Advances in the modeling of stellar spectra, and applications to the Galaxy and its stars

Adam Joseph Wheeler

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2022

© 2022

Adam Joseph Wheeler

All Rights Reserved

Abstract

Advances in the modeling of stellar spectral, and applications to the Galaxy and its stars

Adam Joseph Wheeler

Large stellar surveys are revealing the chemodynamical structure of the Galaxy across a vast spatial extent. However, the many millions of low-resolution spectra observed to date have not yet been fully leveraged. In chapters 2 and 3, we employ data-driven spectroscopic models to the low-resolution LAMOST survey ($' = 1800$). In chapter 2, we employ *The Cannon*, a data-driven approach for estimating chemical abundances, to obtain detailed abundances, using the GALAH survey as our reference. We deliver five (for dwarfs) or six (for giants) estimated abundances representing five different nucleosynthetic channels, for 3.9 million stars, to a precision of 0.05 - 0.23 dex. Using wide binary pairs, we demonstrate that our abundance estimates provide chemical discriminating power beyond metallicity alone. We show the coverage of our catalogue with radial, azimuthal and dynamical abundance maps, and examine the neutron capture abundances across the disk and halo, which indicate different origins for the *in-situ* and accreted halo populations. LAMOST has near-complete *Gaia* coverage and provides an unprecedented perspective on chemistry across the Milky Way.

Stars with unusual levels of enrichment in a particular element are of great interest, but often pose a problem for data-driven methods. In chapter 2, we present a simple method, for the detection of $-$-enriched stars, for arbitrary elements $-$, even from blended lines. Our method does not require stellar labels, but instead directly estimates the counterfactual unrenriched spectrum from other unlabelled spectra. We apply this method to the 6708 Å Li doublet in LAMOST DR5,

identifying 8,428 Li-enriched stars seamlessly across evolutionary state. We comment on the explanation for Li-enrichement for different subpopulations, including planet accretion, nonstandard mixing, and youth.

The Galactic disk exhibits complex chemical and dynamical substructure thought to be induced by the bar, spiral arms, and satellites. In chapter 4, rather than calculating spectroscopic quantities, we use them to understand the Milky Way. We explore the chemical signatures of bar resonances in action and velocity space and characterize the differences between the signatures of corotation and higher-order resonances using test particle simulations. Thanks to recent surveys, we now have large datasets containing metallicities and kinematics of stars outside the solar neighborhood. We compare the simulations to the observational data from *Gaia* EDR3 and LAMOST DR5 and find weak evidence for a slow bar with the "hat" moving group ($250\ \mathrm{km\ s^{-1}} < E_q < 270\ \mathrm{km\ s^{-1}}$) associated with its outer Lindblad resonance and "Hercules" ($170\ \mathrm{km\ s^{-1}} < E_q < 195\ \mathrm{km\ s^{-1}}$) with corotation. While constraints from current data are limited by their spatial footprint, stars closer in azimuth than the Sun to the bar's minor axis show much stronger signatures of the bar's outer Lindblad and corotation resonances in test particle simulations. Future datasets with greater azimuthal coverage, including the final *Gaia* data release, will allow reliable chemodynamical identification of bar resonances.

Finally, in chapter 5, we present KORG, a new package for 1D LTE (local thermal equilibrium) spectral synthesis, which computes theoretical spectra from the near-ultraviolet to the near-infrared, and implements both plane-parallel and spherical radiative transfer. It is compatible with automatic differentiation libraries, and easily extensible, making it ideal for statistical inference and parameter estimation applied to large data sets. We outline the inputs and internals of KORG, and compare its output spectra to those produced by other codes. We use five example wavelength regions across 3660 Å – 15050 Å to show that the residuals between KORG and the other codes are no larger than that between existing codes themselves. We show that KORG is 1–100 times faster than other codes in typical use.

Table of Contents

Acknowledgments xvii

Chapter 1: Introduction 1

1.1 A brief history of early stellar spectroscopy 1

1.2 Modeling stellar spectra today 2

1.2.1 Spectral synthesis 2

1.2.2 Data-driven spectral models 4

1.2.3 Challenges in modelling stellar spectra 6

1.3 Recent, current, and future stellar spectroscopic surveys 7

1.4 Applications of spectroscopically-derived quantities 8

1.4.1 Nucleosynthesis 8

1.4.2 Galactic archaeology 10

1.4.3 Stellar physics 12

Chapter 2: Abundances in the Milky Way across ve nucleosynthetic channels from 4 million LAMOST stars 13

2.1 Introduction 13

2.2 Data 16

2.2.1 Quality cuts and data cleaning 18

2.2.2 Dwarf and giant models 19

2.3 Model 21

2.4 Model Evaluation 25

2.4.1 Model interpretability 27

2.5 Results 27

2.5.1 Catalog 27

2.5.2 Detailed abundances of wide binaries 32

2.5.3 Mapping chemistry in the Milky Way 35

2.5.4 The disk-halo transition seen in chemistry 40

2.6 Discussion & Conclusions 42

Chapter 3: An unsupervised method for identifying Li-enriched stars directly from spectra: Li in LAMOST 45

3.1 Introduction 45

3.2 Data 47

3.3 Methods 47

3.3.1 Tangent Space Projection 48

3.3.2 Identifying Strong Absorption in the Residuals 50

3.3.3 Hyperparameter Selection 52

3.4 Results 53

3.4.1 Planet Accretion on the Subgiant Branch 56

3.4.2 Li-enriched Red Giants 58

3.4.3 Young and pre-Main Sequence Stars 60

3.5 Discussion 61

3.5.1 Blended lines 64

3.5.2 Algorithm 64

3.5.3 What could possibly go wrong? 65

3.6 Conclusions 67

Chapter 4: Chemodynamical signatures of bar resonances in the Galactic disk: current data and future prospects 68

4.1 Introduction 68

4.2 Data 71

4.2.1 The MW disk as seen in this dataset 73

4.3 Simulations 75

4.4 Results 76

4.4.1 Metallicity signatures of bar resonances in action space 76

4.4.2 Action-metallicity signatures in the LAMOST data 81

4.5 Discussion 83

4.6 Conclusions 88

Chapter 5: Korg: a modern 1D LTE spectral synthesis package. 91

5.1 Introduction 91

5.2 Description of code 92

5.2.1 Inputs 92

5.2.2 Chemical equilibrium 94

5.2.3 Continuum absorption 95

5.2.4 Line absorption 96

5.2.5 Radiative transfer 98

5.3 Scienti c veri cation 99

5.3.1 Continuum absorption in the violet and ultraviolet 100

5.3.2 Balmer series 105

5.3.3 C_2 band 105

5.4 Benchmarks 106

5.5 Conclusions 108

Conclusion 110

5.6 Summary of results 110

5.7 Re ections and future directions 111

References 113

Appendix A: Notes on Chapter 2 132

A.1 Linear Coef cients 132

A.2 Radial velocity precision 132

A.3 GALAH windows 132

A.4 Action and azimuthal velocity uncertainty 132

Appendix B: Notes on Chapter 4 137

B.1 Simulated observational error 137

B.2 Axisymmetric approximate actions calculated with the Eilers 2019 potential 137

B.3 Alternate metallicity assignment 137

B.4 APOGEE data 141

Appendix C: Notes on Chapter 5 142

C.1 Numerical evaluation of transfer integrals 142

List of Figures

2.1 Face on (top) and edge on (bottom) Contours in surface density (10^{5} $M_\odot$ kpc^{-2}) shown in heliocentric Galactic coordinates for GALAH and LAMOST, which probes much farther into the outer disk and halo. The Galactic center is at $X_\odot = 8$ kpc, $Z_\odot = 0$. 17

2.2 A comparison of the GALAH and LAMOST spectra for the same star, 2MASS J00010184+0407201, Gaia DR2 2740354684364096000. The top panels show part of the star's GALAH spectrum (S/N = 65) and Cannon model, while the bottom two show the same for LAMOST (S/N = 179). On the left, note the large v_V line and surrounding features, on the right, observe the fit around a known Mn feature (highlighted in red on the left). Both the measured spectra and Cannon models are shown with their 1σ errors, shown with error bars in the right-hand panels and a shaded region in the left-hand panels. 17

2.3 Our 1722 training stars in the LAMOST Kiel diagram space colored by (GALAH) [Fe/H]. Since there are more metal-poor stars in the giant training set, our giant model is unbiased down to lower metallicity. 19

2.4 Our 1722 training stars across the (LAMOST) Kiel diagram in bins of metallicity. For the extreme ends of our metallicity range, the lack of training set coverage may bias our estimated labels. 20

2.5 Mean fractional distance error as a function of Galactocentric radius. Distance errors blur chemical maps of the Milky Way far from the solar annulus. 21

2.6 The Cannon likelihood as a probabilistic graphical model. During training, the latent variables in the “stars” panel (ℓ) are fixed using stars for which labels are known, in our case from GALAH. When inferring stellar labels, the latent variables in the “pixels” panel (ℓ and B) are fixed to their point estimates from training and the maximum likelihood estimates for all are calculated. 24

2.7 Cross-validation recovery of training set labels for the dwarf model. Horizontal lines show boundaries beyond which the model is biased or unprobed by the training set. 25

2.8 Analogous to Figure 2.7. Cross-validation recovery of training set labels for the giant model. Horizontal lines show boundaries beyond which the model fails to extrapolate. While the scatter in [Fe/H] is higher at lower values, the model appears to be nearly unbiased down to [Fe/H] = 1.5. 26

2.9 The precision of each of our abundances relative to the range over which the model is approximately unbiased. We generally infer abundance ratios with precision at the 30-50% level. Smaller values indicate higher relative precision of that abundance and presumably higher discriminating power between stars. 26

2.10 A portion of the best-fit model spectra (black) and real data (colored), both with 1-σ uncertainties, for 6 randomly-chosen stars in our catalog. Though simple, the model is flexible enough to fit the data across the Kiel diagram. 28

2.11 Inferred [Fe/H] vs effective temperature for LAMOST and GALAH dwarfs in Praesepe. The GALAH values come from internal DR2, which used the same analysis pipeline as the DR2 values in our training set. The systematic trend with $T_{\rm eff}$ is spurious, since all stars in Praesepe have the same abundances to below the precision achievable with low-resolution spectra. 31

2.12 Error-weighted chemical distances for EB19 wide binaries and for random field pairs selected with the same cuts and chosen to have the same [Fe/H] distribution as the EB19 sample. The wide binaries are more chemically similar than implied by their similarity in bulk metallicity alone. 33

2.13 The ¹'–I º plane colored by mean label value for nine labels of our 800,000 giant stars, summed over all. The flaring of the disk can be seen in [Eu/Fe], [Mg/Fe], and [Mn/Fe]. 36

2.14 Mean abundances of LAMOST giants in the thin disk (Ÿ 30 kpc km s^{1}), mapped in the ¹-–. º plane.) $_e$ and $_+$ are also mapped for comparison. 38

2.15 Mean-abundance maps of thin disk (Ÿ 30 kpc km s^{1}) LAMOST giants in the ¹'–E$_q$º plane. The dashed box outlines the longest ridge of increased stellar number density. We plot only stars with log¹6º Ÿ 3 to minimize contamination from dwarfs, whose abundances are generally on a different scale and $_e$-have dependent systematic trends, but a vertical feature is still visible at the solar radius, ' = 8 kpc, presumably induced by the selection function. 39

