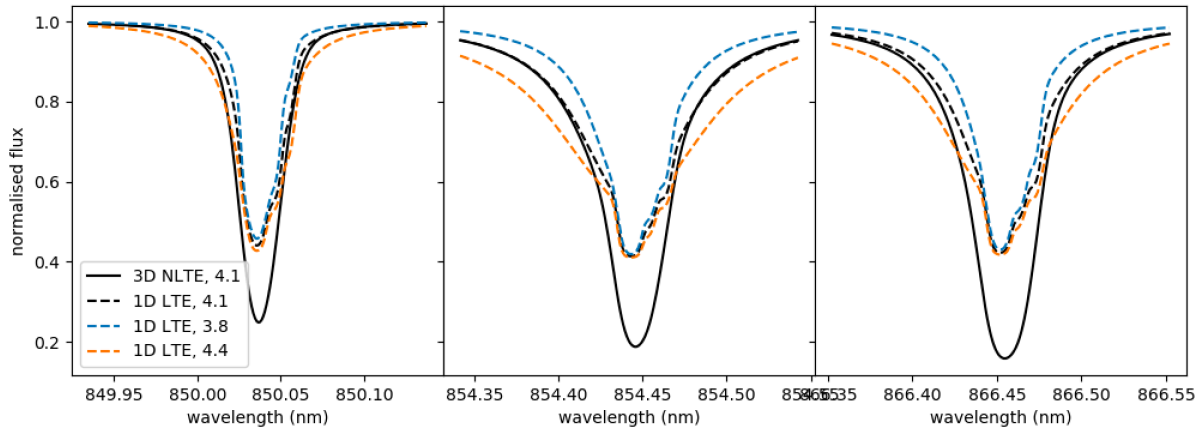

Figure 2.1: 3D NLTE (solid) and 1D LTE (dashed) synthetic profiles under different calcium abundances (A(Ca)), A(Ca) = 3.8 dex (blue), A(Ca) = 4.1 dex (black), and A(Ca) = 4.4 dex (orange). These profiles are computed with stellar parameters $T_{\rm eff}$ = 5800, $\log g = 3.7$, and [Fe/H] = −2.5.

computing the equivalent width (EW), then translating this to a metallicity measurement through formulas (e.g Starkenburg et al. (2010); Da Costa (2016)). Translating the fitted spectra to an EW is traditionally done because of the poor synthesis of the core of the CaT lines where this issue is avoided because EW is dominated by the wings of the CaT; and because most low or medium resolution surveys have large instrumental profiles at the CaT lines, removing all information except for the EW and radial velocity. Fitting synthetic line profiles is difficult due to the differences in shape of the profile computed with different atmospheres and assumptions. Fig. 2.1 shows the comparison of 3D NLTE and 1D LTE profiles of the Ca triplet with different Ca abundances. The core of the 1D LTE profiles are not as deep as the core of the 3D NLTE profile in this figure, and an increase in abundance in 1D LTE only causes the wings to increase in strength with minimal effect on the core. It is thus not possible to predict the true shape and strength of these lines under traditional 1D LTE spectrum synthesis. As a result, using CaT as metallicity indicators requires a more advanced approach.

Ca profiles in particular are hard to get right, owing to the fact that Ca lines are incredibly strong in most stars and form partially in the chromosphere. Strong lines are hard to simulate due to their higher opacity, pushing line formation higher in the atmosphere; as a result, strong enough lines can form outside of the stellar atmosphere bounding box. Not only can strong lines form outside the simulation box, but strong lines also often form in the chromosphere which has drastically different physical properties to the photosphere. Currently, most grids of stellar atmosphere models only simulate the photosphere, neglecting the chromosphere as most lines are not strong enough to form in the chromosphere. The chromosphere is hotter, less dense, and has higher velocity fields compared to the photosphere (Carlsson et al., 2016), see Fig. 2.2. Out of these differences, test calculations

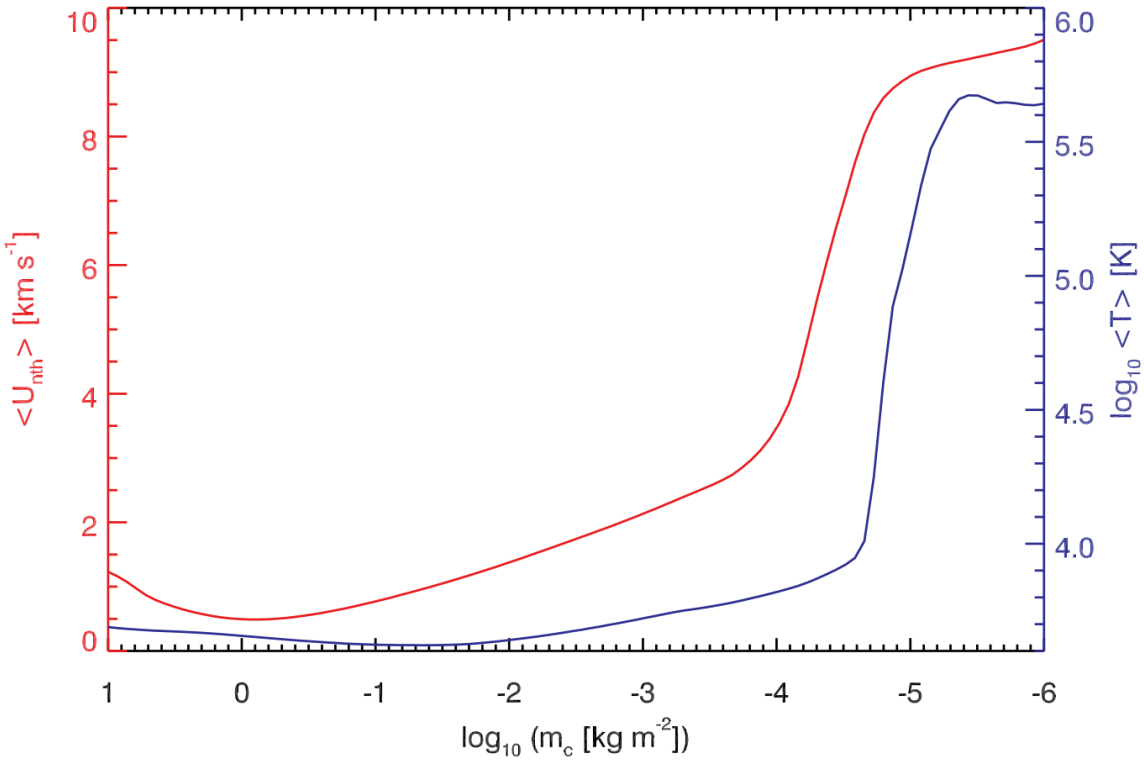

Figure 2.2: A theoretical chromosphere of the Solar chromosphere from Carlsson et al. (2016). Shown in the average velocity field (red, left) and the average temperature (blue, right) as a function of column mass ($m_c = \int \rho dz$, where ρ is density and z is the height).

indicate that the higher velocity field is what should affect the shape of the Ca IRT the most.

Datasets and Packages

I will compute 3D NLTE spectra for Ca, through the use of the Balder 3D NLTE spectrum synthesis code (Amarsi et al., 2016a,b, 2018) and the STAGGER 3D RHD stellar atmosphere grid (Magic et al., 2013). Calculations will be compared to archival spectra, for a set of well-studied benchmark stars, whose fundamental parameters are accurately known from photometric and interferometric methods. I use observed spectra from different instruments in order to compare observational effects, in particular, I will use spectra from: UVES-POP (Bagnulo et al., 2003), and PolarBase (Petit et al., 2014), the Solar Hamburg atlas (Neckel, 1999), PEPSI (Strassmeier et al., 2018a,b), and giraffe (Gilmore et al., 2012). These will be used to verify our methodology, and in particular our implementation of a pseudo-chromosphere (through the introduction of, e.g., artificial velocity fields).

Methodology and Analysis

The steps for this projects are as follows:

- Find rough Ca abundance of benchmark stars (Sun, HD84937, HD140283, Arcturus, and HD 122563) through matching 3D NLTE synthetic spectra against archive spectra. (Jul 2020)
- Check if perturbing the upper layers in the atmosphere with different methods makes the synthetic spectra more close in shape to the observed spectra. (Aug-Sep 2020)
- Compute a 3D NLTE grid of Ca synthetic spectra over FGK-type stars (Sep-Oct 2020)
- Apply this new grid to a science case and publish the results which has the biggest impact (Feb 2021), e.g
  - revising the extra-Galactic globular cluster metallicity scale for the WAGGS sample (Usher et al., 2017)
  - revising the metallicity results for dwarf spheroidal galaxies in the DART sample (Starkenburg et al., 2010)
  - show age metallicity relationship taking into account posteriors

Risks

The computational cost of a grid of 3D NLTE Ca is estimated to be ∼3 million CPU hours, or 6 MSU (assuming 5 snapshots per abundance and 5 abundances over 200 stellar atmosphere models, based upon test calculations). We have access to 25 million SU for the rest of 2020 on NCI, and an application for 2021 is in the works, and will likely yield 20 million SU, in addition to 7 million SU already granted. As a fallback, we also have access to the avatar cluster, which has 10 million SU per year.