2.16 The ([Ba/Fe]-[Fe/H]) plane, colored by azimuthal velocity, E$_y$, and plotted in spatial bins in the Galaxy, with (up to) 700 randomly selected stars plotted in each bin. Because of the $_e$ -dependent systematics in our inferred [Ba/Fe] values, we have only plotted stars with 4800 K Ÿ $_e$ Ÿ 5000 K. With increasing j I j, halo stars (as distinguished by their lower -velocities) become more dominant. 40

2.17 Same as gure 2.16, with [Eu/Fe] in place of [Ba/Fe] and with stars chosen without restriction on $_e$ 41

3.1 Algorithm summary. 49

3.2 Our line model (black) and contaminant models (grey, dashed). Only one single-pixel contaminant model is shown for clarity. 51

3.3 Mean squared error of the predicted spectral region as a function of subspace dimension, @ for different neighborhood sizes, The dashed grey line marks the best j 2• pixel (: = 3) achieved by using the local mean spectrum as a prediction. We use: = 1000, @= 50 for our analysis. 53

3.4 The prediction of the Li absorption region in a spectrum of LAMOST J062739.73+463634.4, a Li-enriched giant. The imputed spectrum for = 30– @= 4 is very similar to that for : = 1000– @= 50. Neither predict the absorption seen in the data. The 30 nearest spectral neighbors are shown in gray. 54

3.5 left: Portions of 21 spectra of Li-enriched stars spaced roughly evenly on the Kiel diagram and sorted by U (6565 Å) amplitude. right: their positions on the Kiel diagram, with mass tracks for 2 , 1.5 " , 1 " , 0.8 " , and 0.5" solar metallicity stars (details in text). 55

3.6 The distribution of excess equivalent widths for all candidates (those with f EEW Ÿ 0•05Å) and for Li-enriched stars (candidates with EEW 0•15Å). 56

3.7 top: The 8,428 Li-enriched stars plotted on the Kiel diagram with MIST solar-metallicity mass tracks. bottom left: fraction all stars with f EEW Ÿ 0•05 Å with are Li-enriched. bottom right: the number of of stars with f EEW Ÿ 0•05 Å. Dashed vertical lines are the same as in Figure 3.9. They mark the range over which young stars are Li-enriched. 57

3.8 The raw count and occurrence fraction of Li-enriched giants as a function of metallicity. Errorbars are determined assuming Poisson uncertainty in raw counts. The fraction of Li-enriched giants increases smoothly with [Fe/H]. 61

3.9 Li EEWs as a function of f e for the Pleiades stars observed by LAMOST. Note that the x-axis is reversed. The grey dashed line marks the threshold at which we consider a star Li-enriched. The decrease in EEW with increasing effective temperature is due to Li's low ionization potential. The decrease with decreasing effective temperature is because those stars have depleted more of their Li. 62

3.10 A portion of a spectrum of an Ba-rich star (top; LAMOST J120744.31+172831.6) and an Fe-rich star (bottom; LAMOST J093835.58-064404.6) as observed by LAMOST. TSP correctly imputes the masked region (marked in grey) for the Fe-rich star, but not the Ba-rich star. Note that the continuum is underestimated to different degrees in each, so the y-axes are not directly comparable. 63

4.1 The radial distribution of stars within 1 kpc from the Galactic midplane in each input catalog, as well as in the crossmatched joint catalog. LAMOST radial velocities can fruitfully augment Gaia in the outer disk. 71

4.2 The LAMOST $_q - \cdot$ and $' - E_q$ planes in stellar density (top row; 10^6 stars), fraction of inwards-moving stars (second row; 10^6 stars), mean [Fe/H] (third row; 10^5 stars) and mean [Mg/Fe] (bottom row; 10^5 stars). The OLR axisymmetric resonance lines for the three pattern speeds we consider are overplotted on the $_q - \cdot$ plane, and contours of constant $_q$ corresponding the circular OLR orbits for the same pattern speeds are overplotted on the $' - E_q$ plane. The top four panels contain 800–000 stars and the bottom four contain 95–000, because we made additional cuts to mitigate abundance systematics. We use fewer stars when plotting metallicities or abundances because we restrict our sample to a small part of log¹6º–) $_e$ space to minimize systematics. 74

4.3 The metallicity signatures at the bar resonances in the intermediate bar test particle simulation. A and B: mean metallicity in the meridional plane for the axisymmetric and barred test particle simulations, respectively. C and D: number density in the $_q - \cdot$ plane for the axisymmetric and barred test particle simulations, respectively. E and F: mean [Fe/H] in the $_q - \cdot$ plane for the axisymmetric and barred test particle simulations. G: mean [Fe/H], in bins of $_q$ and $\cdot$, for the axisymmetric (dashed) and intermediate bar (solid) test particle simulations. The CR resonance attens the mean-metallicity trends, while the higher-order resonances (especially the OLR) create wave-crest signatures. H: differences between the above: the difference between mean [Fe/H] of each bin and the lowest $\cdot$ bin. Error bars in the bottom two panels show the standard error in each bin. Except at the largest values of and $_q$, the standard errors are too small to be visible, indicating that the trends are robust. The colored bands show the range of values each axisymmetric resonance line takes on between 160 kpc km s^{1}. See also Appendix B.1. 77

4.4 Mean [Fe/H] as a function of $_q$ in bins (same as Figure 4.3) of for particles evolved in all three simulations. Error bars show the standard error. The qualitative attening and wave-crest signatures are not sensitive to pattern speed. The vertical bands show the location of axisymmetric resonance lines, as in Figure 4.3. 79

4.5 Mean [Fe/H] as a function of $_q$ in bins (same as Figure 4.3) of for particles evolved with two different bar strengths, but the same pattern speed. Error bars show the standard error. Because the stronger bar traps more stars on resonant orbits, the metallicity signatures at CR and the OLR cover a larger range in $_q$, and are therefore also stronger at larger 80

4.6 Mean [Fe/H] as a function of $_q$ in bins of ' for LAMOST. The errorbars show the standard error on the mean. There are not enough stars with $_q$ 60 kpc km s^{1} to plot trends. Grey vertical bands mark the approximate $_q$ of known moving groups for orientation. Gray dashed lines mark the $_q$ values of circular orbits with ' at the radius where the solar azimuth intersects hypothesized spiral arms (taken from Reid et al. 2014). Observational data is not yet plentiful enough to clearly identify or rule out the resonance signatures described in section 4.4.1, but the attening around 1400 kpc km s^{1} Ÿ $_q$ Ÿ 1700 kpc km s^{1} and the elbow at $_q$ 2200 kpc km s^{1} are suggestive of a slow bar. 82

4.7 The location of observed stars (i.e. the spatial selection function) has a strong impact on the attening and wave-crest signatures. Here, resonance signatures as exhibited by stars in different azimuthal regions of the disk in the intermediate bar simulation are plotted. ' bins above 60 kpc km s^{1} are not plotted because of the lack of suf cient data. Each row shows stars within a cylinder centered at the solar Galactocentric radius at different azimuths: exactly along the minor axis of the bar (top), in the position of the Sun (middle rows), and exactly along the bar's major axis (bottom). The left column shows number density in the plane, i.e. the footprint of each mock data set in the disk. The white line marks the orientation of the bar, and the white star marks the location of the Sun. The right column shows metallicity trends in $_q$ and ' for each region. In row 3, only stars with azimuth farther from the bar azimuth than the Sun are used. The CR attening and the OLR wave-crest are much more prominent in stars anti-aligned with the bar, but could be seen in a large volume of local data by plotting only stars farther from the bar's azimuth than the Sun (row 3). 84

4.8 The “banana-shaped” orbits of stars trapped at CR spend little time near the bar's major axis, and many dwell exclusively near its minor axis, which explains the azimuthal sensitivity demonstrated in Figure 4.7. The relative slope of the [Fe/H] relation around CR with respect to the overall slope for stars with $\dot{Y}$ 60 kpc km s^{-1} as a function of position in the disk for the intermediate bar simulation. The black rectangle indicates the orientation of the bar. The orbits of two stars trapped by the CR resonance are overplotted (in the rotating frame). 86

4.9 Mean [Fe/H] as a function of R_g and ϕ for the intermediate bar simulation with the LAMOST's spatial selection function imposed. top: full LAMOST footprint. bottom: partial LAMOST footprint, with only stars with azimuth less than the Sun's. The attening and wave-crest signatures at CR and the OLR, respectively, are extremely attenuated because stars are concentrated near the Sun. 87

5.1 Derivatives of a synthesized solar spectrum with respect to various abundances, $\frac{m F}{m\ ^{1}\text{-}\ ^{o}}$. Vertical offsets have been applied, but the relative scaling is preserved, which is why some derivative spectra appear at here (but show uctuations at a sub-percent level). 93

5.2 Portions of a synthetic solar spectrum generated with KRG, MOOG, and TURBOSPECTRUM, as well as the continuum generated with an empty linelist from 4000 Å to 10000 Å. All wavelengths are vacuum. The residuals (KRG - other) are shown underneath the recti ed ux for each wavelength region. The level of agreement between KRG and the other codes is similar to their agreement with each other. See text for a discussion of the discrepancies. 101

5.3 Same as Figure 5.2, but showing the synthetic spectrum of Arcturus. 102

5.4 Same as Figure 5.2, but showing the synthetic spectrum of HD122563. 103

5.5 Same as Figure 5.2, but showing the synthetic spectrum of HD49933. 104

5.6 The time taken by KORG, MOOG, and TURBOSPECTRUM for the syntheses in Section 5.3. top: average compute time for the optical wavelength regions, which used very dense linelists. bottom: compute time to synthesize spectra from 15,000 – 15,500 Å with the APOGEE DR16 linelist, a more "realistic" use case. Note that the wavelength regions synthesized were larger than those plotted in Figures 5.2 – 5.5. 107

5.7 top: Time required to compute simultaneous derivative of spectrum with regard to # different element abundances, for varying # 108

A.1 Linear coefficients for the dwarf model, omitting high-uncertainty pixels with $\bar{\sigma}_Y$ 4000. 133

A.2 Same as Figure A.1, but for our giant model. 134

A.3 Reported RV precision of stars in our catalog from Gaia, and LAMOST. While the LAMOST precision is worse, Gaia only measured RVs for approximately one fifth of our catalog. 135

A.4 Median values and uncertainties of the three actions and azimuthal velocity as a function of Galactocentric radius r. Note that the error bars here do not show scatter, but median uncertainty. 135

B.1 Same as Figure 4.3, but with simulated observational radial velocity and distance error applied. The left column shows the unperturbed simulation and the right column shows the perturbed one. Both are barred. Measurement error on radial velocity and distance has only has very small effect on the qualitative signatures discussed in this paper. 138

B.2 LAMOST data just as in Figure 4.6, but with actions calculated using the Eilers et al. 2019 potential. The overplotted resonances are for $R_0 = 36$ kpc km s^{-1}. The CR and 1:1 resonances are not within the plotted J_ϕ range. The use of this potential has no appreciable impact on the interpretation of the observational data. 139

B.3 Similar to Figure 4.3, but comparing the intermediate bar simulation with metallicities painted on according to initial action with two different dependencies on (left and right column, dashed and solid lines, respectively). See text for details. Our proposed signatures are largely the same when metallicities are assigned this way. 140

B.4 Same as the top panel (Ω_b = 33.5 kpc km s^{-1}) in Figure 4.6, but using APOGEE DR16 instead of LAMOST. Note that APOGEE observes more stars towards the Galactic center than LAMOST and populates also Ω_q = 1300km s^{-1} kpc^{-1}. The smaller number of stars (compared to LAMOST) means that it's not possible to draw strong conclusions from the observational data. 141

List of Tables

1.1 Summary of recent, current, and future stellar spectroscopic surveys. Abbreviations are ground-based. For current surveys, “stars observed” refers to the expected number upon completion. Adapted from Buder 2019 and Kollmeier et al. 2017. . . 7

2.1 Catalog schema. Here, x, and X stand for each of the chemical symbols for the elements whose abundances being estimated. Our Galactic coordinate system is right-handed. We also make available the table of per-observation labels. 29

3.1 Catalog schema. Data available at https://doi.org/10.7916/d8-3ap9-qe35 54

5.1 Sources of continuum absorption in DKG. References marked with “*” are via polynomial fits in Gray 2005. Those marked with “†” are via Colgan et al. 2016. When bounds in temperature or wavelength are missing, the absorption coefficient is defined for all positive values. 95

A.1 GALAH line windows for each element 136

Acknowledgements

First of all, thank you to Melissa, for being a great advisor. You remained always-available, encouraging, and insightful, even when on the other side of the world and in enmired in bureaucracy. Thank you for the freedom you gave me as a became a more independent researcher, and for tempering my most self-critical impulses.