We do not anticipate any other risks for this project. No comprehensive Ca grid exists in the literature, the largest grid contains 22 models (Mashonkina et al., 2007), therefore, a 3D hydrodynamic NLTE Ca grid would be a significant improvement regardless of whether the chromospheric effects can be replicated exactly or not.

Expected publications

Paper I: High-resolution 3D NLTE Ca line profile grid

Potential bonus paper: improved metallicity measurements for extra-Galactic globular clusters from the WAGGS sample (Usher et al., 2017), and metallicity results for dwarf spheroidal galaxies in the DART sample (Starkenburg et al., 2010).

2.3 Project 2: GALAH

Aim

The aim of this project is to analyse the spectra taken by GALAH combined with the photometry from SkyMapper (Wolf et al., 2018) and the integrated fluxes from the Gaia mission's BP and RP spectrometers, and the parallax measurements from Gaia, to produce more accurate stellar labels through the combined analysis of multiple data sets. I will produce a grid of 1D NLTE synthetic spectra for all elements captured in the GALAH survey to investigate NLTE effects on these observations. In addition, I will develop an new analysis pipeline incorporating advanced statistical methods which will speed up the time taken to derive stellar labels per star and quantify the systematics in the methods.

Background

There are many ways to observe stars, including but not limited to: spectroscopy, photometry, parallax, interferometry; but the analysis of multiple types of data sets simultaneously tend to be slow. Combined data analysis of all available observations and physical modelling would give more accurate results, and be especially useful in cases where one observation measures a quantity better than another. For example, $T_{\rm eff}$ and $\log g$ are incorrectly estimated from high-resolution spectra using 1D hydrostatic and under LTE assumptions (Lind et al., 2012), whilst photometry measures the shape of the spectral energy distribution, which is closely related to the definition of the effective temperature and yields nearly fundamental estimates (Casagrande et al., 2014); or if the data quality is insufficient to constrain all parameters simultaneously due to partial degeneracies. Stellar parameters are poorly constrained when simultanously determined in spectra with relatively low SNR due to the partial degeneracy and need to be better constrained with photometry.

We intend to use spectroscopy from GALAH. GALAH data is taken by the Anglo-Australian Telescope and has resolving power $\mathrm{R} = \lambda/\Delta\lambda = 28000$, with 4 discrete wavelength channels, covering 4713–4903 Å, 5648–5873 Å, 6478–6737 Å, and 7585–7887 Å, respectively. This high resolution allows for accurate abundance determination and the large wavelength coverage captures many elemental lines shown in Fig. 2.3. GALAH DR3 contains ∼600k stars, and by the DR4 will contain ∼800k stars. We intend to use photometry from SkyMapper (in the uvgriz filters; Wolf et al., 2018) and 2MASS (which has JHK_s bandpasses; Skrutskie et al., 2006). These photometric data sets cover a wide range of wavelengths, from 300nm to 2.3 micrometer, enough to cover the spectral energy distribution of any late-type star.

In addition to using extra information, consistent NLTE effects have not been studied to date. Most synthetic spectra are generated using the assumptions of LTE. GALAH DR3 incorporates departure coefficients which are calculated for each element separately. NLTE effects are driven by the presence of non-local radiation fields. These only propagate as far as the opacities allow them; in some cases line or continuous transitions of one element provide background opacities that block radiative processes in other elements. Such effects

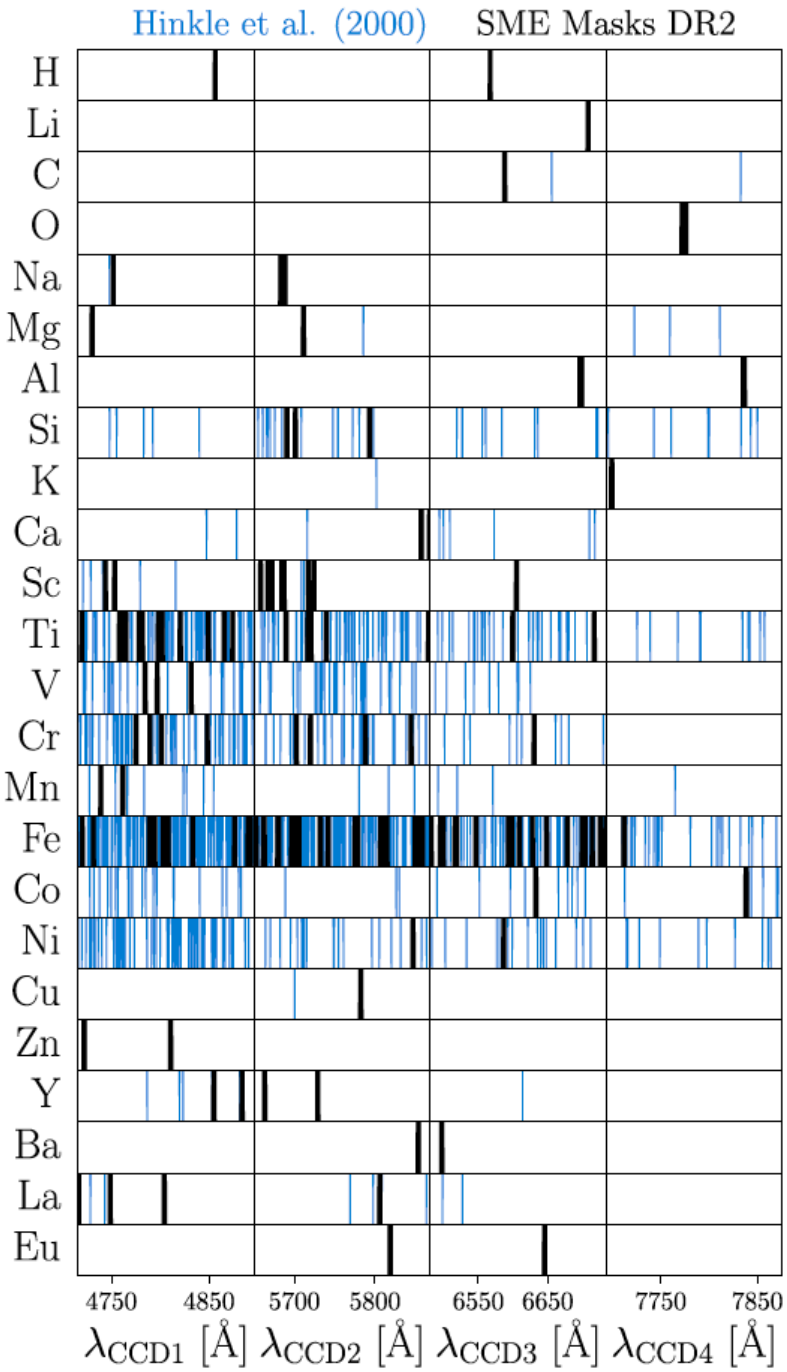

Figure 2.3: Illustration of the spectral lines visible for different elements within the four GALAH wavelength ranges from Buder et al. (2018). Blue and black lines are relatively unblended features with known accurate atomic data and are suitable abundance indicators on an absolute abundance scale.

may yield important second-order effects in cases where the background features have significantly altered strength under NLTE. Generating NLTE synthetic spectra computed for every element consistently will allow us to study NLTE effects consistently for all elements in the data set. Only recently have multi-element NLTE solutions become accessible, Osorio et al. (2020) computed NLTE solutions for 4 elements (Na, Mg, K, and Ca), we intend to do 14 (H, Li, C, N, O, Na, Mg, Al, Si, K, Ca, Mn, Ba, and Fe). Additional elemental abundances (Sc, Ti, V, Cr, Ni, Cu, Zn, Rb, Sr, Y, Zr, Mo, Ru, La, Ce, Nd, Sm, and Eu) provided in GALAH and molecular lines are computed under LTE. To compute a mixed LTE and NLTE synthetic spectrum, we first set up the atoms separately, then solve the statistical equilibrium for all elements at the same time. The LTE elements are assumed to be fixed background opacities in the NLTE solution. Then the emergent intensities are found through the solved level populations. Whilst it would be more accurate to use 3D hydrodynamic model atmospheres, it is currently too computationally intensive for multi-element NLTE solutions. Therefore, this project will make use of 1D hydrostatic model atmospheres.

The advancement in techniques for this project include quantifying the systematic errors caused by interpolation and improving the analysis speed per star. The current analysis method in GALAH only produces point estimates and does not explore the systematic errors. BNN has the ability to include errors within the estimates, and normalising flows will improve the analysis speed per star beyond what MCMC can achieve. This improvement in analysis will also hopefully solve the issue in GALAH DR3, where there is some clustering of stars around atmospheric grid nodes. This is due to some combined effect of poor atmospheric interpolation and the current fitting method, Spectroscopy Made Easy (SME). Fig. 2.4 shows this clustering in the distribution of $T_{\rm eff}$ for GALAH DR3, there are clear spikes in the distribution caused by clustering of stars around atmospheric grid nodes.