Thank you to David Kipping, my rst advisor in the astro department, and to Kathryn Johnston, Marcel Agüeros, and David Schiminovich, who served on my thesis committee.

Thank you to all of my collaborators on the work that went into this thesis. Thanks to the Galah team, especially Sven Buder, for answering my questions and providing useful feedback, and for producing such an important and singular dataset. Thanks to David Hogg, for his seemingly endless enthusiasm and insight. Thanks to Wilma Trick, Francesca Fragkoudi, and Irene Abril Cabezas for their expertise in dynamics, and their patience with the fact that I don't have much. Thank you to Andy Casey for sharing my enthusiasm for scienti c software development, being encouraging and insightful, and caring enough to do thankless infrastructure tasks. An especially big thanks to Matthew Abruzzo, without whom my current research program might not exist, and my code would be much slower and sloppier.

Thanks to Jennifer Johnson and David Weinberg (and everyone at OSU) for their interest in my research and for the vote of con dence in hiring me. Thanks also to Charlie Conroy and Phill Cargile for their generosity with their time, and their help with postdoc applications.

Thanks to all the researchers who have answered my questions via email: Chris Sneden, Samuel Potter, Paul Barklem, Karen Lind, and Thomas Nordlander.

Thank you to all the friends I made at Columbia. Mihir, for many long walks and interesting conversations. Hopefully we will see more of each other in Ohio. Thanks to Aaron for a shared interest in food, and for your inspiring efforts to better the department. Thanks to Matthew, for all of your knowledge and interest about computing, and for providing someone outside my research group with whom to discuss my work (hopefully, I did the same for you). Thanks to Nicole, for giving me an outlet to practice ASL, and for making the department a much more fun place to work. Thank you also to Avery, Moiya, Jorge, Tomer, Aliya, Karen, and all the other grads I overlapped with (and Tom, an honorary grad).

Thanks to my family, especially my parents for caring about my education and for teaching me to value knowledge. Thank you to Steph, for everything.

Chapter 1: Introduction

When we look at a star, we see its outermost layers only. These layers, which contain something like 10^{10} of the star's mass, are called the “atmosphere”, although there is no sharp distinction between atmosphere and not-atmosphere, as there is on a rocky planet. Much of what we know about stars—radial velocities, temperatures, surface gravities, and (centrally to this thesis) compositions—comes from spectroscopic observations of their atmospheres. As discussed below, the elemental abundances of stars are crucial data for understanding physical science across scales from nuclear to cosmological. This thesis is concerned with better learning what stellar atmospheres are telling us (chapters 2, 3, and 5), as well as how we might put that information towards understanding the Milky Way (chapter 4).

1.1 A brief history of early stellar spectroscopy

It's not clear who the rst spectroscopist was, since prisms have been used to split sunlight since ancient times. Newton is sometimes lauded as the originator of the eld, but a stronger contender is Joseph von Frauenhofer. In the early nineteenth century, he was the rst to assemble a spectroscope, and the rst to use a diffraction grating rather than a prism, allowing for the calculation of wavelength. Using his instrument, he observed dark absorption lines and bands in the solar spectrum, to which he assigned single-letter designations, some of which are still used. In 1861, Bunsen and Kirchhoff identi ed these lines with those in emission spectra of the elements, kicking off increased enthusiasm for stellar spectroscopy.[1]

Starting the late nineteenth century, the Harvard Computers examined tens of thousands of spectra. As part of this effort, Annie Jump Cannon, who went on to lead the project, and de-

[1] An interesting bit of trivia is that Lewis Rutherfurd, an important spectroscopist of the era, was a trustee of Columbia and the namesake of Pupin's rooftop observatory.

veloped the OBAFGK system of spectral types. (The OBAFGK system, which categorizes stars into what is now understood to be a temperature sequence, is based on the presence and strength of various absorption features.) The early twentieth century brought the new science of quantum mechanics, putting spectroscopy on rm theoretical ground. Megnad Saha developed a theory of ionization which connected the purely empirical spectral types to atmospheric temperature. Shortly afterward, Cecilia Payne-Gaposchkin's thesis (Payne 1925), further clari ed this correspondence and demonstrated that stars are composed primarily of hydrogen and helium. By the late thirties, the fundamentals of nuclear physics were in place, and stars were known to be a foundry of heavy elements (see below for a discussion of nucleosynthesis). In the 1930s and 1940s, much progress was made on the quantitative modeling of stellar atmospheres, including Subrahmanyan Chandrasekhar's work on H as a source of opacity (Chandrasekhar 1944), and on radiative transfer (e.g. Chandrasekhar 1934; Chandrasekhar 1935; Chandrasekhar 1944). Though much progress has been made in the modern era, it has not yet been neatly mythologized, and I will not attempt to summarize it here.

1.2 Modeling stellar spectra today

1.2.1 Spectral synthesis

The task of starting from a few physical properties (effective temperature, surface gravity, composition, etc) describing a star's atmosphere and predicting the emergent spectrum is commonly referred to as *synthesis*. Since the late 1960s, astronomers have written software to perform this calculation (Sneden 1973b; Kurucz 1993; Plez, Smith, and Lambert 1993; Gray and Corbally 1994; Valenti and Piskunov 1996; Sbordone et al. 2004; Sneden et al. 2012; Plez 2012; Valenti and Piskunov 2012; Piskunov and Valenti 2017), and Chapter xme introduces a new contribution to the genre (and provides more details of physics involved). The two primary inputs to synthesis are the *linelist* and *model atmosphere*. The linelist tabulates the wavelength, oscillator strength,[2]

[2] traditionally recorded as log 6 5 value, where 5 is the oscillator strength and 6 the degeneracy of the lower energy level

and broadening parameters of each transition in the wavelength region of interest. The broadening parameters, which capture how the line pro le changes under pressure, are the most theoretically approximate. The model atmosphere tabulates the properties (density, ionization fraction, pressure, temperature, etc.) of the atmosphere as a function of depth, and is based on 1D hydrostatic equilibrium. (There are exceptions, see below.) It's worth noting that the structure of a star's atmosphere is physically intertwined with the wavelength-resolved ux emerging from it, and that the separation of these two is a matter of computational expediency. There are, in addition, “secret” inputs to spectral synthesis. Though they are typically built in to synthesis codes, they are not known precisely enough for to be neglected as potential sources of error. The partition functions and molecular equilibrium coef cients of each species (sometimes including non-negligible pressure dependence), are not always well-known. As are the wavelength-resolved opacities due to Rayleigh scattering, H, He, H_2 and metals. Also worth mentioning is the fact that numerical techniques differ between spectral synthesis codes, and may introduce bias in unnoticed ways.

Much of the recent work in spectroscopic models has been to relax the assumptions of time-independence, spherical symmetry, and local thermodynamic equilibrium (LTE). Eliminating the former two assumptions means calculating atmospheric structure using the full hydrodynamic equations (e.g. Freytag et al. 2012; Magic, Weiss, and Asplund 2015), or the magnetohydrodynamic equations (e.g. Vögler et al. 2005), if the internal magnetic eld is strong. These are usually done with a "box in a star" domain, in which a small portion of the star is assumed to be representative. Larger "star in a box" simulations have been done, but are more computationally expensive for little gain. As with 1D atmospheres, computing 3D atmospheres is done as a separate step from detailed radiative transfer to produce a model spectrum. The assumption is of LTE is that the kinetic and electron state distributions of small parcels of matter can be described by a single temperature. It is broken when the collision rate is insuf cient to thermalize material in the presence of the (non-thermal) radiation. While LTE often holds true, it has been clear since the work of Thomas 1957 that it sometimes does not. Non-LTE calculations are expensive because they couple the chemical properties of the plasma at all layers of the atmosphere via the radiation

in a nontrivial way. Under LTE, the properties of the plasma can be cleanly separated and addressed one atmospheric layer at a time, with the radiation following from that. Solutions to the spectral synthesis problem not assuming LTE are necessarily iterative, and orders of magnitude slower than the LTE solutions. They also have the disadvantage of requiring more inputs: most notably collisional cross-sections, which are rarely measured in the lab.

Lifting these assumptions is possible, but pushes the limits of current computing power. Some efforts have been made to capture insights from 3D and non-LTE simulations and apply them to faster 1D LTE ones. 3D atmospheres can be approximated by 1D averages, or by the 1.5D approximation, in which only line-of-sight radiation is accounted for (e.g. Amarsi et al. 2016). Likewise, non-LTE electron populations can be approximated with departure coefficients calibrated with non-LTE simulations. These are exciting ways to produce more accurate abundances at the scale required by modern surveys.

1.2.2 Data-driven spectral models

In recent years, statistical and machine learning methods (called *data-driven* hereafter) have been ascendant in spectroscopic analysis. This was kicked off by THE CANNON (Ness et al. 2015; named for Annie Jump Cannon), which approximates the flux at a given wavelength as a quadratic function of stellar parameters and abundances. These methods all employ a flexible model without explicit physics, but they vary in nearly every other possible way. A principal axis of differentiation is the flexibility of the model: compared to a neural network (there are several employing them), the quadratic behavior of THE CANNON is rigid.

Second, most methods are supervised (chapter 3 discusses counterexamples), meaning that they are calibrated with spectra associated with *a priori* correct parameters and abundances, a *training set*. Within this category, we can distinguish between models trained on real observations (e.g. Casey et al. 2016) and those trained on synthetic data (e.g. Ting et al. 2019). Real observations are affected by measurement error and the training abundances and parameters are inexact. Synthetic data may be partially unrealistic, but they can be generated for arbitrary input parameters and

abundances, meaning that training data are as plentiful as necessary in addition to being perfectly precise. On the other hand, models trained on only synthetic data offer little theoretical advantage over the direct use of physical models. (This is not to discount their computational advantages.) Indeed, codes which interpolate a large grid of model spectra (e.g. Recio-Blanco, Bijaoui, and De Laverny 2006; Allende-Prieto and Apogee Team 2015, though there are many others) arguably fall into this category. Models which make use of both observational and synthetic data (e.g. Xiang et al. 2019; O'Briain et al. 2020) attempt to have the best of both worlds.