Data sets and Packages

I will use spectra from the GALAH survey, and photometry from SkyMapper, Gaia BP RP, and 2MASS. I am currently planning on using TensorFlowProbability and TensorFlow for the pipeline.

There has been a lot of work done on the atomic line parameters (Heiter et al., 2015; Buder et al., 2018; Amarsi et al., 2020, Buder et al., in prep), these will be retained in multi-element NLTE solution. The existing selection of lines are selected based upon relatively unblended features with known accurate atomic data (see Fig. 2.3). We will retain the existing line selection, but fit all areas simultaneously as opposed to one line at a time like in SME.

Methodology and Analysis

The steps for this project are as follows:

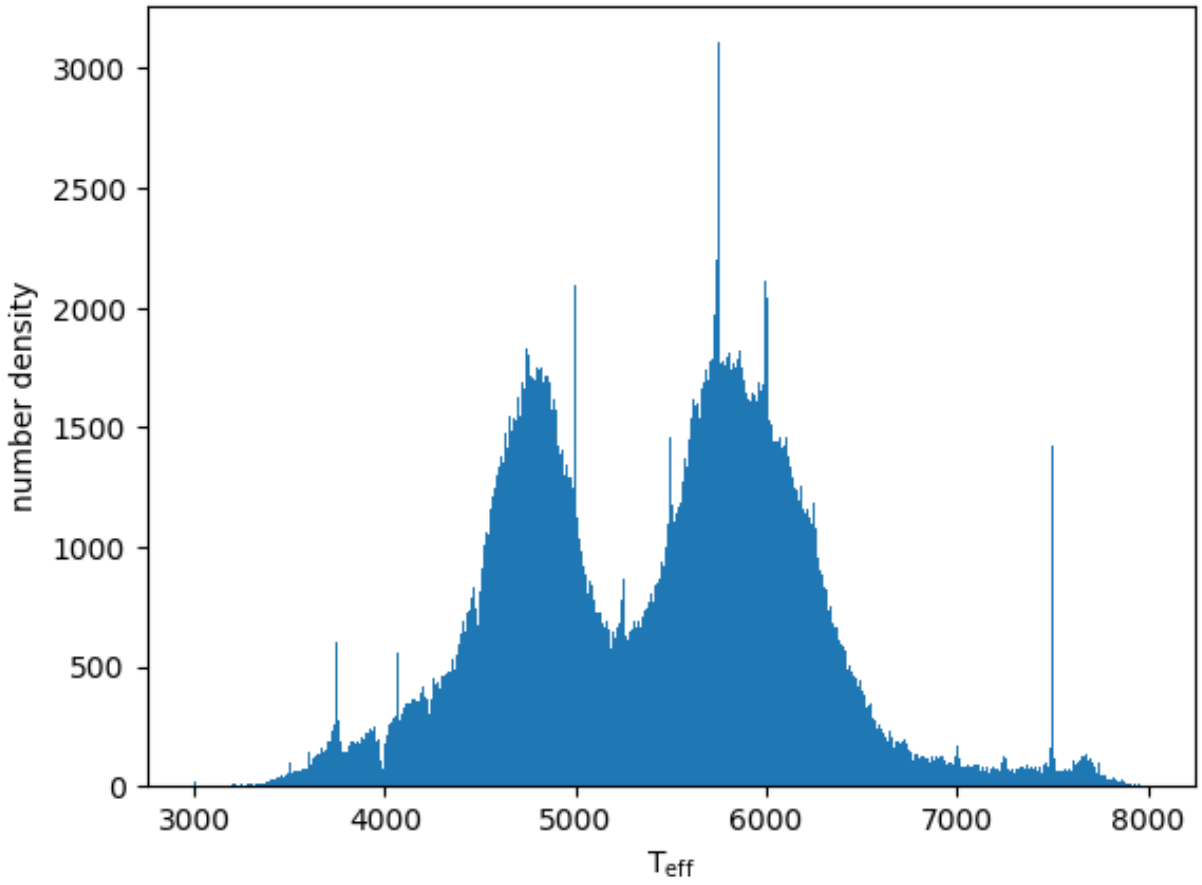

Figure 2.4: The number density of stars at different $T_{\rm eff}$ for GALAH DR3. Note the artificial spikes in the histogram (e.g. 3800 K, 4100 K, 4900 K, etc) caused by clustering of stars around atmospheric grid nodes.

- Generate a grid of 1D NLTE synthetic spectra (Apr-May 2021)
- Interpolate this grid to arbitrary stellar labels through use of ensemble of NN or BNN, picking the method which gives the better results (Jun-Aug 2021)
- Sample posterior of stellar labels through MCMC or normalising flows, using normalising flows if MCMC is too slow per star (Cargile et al. (2019) takes 1-2 hours per star, which implies 1-2 million CPU hours for a single run in GALAH) (Sep-Nov 2021)
- Test the pipeline against benchmark stars, stellar clusters, and co-moving groups (Dec 2021-Feb 2022)
- Apply the pipeline to GALAH data and photometric data (Apr-Jun 2022)
- Apply this new data set to a science case and publish the results which has the biggest impact (Jul-Aug 2022), e.g.
  - Systematic abundance differences amongst stars in comoving pairs and open clusters with different evolutionary state, in the context of atomic diffusion.
  - Star-to-star abundance amongst the light elements in globular clusters at low metallicity
  - show age metallicity relationship taking into account posteriors

Risks and Backup projects

This is an ambitious project both in terms of scope and methodology, with every proposed step in developing the pipeline an improvement upon the old. Given time constraints, we can implement fewer improvements as many improvements are independent of each other; it is also possible to scale back the data products of GALAH. Only improving the stellar parameters through combining spectroscopy and photometry whilst still keeping the analysis for abundances the same will still give more accurate abundances due to the dependence of abundances on stellar parameters.

1D NLTE solution and spectral calculation: ∼50000 CPU hours, or ∼100000 SU. 2.5 GPU years for experimenting with which interpolation model to use, or ∼1 million SU. For a total of ∼1.1 million SU. For 2021 and 2022, we expect to have similar time to 2020 on NCI Gadi and avatar, and potentially GPUs on Amazon.

Expected publications

Paper II: GALAH DR4, description of new analysis with posteriors of ages, stellar parameters and abundances for 500k+ stars, and the 1D NLTE grid of synthetic spectra.

In addition, every major GALAH data release is accompanied by a number of 5-10 supplementary science papers. As a core builder of the catalog, I will be working closely on

other papers in the GALAH collaboration which uses my data set to ensure the validity of my results.

2.4 Project 3: Gaia RVS

Aim

The aim of this project is to improve stellar labels and radial velocities measured from Gaia RVS spectra. I will improve metallicity and radial velocity measurements through using our 3D NLTE Ca grid from project 1 with Gaia RVS spectra. In addition, I will use the new pipeline developed in project 2 to improve stellar labels from the Gaia RVS data set.

Background

The Gaia RVS data set has resolving power R = 11700 and covers the wavelength range 8450–8720 Å. This wavelength range includes lines from Fe, Ca, Mg, Ti, and Si, which we can measure elemental abundances for. Gaia BP operates in the wavelength range 3300-6800 Å, and RP operates in the range 6400-10500 Å, capturing the same stars as the Gaia RVS instrument.

Gaia RVS spectroscopy will have ∼150 million stars (making this the largest ever spectroscopic dataset), out of which we can determine metallicities for ∼5 million and detailed abundances for the ∼2 million brightest stars (Starkenburg et al., 2010). At the same time, Gaia will capture exquisite photometry of these same stars through the use of the BP and RP photospectrometers, allowing us to apply the same methodology established in project 2 to improve the stellar parameters and abundances measured. In addition to photometry, Gaia will also provide parallax measurements for these same stars. Gaia RVS is a larger data set than GALAH but covers a smaller wavelength range. With this larger data set, there is more we can learn about Galactic archaeology. However, due to the smaller wavelength range of Gaia RVS, there are less elemental lines captured, meaning modifications will need to be made to the pipeline.

Elemental abundance measurements are directly tied to the assumed stellar parameters, therefore, improving stellar parameter determination should improve the accuracy of abundance measurements. The new 3D NLTE Ca grid from project 1 will provide more accurate metallicity measurements, in addition to more accurate Ca abundance measurements. Project 1 will also provide more accurate radial velocity measurements. 3D effects change radial velocity measurements. The wings of strong lines form deeper in the atmosphere whilst the core of strong lines form higher. Deeper in the atmosphere, convection has a net velocity upwards, whilst higher in the atmosphere the opposite is true. This causes a shift in the core of the line, resulting in a net shift of the zero point of the measured radial velocity.

It is extremely difficult to get any robust abundances within the Gaia RVS wavelength range. Through my PhD studies, we will push the boundary of low-resolution spectroscopy analysis, and demonstrate that we can extract multiple abundances robustly from Gaia RVS, as well as quantifying their uncertainties. This will improve the Galactic archeology science that comes from this data set.