All of the methods cited so far are generative ("forward", as astronomers say), meaning that they map stellar parameters and abundances to spectra, not the other way around. To estimate the parameters and abundances of a star, we combine these methods with statistical inference. Maximum likelihood or maximum *a posteriori* estimation is nearly always used, though in principal more sophisticated approaches are entirely possible. While this scheme is potentially more reliable, the necessity of inference make generative models slower than discriminative ones, which map spectra to parameters and abundances directly. All discriminative models (e.g. (Xiang et al. 2017; Leung and Bovy 2019; Olney et al. 2020)) are trained on observational data, as they are more sensitive instrument peculiarities and must be calibrated with data as similar as possible to the data on which they will be deployed (all data-driven models share this characteristic, but it is strongest with this subset).

One of the chief virtues of data-driven models is speed. This is an important advantage now that industrial surveys supply a unwieldy quantity of data. But data-driven models trained on observational data can also—in principle though not always in practice—bootstrap themselves to physics missing from traditional models. An obvious way in which this happens is that data-driven methods can use training data (abundances and parameters) calculated from well-understood lines to empirically construct a model of all wavelengths observed. In application to new spectra, this enables them to infer more precise parameters and abundances at a given signal-to-noise ratio (S/N), and to work at all in application to low S/N data. The downside is that data-driven models can easily "learn" incorrect physics. Of course if the training data are biased, data-driven models

will inherit those biases (for example see section 2.5.1), though note there are some efforts to ameliorate problems like this (e.g. Eilers et al. 2022). Worse, if the training data are too noisy, too few, or lacking in part of parameter space, data-driven models can easily give rise to new unphysical behavior. Another downside of data-driven models (though one which may be overstated) is that their behavior may not be straightforward to understand. Chapter 3 is an example of an application where this presents a challenge to interpretation.

1.2.3 Challenges in modelling stellar spectra

While it's difficult to overstate the impact of spectroscopically-derived quantities on astronomy, they are often affected by serious biases. Unaccounted-for physics (e.g. magnetic fields, see Spina et al. 2020) affects a subset of stars, and uncertainties in input parameters affect them all. Jofré et al. 2015 compared analyses done on the same high-quality, high-S/N data, and found that inferred abundances vary by 0.03 – 0.4 dex[3], a serious problem when the abundances of most element range 1 dex or less for disk stars. Efforts to improve accuracy by careful selection of input parameters or by differential analysis, reduce this to the roughly 0.03–0.15 dex level (see e.g. Jofré et al. 2015; Buder et al. 2021), but further improvement is still necessary. This problem is endemic enough that different surveys are frequently referred to as having different "abundance scales" due to their differing analysis pipelines. Consequentially, it's crucial that abundances used in calculations be derived as similarly as possible, and researchers are often forced to constrain their analyses to a small part of the $T_{\rm eff}$ – $\log g$ plane, with the hope that systematics will be largely correlated with effective temperature and surface gravity. To make matters worse, the recent consensus is that, after accounting for a few degrees of freedom (corresponding to enrichment from type Ia supernovae, core-collapse supernovae, and asymptotic giant branch stars), the remaining variations in detailed abundances are at the 0.02 dex level (Weinberg et al. 2019; Ting and Weinberg 2022; Ness et al. 2022), well below the potential model bias and close to the noise floor of large surveys. There is

[3] The dex is a nonunit attached to the logarithm of a unitless number by astronomers. In the case of elemental abundances, the unitless number is a ratio of number densities. A difference of 0.4 dex is a factor of 2.5 in the amount of an element. A delta of 0.05 dex is equivalent to 12% difference.

reason to believe that reduction from images to 1D spectra, ux as a function of wavelength, also poses a problem. For example, in APOGEE, the scatter of derived abundances across different observations is signi cantly larger than predicted from measurement uncertainty (Jönsson et al. 2020). Starting to address these problem is what motivates my current work, but is not the main subject of this thesis.

1.3 Recent, current, and future stellar spectroscopic surveys

survey	stars observed	wavelengths [nm]	resolving power
RAVE	$4.5 \quad 10^5$	841 - 879.5	7500
SEGUE	$3.5 \quad 10^5$	385 - 920	2000
APOGEE	5^5	1510 - 1700	22,000
LAMOST	$9 \quad 10^6$	400 - 900	1800
Gaia-ESO	10^5	target-dependent (visible)	17,000
GALAH	$8 \quad 10^5$	400 - 850	28,000
4MOST	10^7	400 - 900	target-dependent
MWM	$7 \quad 10^6$	1510 - 1700	22,000
Gaia DR3 (RVS)	10^6	845 - 872	11,500
Gaia DR3 (XP)	$2.2 \quad 10^8$	350 - 1050	70
MOONS	$2 \quad 10^6$	600 - 1800	20,000
4MOST	10^7	400 - 900	5000 ; 20,000

Table 1.1 Summary of recent, current, and future stellar spectroscopic surveys. All but Gaia are ground-based. For current surveys, “stars observed” refers to the expected number upon completion. Adapted from Buder 2019 and Kollmeier et al. 2017.

There are several surveys relevant to this thesis; Table 1.1 presents an overview of 10^7, the observations (not counting the Gaia XP spectra). Large (here meaning 10^5 stars) stellar spectroscopic surveys were made possible by advances in multi-object ber-fed spectroscopy, pioneered at the Anglo-Australian Observatory (AAO) and Steward observatory (see Hill 1988 for a review). The technology was driven by the desire to observe more galaxies, but it was quickly applied to individual stars. Interestingly, both manually-placed bers and robotic positioners have coexisted in major observing programs for decades, with the Sloan Digital Sky Survey (SDSS; York et al. 2000) as the primary user of hand-plugged plates. As of June 2022, SDSS has almost completed its transition to robotic positioners. Gaia-ESO (Gilmore et al. 2012; not af liated with Gaia, but named

for it's complementarity) is notable for it's commitment to analysis pipeline pluralism. GALAH (GALactic Archaeology with HERMES; De Silva et al. 2015), was conceived of as a chemical tagging project (discussed below), and thus observes at the highest resolution in order to measure the abundances of up to 30 elements. LAMOST (Large Sky Area Multi-Object Fiber Spectroscopic Telescope; Cui et al. 2012; Deng et al. 2012) is the current record holder for most targets observed (excluding the *Gaia* BP/RP spectra). By producing the largest catalog of low-resolution spectra, LAMOST has motivated much innovation in data-driven spectroscopic models.

As I write this, *Gaia*'s third data release Gaia Collaboration et al. 2016; Gaia Collaboration et al. 2016, bringing spectra from its radial velocity spectrograph (RVS), is hot off the presses.[4] Like RAVE, the RVS observes the calcium triplet only. Though so low-resolution as to not qualify as spectra in the traditional sense, the *Gaia* XP spectra are so incredibly numerous that they merit mention as well. Just now ramping up, Milky Way Mapper (MWM; part of SDSS-V; Kollmeier et al. 2017) will observe nearly 10^7 stars at the same resolution as APOGEE. This is being achieved by 1) the switch to a robotic positioning system, and 2) shorter observations resulting in noisier spectra (motivated largely by the success of data-driven modes on low signal-to-noise spectra).

Two upcoming instruments are also worth mentioning. The MOONS (Multi-Object Optical and Near-infrared Spectrograph; Cirasuolo et al. 2011) will operate in the infrared, executing a survey at larger magnitudes than APOGEE or MWM, providing an unprecedented view of the Milky Way's bulge. 4MOST (the 4-metre Multi-Object Spectroscopic Telescope; de Jong et al. 2019) will run five stellar surveys in the halo, disk, and bulge and will include both low- and high-res observations.

1.4 Applications of spectroscopically-derived quantities

1.4.1 Nucleosynthesis

Among the most important applications of stellar spectroscopy is the observational constraint of nucleosynthetic models. The modern understanding of astrophysical nucleosynthesis owes an

[4] If this were not a metaphor, *Gaia* DR3 would require roughly 10^8 pages of newsprint.

enormous debt to the well known B^{2}FH paper (Burbidge et al. 1957), and in many ways, the general picture has not changed much since then. Big Bang nucleosynthesis (the coalescence of nuclei as the early universe cooled) is the source of nearly all hydrogen and helium, and much of the present-day supply of lithium (the third element is an interesting case–chapter 3 discusses other production mechanisms). Beryllium and boron (atomic number Z = 4–5) are formed primarily through cosmic ray spallation (induced fission) of material in the interstellar medium (ISM; a bit of lithium is also formed this way). The elements with $Z > 5$ are produced by mechanisms which can be divided into two categories: charged-particle reactions and neutron-capture reactions. Charged-particle reactions require the extreme temperatures and pressures of stellar interiors to overcome the Coulomb barrier between particles. These are the reactions which power stars. Neutron-capture reactions can happen at lower temperatures but require a flux of free neutrons, which are unstable.

The most important progenitors of the elements produced in charged-particle reactions are core-collapse supernovae (CCSNe) and type Ia supernovae (SNe Ia). CCSNe release every element from carbon (Z = 6) to zircon (Z = 40) into the ISM. Of these, the alpha elements, Ne, Mg, Si, S, Ar, Ca, and Ti, are produced by the agglomeration of helium nuclei (alpha particles), and are well-correlated with each other and particularly abundant. SNe Ia release a smaller range of elements, from silicon (Z = 14) to zinc (Z = 30), with the strongest contribution being to elements heavier than scandium (Z = 21), the iron-peak elements. The light elements (carbon, nitrogen, and oxygen) necessitate a complication of this simple picture. They are released by winds from AGB stars and massive stars, in addition to CCSNe.

Elements heavier than rubidium (Z = 37) are formed via neutron capture, the growth of nuclei one neutron at a time. Neutron capture can be divided into the S (slow)-process, wherein the timescale of neutron bombardment is slower than that of beta decay, and the R (rapid)-process, wherein it is faster. The S-process and R-process produce the elements in different amounts. The former is understood to occur primarily in asymptotic giant branch (AGB) stars, but the site of the R-process is less certain. Follow-up of the GW170817 neutron star merger (Abbott et al. 2017) detected signatures of the R-process, and the enrichment pattern of Reticulum II (Ji et al. 2016;

Roederer et al. 2016) substantiates the paradigm of NSMs as the primary enrichment site. Other sites are not ruled out, however. Magnetorotational core-collapse supernovae and accretion disk out ows from collapsars are two of the most promising, and emerging evidence from elemental abundances suggests that they play an important role (see Cowan et al. 2021 or Kajino et al. 2019 for a review).

There are many gaps in our understanding of nucleosynthesis. For elements with $29 \leq Z \leq 50$, contributions are expected from both charged-particle reactions and neutron capture. These are the elements whose origins are most uncertain. Weinberg et al. 2019 empirically characterized the yields of CCSNe and SNe Ia, and found a variety of disagreements with theoretical predictions, even for relatively well-understood elements. It may be that systematics in abundances are a major reason for the tension. Theoretical predictions of detailed yields remain largely observationally uncon rmed regardless. More detailed models incorporating yields depending on progenitor mass or metallicity are even more dif cult to observationally constrain, and disentangling the effects of AGB stars from SNe remains dif cult (Grif th et al. 2021).