Data sets and Packages

I will use Gaia RVS spectroscopy and Gaia BP RP and 2MASS photometry, improving metallicity measurements with the 3D NLTE Ca grid from project 1. I will also use the pipeline developed in project 2, which uses TensorFlowProbability and TensorFlow.

Methodology and Analysis

- Generate synthetic spectra for the Gaia RVS wavelength range (Sep-Oct 2022)
- Adapting the previously developed pipeline to Gaia data (Nov 2022-Jan 2023)
- Test pipeline on globular clusters, co-moving groups, and interferometric calibration stars (Mar-May 2023)
- Apply pipeline to Gaia RVS and photometric data (Jun-Jul 2023)
- Application to one science case and publish the results which has the biggest impact (Aug 2023)
  - Systematic abundance differences amongst stars in comoving pairs and open clusters with different evolutionary state, in the context of atomic diffusion
  - Flagging stars with elevated / depleted abundances of key elements that indicate particular nucleosynthetic processes
  - show age metallicity relationship taking into account posteriors

Risks and Backup data sets

Gaia RVS data set is not available yet, and has already been delayed by COVID-19. If Gaia RVS is not released in time, we can use the Rave DR6 data set which was recently released (Steinmetz et al., 2020). RAVE DR6 is similar to Gaia RVS in wavelength range, including an extra aluminium line, but with a lower resolving power of R = 7500, and only containing 450k stars. The extra aluminium line allows us to measure aluminium abundance and opens up extra science cases I could apply this data set to.

Analysing both Gaia RVS spectroscopy and Gaia BP RP photometry relies on the methodology for GALAH to work. Regardless of if I do successfully develop a new pipeline for GALAH or if I can apply it to Gaia data, we can still improve the metallicity measurement using our new 3D NLTE Ca grid.

We plan on retaining much of the analysis from project 2, so the estimated computational cost is lower. The testing of the interpolation model will cost 1.5 GPU years, or ∼375000 SU. We expect that we will still get NCI Gadi and avatar time in 2023 as well.

Expected publications

Paper III: Gaia RVS 3D NLTE analysis

Potential bonus paper: re-analysis of RAVE data

Chapter 3

PhD timeline

3.1 Breakdown of the projects

3.2 Research collaboration and travels

I will visit Stockholm University, Sweden to collaborate with Karin on the 3D NLTE Ca project, planned 2021 summer (if international travel opens up again). This planned visit is to work on the additional bonus paper for project 1, so whilst not critical to the completion of the PhD, it would be a good opportunity.

If there are no relevant international conferences in 2021, I will attend the ASA.

Future travels/online attendance for conferences and summer/winter schools will be chosen as they are announced.

Table 3.1: A proposed timeline for the PhD projects

Project	Timeline	Description
Project 1	Mar-Jun 2020	Literature review and test Ca atom and find Ca abundance on benchmark stars
	Jul-Aug	Check the effects of adding microturbulence and perturbing the velocity field of the topmost layers and write thesis proposal
	Sep-Oct	Compute grid of 3D NLTE Ca and give thesis proposal talk
	Nov-Jan 2021	ADACS internship: done during program leave, to learn proper programming practices (e.g: agile programming) and concepts (GPUs).
Project 1	Feb	Apply 3D NLTE Ca grid to other benchmark stars and field stars (metal-poor giants)
	Apr-Jun	Write up 3D NLTE Ca paper
Project 2	Apr-May	Create grid of 1D NLTE synthetic spectra
	Jun-Aug	Experiment with interpolating this grid to arbitrary stellar labels, testing both ensemble of NN and BNN
	Sep-Nov	Test speed of MCMC and normalising flows on sampling posterior of stellar labels
	Dec-Feb	Test the pipeline against benchmark stars, stellar clusters, and co-moving groups
	Apr-Jun 2022	Apply this pipeline to GALAH data and photometric data
	Jul-Aug	Apply this new data set to a science case
	Sep-Nov	Write up GALAH paper
Project 3	Sep-Oct	Generating synthetic spectra for Gaia RVS wavelength range
	Nov-Jan 2023	Adapting the GALAH pipeline to Gaia data
	Mar-May	Test on stellar clusters, co-moving groups, and interferometric calibration stars
	Jun-Jul	Generating the data set
	Aug	Apply data set to a science case
	Aug-Oct	Write up Gaia paper
	Nov-Dec	Write thesis and presentation

Bibliography

Amarsi A. M., Asplund M., Collet R., Leenaarts J., 2016a, , 455, 3735

Amarsi A. M., Lind K., Asplund M., Barklem P. S., Collet R., 2016b, , 463, 1518

Amarsi A. M., Nordlander T., Barklem P. S., Asplund M., Collet R., Lind K., 2018, , 615, A139

Amarsi A. M., et al., 2020, arXiv e-prints, p. arXiv:2008.09582

Armillotta L., Krumholz M. R., Fujimoto Y., 2018, , 481, 5000

Asplund M., Nordlund Å., Trampedach R., Allende Prieto C., Stein R. F., 2000, , 359, 729

Bagnulo S., Jehin E., Ledoux C., Cabanac R., Melo C., Gilmozzi R., ESO Paranal Science Operations Team 2003, The Messenger, 114, 10

Bland-Hawthorn J., Krumholz M. R., Freeman K., 2010, The Astrophysical Journal, 713, 166

Böhm-Vitense E., 1958, , 46, 108

Buder S., et al., 2018, Monthly Notices of the Royal Astronomical Society, 478, 4513

Cargile P. A., Conroy C., Johnson B. D., Ting Y.-S., Bonaca A., Dotter A., 2019, arXiv:1907.07690 [astro-ph]

Carlsson M., Hansteen V. H., Gudiksen B. V., Leenaarts J., De Pontieu B., 2016, Astronomy & Astrophysics, 585, A4

Casagrande L., et al., 2014, Monthly Notices of the Royal Astronomical Society, 439, 2060

Charnock T., Perreault-Levasseur L., Lanusse F., 2020, arXiv:2006.01490 [astro-ph, stat]

Da Costa G. S., 2016, Monthly Notices of the Royal Astronomical Society, 455, 199

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, , 125, 306

Francois P., Matteucci F., 1993, , 280, 136

Freeman K., Bland-Hawthorn J., 2002, Annual Review of Astronomy and Astrophysics, 40, 487

Gaia Collaboration et al., 2016, Astronomy & Astrophysics, 595, A1

Gilmore G., et al., 2012, The Messenger, 147, 25

Goodfellow I., Bengio Y., Courville A., 2016, Deep Learning. MIT

Heiter U., et al., 2015, , 90, 054010

Ksoll V. F., et al., 2020, arXiv:2007.08391 [astro-ph]

Lind K., Bergemann M., Asplund M., 2012, Monthly Notices of the Royal Astronomical Society, 427, 50

Magic Z., Collet R., Asplund M., Trampedach R., Hayek W., Chiavassa A., Stein R. F., Nordlund Å., 2013, , 557, A26

Mashonkina L., Korn A. J., Przybilla N., 2007, , 461, 261

Mott A., Steffen M., Caffau E., Strassmeier K. G., 2020, , 638, A58

Neal R. M., 1996, Bayesian Learning for Neural Networks, 1 edn. Lecture Notes in Statistics 118, Springer-Verlag New York

Neckel H., 1999, , 184, 421

Ness M., Hogg D. W., Rix H. W., Ho A. Y. Q., Zasowski G., 2015, , 808, 16

Nordlund A., 1982, , 107, 1

Nordlund A., Dravins D., 1990, , 228, 155

Osorio Y., Allende Prieto C., Hubeny I., Mészáros S., Shetrone M., 2020, Astronomy & Astrophysics, 637, A80

Pereira T. M. D., Asplund M., Collet R., Thaler I., Trampedach R., Leenaarts J., 2013, , 554, A118

Petit P., Louge T., Théado S., Paletou F., Manset N., Morin J., Marsden S. C., Jeffers S. V., 2014, , 126, 469

Rezende D. J., Mohamed S., 2016, arXiv:1505.05770 [cs, stat]

Rosenblatt F., 1958, Psychological Review, 65, 386

Sellwood J. A., Binney J. J., 2002, , 336, 785

Skrutskie M. F., et al., 2006, The Astronomical Journal, 131, 1163

Starkenburg E., et al., 2010, Astronomy and Astrophysics, 513, A34

Steinmetz M., et al., 2020, The Astronomical Journal, 160, 83

Strassmeier K. G., Ilyin I., Steffen M., 2018a, , 612, A44

Strassmeier K. G., Ilyin I., Weber M., 2018b, , 612, A45

Ting Y.-S., Conroy C., Rix H.-W., Cargile P., 2019, The Astrophysical Journal, 879, 69

Usher C., et al., 2017, Monthly Notices of the Royal Astronomical Society, 468, 3828

Wenzel F., et al., 2020, arXiv:2002.02405 [cs, stat]

Wolf C., et al., 2018, Publications of the Astronomical Society of Australia, 35, e010

Šukys J., Kattwinkel M., 2017, arXiv:1711.01410 [cs, stat]