1.4.2 Galactic archaeology

As noted in the rst paragraph of this thesis, when we measure stellar composition spectroscopically, we are measuring the make-up of the atmosphere, not the whole star. While the composition of the stellar interior changes due to nuclear burning, that of the atmosphere remains relatively constant. This useful fact leads directly to the idea of chemical tagging rst put forward by Freeman and Bland-Hawthorn 2002, that the open clusters in which all disk stars were presumably born might be reconstructed by clustering in abundance space. Though this vision catalyzed the collection of much data (most notably the GALAH survey), there have been a few impediments to the success of the program.

First, variations in detailed elemental abundances are not as high-dimensional as anticipated. Second, there are problems with spectroscopic models, discussed above. The last major barrier is that natal clusters may not be suf ciently homogeneous to neatly cluster in chemical space, even

if we could measure abundances with perfect precision and accuracy. Studies of atomic diffusion, the strati cation across depth of different species, in open clusters indicate that it may modify atmospheric abundances at the 0.1 dex level (e.g. Bertelli Motta et al. 2018; Gao et al. 2018; Souto et al. 2019; Liu et al. 2019). Other mixing mechanisms (see Pinsonneault 1997 for a review) may be equally important. While problems with spectroscopic models and atomic diffusion can in principle be overcome by more sophisticated models, the low dimensionality of chemical space may prove to be a fundamental barrier. Blanco-Cuaresma et al. 2015 examined abundances in 31 open clusters and found a within-cluster scatter and between-cluster similarity which indicated that chemical tagging is infeasible. Reconstructing a small subset of natal clusters is nevertheless useful, and may be possible.Prospects of reconstructing accreted populations in the halo, where natal groups are fewer, less mixed, and more chemically varied, are much better.

Besides chemical tagging, there are many other applications of elemental abundances to the understanding of the Milky Way. Chapter 4 is an example of stellar chemistry in the aid of Galactic dynamics, wherein we take advantage of the fact that abundances roughly trace birth radius to identify resonances of the Galactic bar. There are too many works in this category to provide an complete overview, but one important example is the case of the high- and low-U sequences. Gilmore and Reid 1983 were to the rst to deduce that the Galactic disk is make up of two distinct populations. At the time, these were kinematically identi ed: the "thick disk" and "thin disks". It wasn't until roughly 15 years later that Fuhrmann 1998 was able to deduce that the two population were chemically distinct, with different levels of alpha-element enrichment. Much work has gone into further elucidating things, and abundances have allowed us to separate two dynamically-entangled populations more clearly. Notably, Bensby et al. 2011 found that the two populations formed disks with different scale lengths (that is, they dropped off at different rates with distance from the center of the Galaxy). This picture was con rmed by APOGEE (Hayden et al. 2015), which also provided a clearer picture of the z-dependence (thickness) of the two disks as a function of Galactocentric radius Bovy et al. 2016.

1.4.3 Stellar physics

Finally, there is the application of stellar spectroscopy in the understanding of stars themselves. As mentioned already, changes in atmospheric abundances (e.g. of lithium) are useful probes of stellar interiors, as are changes due to various mixing mechanisms. Stellar ages, which are notoriously dif cult to measure, can be constrained in various way by abundances (not to mention that magnetic activity is spectrally diagnosed and related to youth). For example, the abundances of unstable but long-lived r-process isotopes ^{238}U and ^{232}Th, if measurable, indicate age contingent on an initial abundance (see Soderblom 2010 for a review). Martig et al. 2016 and Martig et al. 2016 showed that for red giants, [C/N] can be used to infer mass (and thus age), via mass-dependent dredge-up. There is a much ongoing work linking ages to other abundance variations (Bedell et al. 2018; Hayden et al. 2020; Nissen et al. 2020; Casali et al. 2020). Other spectroscopically derived quantities, e.g. effective temperature, surface gravity, and activity indicators are also valuable, and stellar atmospheres are of course interesting objects of study in their own right.

Chapter 2: Abundances in the Milky Way across five nucleosynthetic channels from 4 million LAMOST stars

2.1 Introduction

Large stellar surveys such as Gaia (Gaia Collaboration et al. 2016), APOGEE (Majewski et al. 2017; Holtzman et al. 2018; García Pérez et al. 2016), GALAH (De Silva et al. 2015; Martell et al. 2017), Gaia-ESO (Gilmore et al. 2012), RAVE (Steinmetz et al. 2006), LAMOST (Deng et al. 2012; Zhao et al. 2012) and SEGUE (Yanny et al. 2009) are providing the data to empirically characterize the Milky Way disk and infer the primary drivers of its formation and evolution (Freeman and Bland-Hawthorn 2002; Bland-Hawthorn and Gerhard 2016).

Detailed chemical abundances are one of the primary measurements made from stellar spectra. Their determination is a primary motivation for medium- and high-resolution spectroscopic surveys for several reasons: they provide effective chemical fingerprints of stars, link directly to the environment in which they were born, and describe the chemical diversity of the disk (e.g. Weinberg et al. 2019) and the chemical pathways of enrichment (e.g. Rybizki et al. 2020). Combined with stellar kinematics, abundances are core to the pursuit of Galactic archaeology.

Conventionally, detailed abundances have been derived from medium- and high- resolution stellar spectra (e.g. APOGEE: $R = 22{,}500$, GALAH: $R = 28{,}000$, and RAVE: $R = 7500$, Gaia-ESO: at least $R = 20{,}000$). Until recently, the inferences from low-resolution spectra, such as LAMOST and SEGUE, were typically limited to stellar parameters and enhancements ($T_{\rm eff}$, $\log g$, [Fe/H], [α/Fe]) (e.g. Lee et al. 2011). Ting et al. 2018 have shown that oxygen abundances can be inferred from spectra in wavelength regions containing no atomic oxygen lines through the features of species in the CNO atomic-molecular network. Indirectly inferred abundances also have a long empirical history. The Ca II triplet, for example, is an often used metallicity index

(Armandroff and Zinn 1988; see Vásquez et al. 2015 for a recent calibration). To date, with the exception of Xiang et al. 2019, the efforts to extract individual abundances from LAMOST have largely focused on a few elements, namely an integrated-element abundance and the elements C and N, which are particularly important, as these elements can indicate age (e.g. Li et al. 2016; Xiang et al. 2017; Ho et al. 2017b; Ho et al. 2018; Zhang et al. 2019).

In this work, we use a data-driven approach to label low-resolution LAMOST spectra with several abundances. LAMOST is one of the largest stellar surveys to date, with over 10^6 publicly available spectra, at $R = 1800$. The survey has extensive coverage of the Milky Way's disk, halo and, in particular, the outer disk, the detailed chemodynamics of which are largely unexplored. Specifically, we employ *The Cannon* (Ness et al. 2015), a model characterized in large part by its simplicity, to derive individual abundances from LAMOST. Other data-driven methods include *The Payne* (Ting et al. 2019), which, like *The Cannon*, works by explicitly modelling spectra as a function of labels (stellar parameters and abundances), and that of Leung and Bovy 2019, which uses a convolutional neural network to estimate labels directly from spectra without explicit inference. Xiang et al. 2019 recently released a catalog of 16 abundances (C, N, O, Na, Mg, Al, Si, Ca, Ti, Cr, Mn, Fe, Co, Ni, Cu, and Ba) for LAMOST DR 5 using a neural-net-based model calibrated by both labelled spectra (using overlap between LAMOST and both APOGEE and GALAH) and physical modelling. This work has many common aspects with our own, but is different in detail. Both calibrate flexible spectral models (a shallow neural network, in the case of Xiang et al. 2019) with labels from high-resolution surveys, but Xiang et al. 2019 also employ gradients of *ab-initio* models. An advantage of using model gradients is that physical expectations are incorporated into the label derivation. Our approach, however, prioritises the data alone in specifying the model, which can be advantageous when physical models are lacking. Differences between the catalogues for those elements trained using the GALAH labels will help reveal the biases of each approach.

Our approach requires reference objects, stars with high-quality spectra and precise labels (stellar parameters and abundances), that are representative of the survey objects. They are used to

calibrate a model that produces synthetic spectra from stellar labels. This model is then used to estimate labels for the full set of survey stars, in our case, the LAMOST catalog. Both the APOGEE and GALAH surveys have stars in common with LAMOST which can serve as possible reference objects. APOGEE provides higher precision abundance measurements than GALAH, which enables, for example, the clear disambiguation of the the low- and high-α sequences, as seen in the radial maps of Hayden et al. 2015 and Nidever et al. 2014. However, the dimensionality of the abundance space measured by APOGEE is low (Ness et al. 2018; Price-Jones and Bovy 2018; Ness et al. 2019) (although note that weak lines of neutron capture elements have been identified in this region (Cunha et al. 2017; Hasselquist et al. 2016)). GALAH, on the other hand, provides abundance measurements across a more extensive set of nucleosynthetic channels, including the neutron-capture A (and B) processes. The neutron-capture element enhancements have been previously explored only through boutique analyses of small samples of stars observed at high resolution (e.g. Bensby, Feltzing, and Oey 2014; Spina et al. 2017) and in the solar neighbourhood, to which GALAH is largely confined (e.g. Buder et al. 2019; Schönrich and Weinberg 2019). GALAH also provides abundances for main-sequence stars, allowing us to extend our modelling to that regime.

We want to explore the promise of the largest number of element abundance families as possible, so we took the roughly 10,000 stars in common between GALAH and LAMOST to build a model using the LAMOST spectra and GALAH stellar parameters and abundances. While the GALAH labels are less precise than those from APOGEE and thus yield less precise LAMOST labels, the LAMOST catalog is large enough to enable very precise mean estimates of abundances on a population basis (e.g Ness et al. 2019; Blancato et al. 2019). Using GALAH as a source for our input labels allows us to propagate A-process and B-process abundances to the outer disk and halo.

In deriving a set of individual abundances for LAMOST, this work complements the LAMOST catalogue, which provides stellar parameters and bulk metallicity (a term used interchangeably with [Fe/H] in this work) only. We deliver inferred abundances for elements from five nucleosynthetic families: light elements, which are dispersed by asymptotic giant branch (AGB) stars and

core-collapse supernovae (CCSN), and whose atmospheric abundances can change due to dredge-up; U-elements, which are dispersed primarily by CCSN; Sc, iron-peakelements, which are dispersed by both CCSN and type Ia supernovae (SNIa); odd-Zelements, which are dispersed by both CCSN and SNIa and expected to display similar trends to U-elements; B-process elements, which are thought to be produced and dispersed in AGB stars; and A-process elements, which are produced in extremely neutron-rich environments. It is not clear at present whether neutron-star mergers are the primary site of the A-process, or if other sites make appreciable contributions (e.g. Arnould, Goriely, and Takahashi 2007; Côté et al. 2018; Hansen et al. 2018; Siegel 2019; Sakari et al. 2018; Sakari 2019). For each star, we deliver five (for dwarfs) or six (for giants) abundances of O (light), Eu (A-process), mean U, Sc (iron-peak), mean B-process, Mg (U), Al (odd /), Mn (iron-peak), and Ba (B-process). Having derived these abundances, we demonstrate the scientific value of multi-element abundances of large numbers of stars. We do this using pairs of stars across the disk and halo, examining the abundance similarity of wide binaries, that have been identified by their kinematics alone. We also map the chemodynamical abundance structure of the disk and halo, making links to signatures of evolution such as radial migration and Galaxy assembly.

In §2 we describe the GALAH and LAMOST data and the quality cuts we applied. §3 provides a brief overview of The Cannon. In §4 we discuss model checks and evaluate the error of our label estimates. §5 discusses our public catalog and key scientific results, and §6 discusses their implications.

2.2 Data

Our data comprise the R=1800 DR 4 v2 LAMOST spectra, the R=28,000 DR 2.1 GALAH spectra and stellar parameter and abundance labels (Buder et al. 2018), as well as the Gaia proper motion and parallax measurements for our stars. From GALAH we use T_{eff}, $\log g$, v_{mic}, and [Fe/H], along with abundances with respect to Fe, of O, Si, Ca, Ti, Eu, Sc, Y, Mg, Al, Mn, and Ba. Figure 2.1 shows the Galactic footprints of GALAH and LAMOST. A portion of each survey's spectrum for a typical training set star is shown in Figure 2.2.

Figure 2.1 Face on (top) and edge on (bottom) Contours in surface density (5×10^{-4}, 3×10^{-2}, 1×10^{-1} kpc^{-2}) shown in heliocentric Galactic coordinates for GALAH and LAMOST, which probes much farther into the outer disk and halo. The Galactic center is at $x = 8$ kpc, $y = 0$.

Figure 2.2 A comparison of the GALAH and LAMOST spectra for the same star, 2MASS J00010184+0407201, Gaia DR2 2740354684364096000. The top panels show part of the star's GALAH spectrum (S/N = 65) and Cannon model, while the bottom two show the same for LAMOST (S/N = 179). On the left, note the large v_V line and surrounding features, on the right, observe the fit around a known Mn feature (highlighted in red on the left). Both the measured spectra and Cannon models are shown with their $\pm 1\sigma$-errors, shown with error bars in the right-hand panels and a shaded region in the left-hand panels.

2.2.1 Quality cuts and data cleaning

One of the formal assumptions of The Cannon is that the training labels are known exactly, so constructing a high-fidelity training set is crucial. To build our training set, we first determined the set of stars in common between GALAH and LAMOST. We performed a 1.00″ sky match between GALAH DR 2.1 and LAMOST DR 4 v2 to identify these reference object candidates, of which there were roughly ten thousand. We then removed all stars from the potential training set with signal to noise ratio (S/N) less than 30 in either the LAMOST I band (snr_z) or the GALAH blue channel (snr_c1). We also removed any star for which chi2_cannon (a column in the GALAH catalog, not a product of our analysis) was greater than 4, which indicates that the best-fit spectral model is a poor fit to the whole spectrum, and any star for which flag_cannon was nonzero, which can indicate a variety of problems with abundance determination. These cuts removed roughly half of the stars from consideration.

We found that cutting on the reported GALAH label errors did not improve our performance against the validation set. To further exclude low-quality measurements from our training set, we therefore generated and evaluated the fit of the best-fit Cannon model spectrum for each reference stellar spectrum, for every element in the GALAH catalog. The GALAH pipeline uses separate Cannon models for each elemental abundance in order to restrict each model to the wavelengths of unblended lines. Each model has different best-fit parameters, which we were not able to retrieve. They are, however, within the errors of the mean reported stellar parameters for each star (Buder et al. 2018). For the stellar parameter labels, we used the values in the GALAH DR2.1 catalog[1], along with E(J−K) values calculated with the Rayleigh-Jeans color excess (RJCE) method (Majewski et al. 2017) applied to ALLWISE (Wright et al. 2010; Mainzer et al. 2011) and 2MASS (Skrutskie et al. 2006) broadband photometry, as was done for the GALAH models. We calculated χ^2 between the best-fit GALAH model and the observed GALAH spectrum for every star in our training set in the region of the strongest lines of each element (the chi2_cannon flag pertains to the global fit). Appendix A.3 lists the wavelength regions used, which are the same windows used in the

[1] available at https://docs.datacentral.org.au/galah/ .

Figure 2.3 Our 1722 training stars in the LAMOST Kiel diagram space colored by (GALAH) [Fe/H]. Since there are more metal-poor stars in the giant training set, our giant model is unbiased down to lower metallicity.

GALAH pipeline. The distribution of χ^2 values for some elements peaked lower than expected from nominal measurement error alone by a factor of 2-3, meaning that a cut on some multiple of $\chi^2 \cdot$dof was not theoretically justified. We removed all stars with χ^2 values above the 85th percentile, for any of its abundances. This led to a significant improvement in our cross-validation results, as discussed in our methods, on the order of 15%-40% percent). Using the 75th percentile, as a more conservative cut, gave us no improvement in cross-validation tests. These cuts leave 1722 stars in the training set. We do not exclude stars flagged in GALAH based on flag_x_fe because we performed our own per-abundance χ^2 cut and because removing stars where the GALAH model may be extrapolating reduces the size of our training set too drastically. We textitasize however that The Cannon is likely to extrapolate well (in a well understood way using the simple polynomial model we employ) for many abundances.

2.2.2 Dwarf and giant models

After the quality cuts described in Section 2.1, we were left with a training set that spans the Kiel diagram and metallicity (Figures 2.3, 2.4). We modeled giants and dwarfs separately, with the

Figure 2.4 Our 1722 training stars across the (LAMOST) Kiel diagram in bins of metallicity. For the extreme ends of our metallicity range, the lack of training set coverage may bias our estimated labels.

division between models given in terms of LAMOST $\log g$ and T_{eff} by

$$\log g = \begin{cases} 4.18 & T_{\mathrm{eff}} > 5200\,\mathrm{K} \\ 1.6 \times 10^{-4} T_{\mathrm{eff}} + 7.3 & T_{\mathrm{eff}} \leq 5200\,\mathrm{K} \end{cases} \tag{2.1}$$

The split gives us 532 giants and 1190 dwarfs as our reference objects. The values in Equation (2.1) to separate dwarfs and giants are somewhat arbitrary. We find that the performance of each model is not not sensitive to these precise values. We decided which elements to infer for each model by balancing our ability to recover each abundance in cross-validation (§2.4) with the objective of having several elements across nucleosynthetic channel. For both models, we include T_{eff}, $\log g$, E_{mic}, [Fe/H], [O/Fe], and [Eu/Fe] as labels. For the dwarfs, we also used iron-relative abundances of: error-weighted mean α (from Mg, Si, Ca, and Ti), Sc, and error-weighted mean s-process (from Ba and Y), while for the giants we also used Mg, Al (odd-Z), Mn (iron-peak), and Ba (s-process). Training a model without O or Mg yields a systematic offset in inferred neutron-capture abundances, likely because the model will exploit correlations between these nucleosynthetic families if they are not controlled for.

We only used mean abundances in the same nucleosynthetic family if they appeared strongly correlated in the training set. We tried using dereddened Gaia G band magnitude instead of $\log g$, which would allow us to apply a prior at test time, but we found that this did not improve our

Figure 2.5 Mean fractional distance error as a function of Galactocentric radius. Distance errors blur chemical maps of the Milky Way far from the solar annulus.

results in practice. The model had trouble predicting extinction, partially because our training sets do not include any high-extinction stars. Including extinction as a label did not improve our ability to predict any of the abundances, so we opted not to.

For subsequent analysis, we cross-matched with Gaia by taking the source within 1 arcsecond of the LAMOST star with the lowest G-band magnitude. Throughout this paper, we use the parallactic distance estimates from Bailer-Jones et al. 2018, which makes use of a prior incorporating the expected galactic spatial distribution. Other distance catalogs (e.g. Anders et al. 2019) will have different biases for stars with uncertain distances (e.g. those far from the Sun), so quantitative results derived from our catalog will be conditioned upon the assumptions of Bailer-Jones et al. 2018. Our results below are largely qualitative, and unlikely to be strongly dependant on choice of distance catalog. Figure 2.5 shows the mean fractional distance error as a function of Galactocentric radius. Distance errors blur our maps of chemistry across the Galaxy (Section 2.5.3), particularly far from the solar annulus, where they reach 20%.

2.3 Model

The following is a brief description of The Cannon (see Ness et al. 2015, for a more extended discussion.). For this work we build a JULIA-based implementation of The Cannon, which is

documented and available at this URL[2] and via the JULIA package manager. The source code uses the same nomenclature as the description here, and allows for optional masking of labels (self-consistent training with the model constrained so that each label is only "on" at specified wavelengths).

For each star n, we take the flux value in the spectral pixel with wavelength λ to be $f_{n,\lambda}$ and its (Gaussian, independent) measurement uncertainty to be $\sigma_{n,\lambda}$. To prepare the spectra for The Cannon we first redshift-corrected the spectra using the value provided in the LAMOST data table and interpolated each star to a common wavelength grid. We then continuum-normalized the spectra by dividing out the continuum, approximated by smoothing the spectra with a Gaussian kernel with a 50 Å standard deviation, truncated at 150 Å from the center, in the same manner as Ho et al. 2017a; Ho et al. 2018. This normalized flux is then near unity in the absence of emission or absorption features. For each reference star, we also define the vector containing its physical parameters and abundances (labels) ℓ_n. These are the quantities we ultimately wish to infer for the rest of the LAMOST spectra, at test time.

Our labels vector ℓ_n for each reference stars is:

$$\ell_n = \left[T_{\rm eff} \quad \log g \quad v_{\rm mic} \quad [{\rm Fe/H}] \quad [X_1/{\rm Fe}] \quad \cdots \quad [X_N/{\rm Fe}] \right] \tag{2.2}$$

where $X_1 \ldots X_N$ are the elements whose abundances we wish to determine. It is good practice for both numerical stability and model flexibility to express all labels in units such that they are distributed around zero and have similar magnitudes. We do this by subtracting from each label its (training set) mean and dividing it by its (training set) dispersion. This transformation is then undone after the inference has taken place.

In numerous published uses of The Cannon (including this one), the flux in each pixel is described by a 2nd degree polynomial of the elements of the label vector whose coefficients, θ_λ, are determined by a training set of spectra for which, ideally, both accurate and precise labels are available. For a given spectral pixel and star, we then have our spectral flux, $f_{n,\lambda}$, for our reference

[2] github.com/ajwheeler/TheCannon.jl

objects at each wavelength, λ, defined as:

$$
\begin{aligned}
\mathbf{y}_\lambda &= \mathbf{w}_{\lambda 0} && \text{(constant term)} \\
&+ \mathbf{w}_{\lambda 1} \ell_1 + \cdots + \mathbf{w}_{\lambda \mathbf{y}} \ell_{\mathbf{y}} && \text{(linear terms)} \\
&+ \mathbf{w}_{\lambda 2} \ell_1^2 + \cdots + \mathbf{w}_{\lambda \mathbf{y}+1} \ell_{\mathbf{y}}^2 && \text{(squared terms)} \\
&+ \mathbf{w}_{\lambda 3} \log(\ell_1) + \cdots \\
&+ \mathbf{w}_{\lambda \mathbf{y}+2} \ell_1 \ell_2 + \cdots + \mathbf{w}_{\lambda \mathbf{y}+\mathbf{y}-1} \ell_{\mathbf{y}-1} \ell_{\mathbf{y}} && \text{(cross-terms)} \\
&+ \text{error}.
\end{aligned}
$$

To specify an error model, we can write the above as a likelihood function

$$
\mathbf{y}_\lambda \mid \mathbf{f}(\mathbf{x}_\lambda, \mathbf{w}_\lambda) \sim \mathcal{N}(\mathbf{f}(\mathbf{x}_\lambda, \mathbf{w}_\lambda), \sigma_\lambda^2) \tag{2.3}
$$

where $\mathcal{N}$ is the normal distribution, σ_λ is model uncertainty (either inherent stochasticity or physics that hasn't been captured by the model) at wavelength λ, $\mathbf{w}_\lambda$ is the vector of coefficients describing how the flux at λ varies with label value, and $\mathbf{f}$, the quadratic expansion (called the vectorizing function in case study 1.6), maps from labels to every 0th, 1st, and 2nd order combination of components of the label vector,

$$
\mathbf{f}(\mathbf{x}_\lambda) = \left[1, \ldots, \ell_1, \ldots, \ell_{\mathbf{y}}, \ell_1^2, \ldots, \ell_{\mathbf{y}}^2, \log(\ell_1), \ldots, \ldots\right]. \tag{2.4}
$$

If $\mathbf{x}_\lambda$ is a vector of length $\mathbf{y}$, $\mathbf{f}(\mathbf{x}_\lambda)$ is a vector of length $1 + \mathbf{y} + \mathbf{y}^2 + 3\mathbf{y} + \mathbf{y}(\mathbf{y}-1)/2$. A more flexible model could be constructed by replacing $\mathbf{f}$ to an expansion with higher order terms, or to other combinations of labels. However, quadratic models have been shown to be sufficient in practice (Ness et al. 2015; Ness et al. 2016; Ness et al. 2019; Ho et al. 2017a; Ho et al. 2018). In fact, a linear model is often

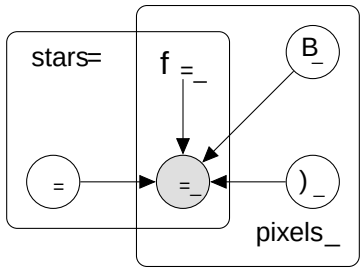

Figure 2.6 The Cannon likelihood as a probabilistic graphical model. During training, the latent variables in the "stars" panel ( ) are fixed using stars for which labels are known, in our case from GALAH. When inferring stellar labels, the latent variables in the "pixels" panel (and) are fixed to their point estimates from training and the maximum likelihood estimates for are calculated.

all that is needed (Birky et al. 2020; Hogg, Eilers, and Rix 2018). The combinatoric increase in model parameters that would be necessary for a higher order polynomial is undesirable.

Figure 2.6 shows the likelihood function as a probabilistic graphical model, which depicts the relationships between observed and latent quantities. Ideally, the full joint distribution over training data and output labels would be sampled from directly (with e.g. Markov-chain Monte-Carlo), but such an approach is not computationally feasible. Instead the problem is divided into a training step, in which a point estimate of each and is estimated from the training set, and an inference step, in which the labels of each star are estimated.

During the training step, each and is jointly fixed to its maximum-likelihood estimate (MLE), given the labeled spectra in the training set. For fixed the model is linear with fixed Gaussian error in , so its MLE, , can be calculated analytically. Finding is then a matter of numerically maximizing the log-likelihood,

$$\log L \left(\ldots \right) = \sum \left[\frac{1}{2} \frac{\left(\ldots \right)^2}{\ldots} - \ln \left(\ldots \right) - \text{const} \right] \tag{2.5}$$

in one dimension. During the inference step, the MLE of is calculated with and fixed to their point estimates. There is no trick to get us out of multivariate optimization here, since is nonlinear.

Figure 2.7 Cross-validation recovery of training set labels for the dwarf model. Horizontal lines show boundaries beyond which the model is biased or unprobed by the training set.

2.4 Model Evaluation

We use 12-fold cross-validation (CV) in order to verify that the model is able to recover stellar labels. We partition the reference objects into twelve random subsets, then predict the labels of each subset using the other eleven as training data. This gives us a prediction for each reference star that has not leveraged its GALAH labels. Figures 2.7 and 2.8 show CV performance for the giants and dwarfs respectively, along with the scatter, bias, and correlation coef cient for each label. Our CV-assessed abundance precision ranges from 0.05 to 0.23 dex for dwarfs and 0.07 to 0.22 dex for giants. Examination of the labels inferred for spectra from repeat observation of the same star show differences consistent with CV-precision.

We also use CV to identify the thresholds beyond which our model is highly biased or unprobed by the training data. We say that the model is in this regime when it under- or over-predicts the label being considered 90% of the time in CV. The speci c calculation is as follows: For a given label, ; (e.g. ; =) e), we approximate ?¹; true − ; inferred ° with a kernel-density estimate (KDE) with bandwidth chosen by Silverman's rule (Silverman 2018), then use this approximate distribution to nd the values of ; inferred at which ?¹; true j ; inferred ° excludes ; inferred at the 90% level. These boundaries are shown as horizontal lines in Figures 2.7, and 2.8. Stars that fall beyond these

Figure 2.8 Analogous to Figure 2.7. Cross-validation recovery of training set labels for the giant model. Horizontal lines show boundaries beyond which the model fails to extrapolate. While the scatter in [Fe/H] is higher at lower values, the model appears to be nearly unbiased down to »Fe• H¼ 1•5.

Figure 2.9 The precision of each of our abundances relative to the range over which the model is approximately unbiased. We generally infer abundance ratios with precision at the 30-50% level. Smaller values indicate higher relative precision of that abundance and presumably higher discriminating power between stars.

boundaries are flagged in our catalog. Figure 2.9 shows the precision (twice the scatter in Figures 2.7 and 2.8) of each of our abundances relative to the range over which the model is roughly unbiased. This quantity is often what is relevant when comparing the labels of different stars, rather than characterizing a single star.

2.4.1 Model interpretability

The Cannon is simple enough that its parameters are open to direct interpretation. This sets it apart from more complex modeling approaches such as, e.g. neural networks.

It is clear, for example, that our model learns T_{eff} in large part from the Balmer series, as this is where $\sigma_{T_{\mathrm{eff}}}$ becomes large. By examining model scatter $\sigma_{\mathbf{B}}$ as a function of wavelength, we can tell that our model is less precise in the regions of CN bands, at the beginning of the spectral region of LAMOST. In some wavelength regions $\mathbf{B}$ drops to 0, likely because continuum-normalization introduces small correlations between nearby pixels that are not accounted for by the model. It is also apparent that the model is leveraging the whole spectrum to predict abundances, rather than strong lines only. We performed tests by isolating only regions where individual abundance features are present in the spectra, forced did not allow the coefficients to vary from zero at training time outside of these regions. This is approach is similar to the way in which GALAH determined their abundance labels, using so called abundance windows. For the LAMOST spectra, this approach of using windows fails to recover abundance ratios in cross-validation. Section 2.5.1 discusses some implications of this fact, and our scatter and linear coefficients are plotted as a function of wavelength in Appendix A.1.

2.5 Results

2.5.1 Catalog

We produce a catalog of stellar parameters and individual abundances for 4,541,883 observations of 3,744,284 stars across the Kiel diagram (Figure 2.10), which is available at this URL, along with our model coefficients and training set. We combine observations of the same star by

Figure 2.10 A portion of the best- t model spectra (black) and real data (colored), both with 1- uncertainties, for 6 randomly-chosen stars in our catalog. Though simple, the model is exible enough to t the data across the Kiel diagram.

reporting (-band)(• # -weighted averages of their labels. Along with our inferred stellar parameters and abundances, we provide the LAMOST and Gaia identi ers for each star, as well as its Galactic position, radial velocity and estimated actions. We also provide windowed and whole-spectrum j ^{2} values and ags to tell when the model is extrapolating. For each star, we calculated approximate actions with galpy (Bovy 2015) using the Stäckel fudge (Binney 2012; Bovy and Rix 2013) and MWPotential2014 , applied using Gaia distances (Bailer-Jones et al. 2018) and proper motions, and LAMOST radial velocities (RVs). While Gaia achieves a better RV precision (see Figure A.3), Gaia RVs are only available for approximately one fth of our catalog. We assumed that the Sun sits at = 8 kpc, I = 0•025 kpc (Jurić et al. 2008) and is moving with E = 11•1 km s ^{1}, E = 232•24 km s ^{1}, E = 7•25 km s ^{1} (Schoenrich and Binney 2009). Table 2.1 provides the full catalog schema. To understand the effect of RV and distance uncertainty on our estimated actions, we sampled their values from their error distribution and calculated Galactocentric coordinates and actions for each iteration, performed 20 times. The median uncertainties for , q, and , respectively were 5 kpc km s ^{1}, 21 kpc km s ^{1}, and 1 kpc km s ^{1}. Appendix A.4 explores these errors in more detail.

Here we highlight several caveats to the use of these data:

column name	type	unit	description
source_id	integer		Gaia DR2 source id
designation	string		LAMOST unique star identi er
giantmodel	boolean		true if labels were estimated with giant model
teff	oat	K	) e
logg	oat		log¹6º
vmic	oat		E mic
kiel_extrap	boolean		true if) e or log¹6º (the axes of the Keil diagram) are in regime where model fails to extrapolate for any observation
chi2	oat		whole-spectrum j 2
fe_h	oat		[Fe/H]
fe_h_extrap	boolean		true if [Fe/H] value is in regime where model fails to extrapolate for any observation
x_fe	oat		[X/Fe]
chi2_x_fe	oat		j 2 calculated in windows around strong lines of X
x_fe_extrap	boolean		true if [X/Fe] value is in regime where model fails to extrapolate for any observation
snrz	oat		LAMOST I-band(• #
ra	oat	deg	right-ascension
dec	oat	deg	declination
R	oat	kpc	' , in Galactic cylindrical coordinates
phi	oat	rad	q, in Galactic cylindrical coordinates
z	oat	kpc	I, in Galactic cylindrical coordinates
vR	oat	km s 1	' -velocity
vT	oat	km s 1	q-velocity
vz	oat	km s 1	I-velocity
JR	oat	kpc km s 1	radial action
Jphi	oat	kpc km s 1	angular momentum
Jz	oat	kpc km s 1	vertical action

Table 2.1 Catalog schema. Here, x and X stand for each of the chemical symbols for the elements whose abundances being estimated. Our Galactic coordinate system is right-handed. We also make available the table of per-observation labels.

- We allow our model to take leverage of the full information content of the spectrum. It therefore not only learns from the most fundamental features of each label, but from correlated features (as we can identify using our model coef cients, which is an advantage of a simple interpretable model). Examination of the model's coef cients reveals that the whole spectrum is leveraged in order to predict each abundance. Our CV tests show that our model works. It performs well with no hyperparameter tuning, and our analysis of wide binaries in the solar neighborhood (El-Badry et al. 2019) is indicative of the additional discriminating power beyond an overall metallicity these abundances provide (see Section 2.5.2). However, the abundances are not being measured directly. The delity of our predicted labels relies on our reference objects (con ned to the solar neighborhood) being representative of the survey data. For this reason, our reported abundances may be more accurate for disk stars than halo stars. In order to identify many case where the model fails to generalize from the training set, we provide χ^2 values calculated across the whole spectrum and individually in narrow windows centered on strong lines, for each element. If the best- t spectrum is a poor t around known features of a given element, it is likely highly enriched or depleted in that element. In fact, this approach is a good way to nd such stars with anomalous abundance patterns. Indeed, Casey et al. 2019b, Kemp et al. 2018, and Norfolk et al. 2019 have used the departure of a basic stellar parameter model generated with The Cannon, from the spectra, to nd LAMOST stars that are enhanced in Li, K, and Ba and Sr, respectively.

- A caveat which is general to data-driven methods is that the model will not necessarily extrapolate correctly outside the parameter space spanned by the training set (see Figure 2.4). We provide ags to indicate when individual abundances are in the regime where they may be incorrectly extrapolated, as well as a ag indicating when T_{eff}, $\log g$, E_{mic}, or [Fe/H] may be incorrectly extrapolated (Table 2.1). We determine when our model is extrapolating as described previously, in Section 2.4.

- While error estimates for each abundance ratio are desirable, producing accurate ones would be prohibitively costly with our current inference infrastructure. We advise the user to utilize

Figure 2.11 Inferred [Fe/H] vs effective temperature for LAMOST and GALAH dwarfs in Praesepe. The GALAH values come from internal DR2, which used the same analysis pipeline as the DR2 values in our training set. The systematic trend with $T_{\rm eff}$ is spurious, since all stars in Praesepe have the same abundances to below the precision achievable with low-resolution spectra.

our CV-assessed error and caution them to be aware that treating our abundance estimates as homoscedastic is a necessary compromise.

- Examination of open clusters in our catalog reveals that our inferred abundance ratios for dwarf stars are subject to strong systematics as a function of $T_{\rm eff}$. There are astrophysical explanations for weak abundance trends with $T_{\rm eff}$ and $\log g$, such as atomic diffusion (Dotter et al. 2017; Gao et al. 2018; Souto et al. 2019), but not trends of this magnitude. Similar systematics are present in the GALAH DR2 internal catalog (which employs the same analysis pipeline as the public release), as well as the of cial LAMOST [Fe/H] values for dwarfs, suggesting that these trends are not introduced by our label transfer, but are present in ab-initio stellar models and possibly inherited via our training set. There are no obvious systematics in the red giant stars in our catalogue, save for [Ba/Fe], discussed below. However, LAMOST does not contain enough red giants in known open clusters or wide binaries to determine the presence and magnitude of any systematics conclusively.

  Figure 2.11 shows systematic trends in Praesepe in [Fe/H] as a function of $T_{\rm eff}$ by plotting our inferred values (for stars selected by gaia_clusters:2018) alongside GALAH internal DR2 values for the same open cluster, which is expected to be chemically homogeneous to a level well

below our precision. GALAH internal DR2 includes stars not part of the public DR2, but it employs the same analysis pipeline. Trends of the type exhibited in Figure 2.11 are reduced but not eliminated in GALAH DR3 (Buder et al. in prep). Other abundances exhibit similar behavior. This indicates that systematic error as a function of stellar parameters is a major contributor to our abundance error (see also Section 2.5.2). To the extent that these trends are physical, the recommendation that the catalog user compare stellar abundances within a narrow range of T_{eff} remains.

The systematic trends we see in dwarf abundances could be "calibrated out" using the nearly 3000 stars in LAMOST DR4 open clusters (with two or more targets), (Cantat-Gaudin et al. 2018), and 142 known wide binaries (El-Badry et al. 2019). Instead of applying post hoc correction, they could also be used to constrain the model at training time. Correcting for these systematics in the dwarf population is beyond the scope of our analysis. Despite this systematic effect, our abundances for dwarf stars are still useful for conducting analyses in restricted temperature ranges (see, for example, our examination of abundances of wide binaries in Section 2.5.2). When examining the abundance trends across the disk, we exclude the dwarf stars, and focus on the 1×10^6 red giant stars in our catalogue. These giants span a vast spatial extent, and alone demonstrate the scientific potential of the distribution of stellar abundance data across the Galaxy.

Unless otherwise stated, in the sections below we employ stars in our catalog for which for which chi2 is less than 7000. Other cuts were not found to have an effect on the results presented below.

2.5.2 Detailed abundances of wide binaries

El-Badry et al. 2019 (hereafter EB19) used Gaia to identify wide binaries in the solar neighborhood and examined their properties as a function of [Fe/H]. We examined the detailed abundances of those present in LAMOST. For ease of analysis, we excluded pairs for which $\log g$ or T_{eff} were not available and those containing at least one giant (a total of 8 pairs). Because of the strong

Figure 2.12 Error-weighted chemical distances for EB19 wide binaries and for random field pairs selected with the same cuts and chosen to have the same [Fe/H] distribution as the EB19 sample. The wide binaries are more chemically similar than implied by their similarity in bulk metallicity alone.

systematic trends with T_{eff} that are present in our dwarf abundances, we constrain our analysis to wide binaries with $\Delta T_{eff} < 250$ K, for which both stars are LAMOST dwarfs.

To confirm the additional discriminating power of our inferred abundances, we examined the abundance similarity of wide binaries compared to a reference sample of non-binary pairs. We constructed a set of random pairs of field stars, where each pair has the same metallicity as the binary pair. The reference stars also conform to the quality cuts made in EB19, with $\Delta T_{eff} < 250$ K. We used rejection sampling to ensure that they had as closely as possible the same [Fe/H] distribution as the EB19 sample. By comparing the abundance distribution of the random field pairs with the wide binaries, we can characterize the amount of information contained in our detailed abundances above and beyond that contained by the bulk metallicity, [Fe/H]. To capture the difference in chemistry between stars we use precision-scaled Euclidean distance,

$$\sqrt{\sum_{i} \left(\frac{x_{i} - x_{i}'}{\sigma_{i}} \right)^{2}} \qquad (2.6)$$

where the x_i's are the elements estimated, and the σ_i's are their CV-assessed uncertainty. Figure 2.12 shows the distribution of these chemical distances for both the wide binaries and the field pairs with the same [Fe/H] distribution. The difference between these distributions shows that detailed chemical abundances provide additional information about star's birth sites. Each abun-

dance included pushes the chemical difference distribution of the binaries and random pairs further apart. The wide binaries peak at a smaller chemical distance than the reference pairs. Wide binaries peak at a distance of 0.8 and reference stars peak at a distance of 2.5. This is consistent with findings that the majority of wide binaries are chemically identical to at least the 0.1 dex level (Andrews, Chanamé, and Agüeros 2018; Andrews et al. 2019; Hawkins et al. 2019). We did not find that binaries with a larger separation are more chemically different, in contrast with the results of Ramírez et al. 2019.

If a cut in $T_{\rm eff}$ is not made, the systematic error in each abundance becomes similar in magnitude to the dispersion of chemistry in the solar neighborhood (0.1 – 0.5 dex, depending on abundance). Without this $T_{\rm eff}$ cut, and with this subsequent high systematic error, random field pairs and wide binaries appear to have very similar chemical difference distributions. Even more stringent requirements for $T_{\rm eff}$ result in even more distinct chemical distance distributions for the wide binaries, but at the expense of the number of qualifying wide binaries. In fact, the chemical differences we see in the wide binaries are much smaller than the error we get in cross-validation. This suggests that systematic $T_{\rm eff}$-dependent effects dominate our CV-assessed errors (see Figure 2.11). If, in future work, we were able to reduce or eliminate this effect, perhaps by conditioning a model on chemically homogeneous open clusters, we could produce much higher-fidelity detailed abundances. Currently, scientific exploitation of our 3 million dwarf stars should employ narrow ranges of $T_{\rm eff}$.

Our tests of the chemical differences between wide binary stars indicate not only that the detailed chemistry provides evidence of a common birth site. They also show that systematic effects are a large fraction of our error budget–a promising sign that we can do better with low-resolution spectra in the future.

We also similarly investigated the chemical differences for a sample of co-moving pairs in Kamdar et al. 2019 compared to a reference set of field stars at the same $x_{\rm Fe}$ and found that they were also chemically more similar than an equivalent set of field pairs, although less so than the wide binaries. Simpson et al. 2019 used GALAH abundances to determine whether 15 co-

moving pairs found in Gaia were co-natal; the same approach could be used here.

2.5.3 Mapping chemistry in the Milky Way

We have one of the largest homogeneous samples of stellar abundances. This sample is ideal for mapping the abundance distribution of the Milky Way across a large spatial extent. First, we map the disk across the meridional plane, (R,z), to characterize the spatial abundance trends in that plane. Similar maps can be created with a different set of abundances using APOGEE but that data set is most concentrated to the disk and the inner Galaxy, while the LAMOST giants more extensively span the halo and outer disk. APOGEE data clearly reveal the aring in intermediate-age populations in the '– I º plane (e.g. Ness et al. 2016 This is presumably a consequence of radial migration (e.g. Roškar et al. 2008a), whereby stars increase in scale height as they move outward in the disk (Minchev et al. 2012). Due to the correlations between abundances and ages (Bedell et al. 2018; Feuillet et al. 2018; Feuillet et al. 2019; Ness et al. 2019), we might expect to also see such aring in mean-abundance maps, although this is potentially confounded by the metallicity dependence of the age-abundance relationships (Ness et al. 2021). Detailed analyses of the chemodynamical distribution across '– I º that seek to make any quantitative claims require a careful consideration of the LAMOST selection function. Characterisation of the aring pro le of the disk also require stellar ages, as noted by Minchev, Chiappini, and Martig 2014; Minchev et al. 2018. Here, we aim to show the potential of these data for more in-depth analysis, that accounts for the selection function.

Figure 2.13 shows the '– I º plane colored by mean label value for nearly 800,000 giant stars, for abundance ratios of Fe, O, Eu, Mg, Al, Mn, and Ba as well as $\log_{10}$ and $\log_{10}$. These maps span 4 kpc Ÿ I Ÿ 4 kpc and 7•5 kpc Ÿ ' Ÿ 15 kpc. The disk is clearly distinct from the halo. At ' = 8 kpc, for example, the halo transition appears as a smooth mean abundance change centered on jIj 2 kpc. Flaring is seen in the individual elements, particularly for O, Mg, Eu and Ba. All of these elements show different aring, of varying strength and pro le. All abundances increase or decrease monotonically with jIj at xed ' , except for [Al/Fe], which increases with jIj until

Figure 2.13 The $l' - l^{o}$ plane colored by mean label value for nine labels of our 800,000 giant stars, summed over all b. The aring of the disk can be seen in [Eu/Fe], [Mg/Fe], and [Mn/Fe].

$|I| \quad 2$ kpc, beyond which it decreases with $|I|$.

The apparent barium-depleted "cone" centered on the sun is caused by systematic trends in [Ba/Fe] as a function of $)_e$ and $\log^1 6^o$, in combination with LAMOST's selection function. If we plot only the red clump stars as identi ed by Ting, Hawkins, and Rix 2018 (roughly 2×10^5 stars), which exhibit a narrow range of stellar parameters and which have very precise photometric distances, this feature disappears. The shape and morphology of aring in the elements is preserved when examining the red clump stars only. Finally, we note that both [Ba/Fe] and [O/Fe] appear to be asymmetrically distributed about the Galaxy's midplane. This asymmetry in the mean abundance value around the midplane persists in maps of the $I' - I^o$ plane made with only red clump stars, suggesting that they are not related to $)_e$-dependent systematics. As seen in Figure 2.13, this feature doesn't correlate clearly with $)_e$ or $\log^1 6^o$; nor does it trace extinction as traced by dust maps, or mean signal to noise of the stars. We do not rule out the possibility that the midplane asymmetry seen in these elements is caused by selection effects, particularly in light of the fact that these asymmetrical features are stretched along lines of sight.

Figure 2.14 shows mean-abundance maps in the $I' - I^o$ plane for kinematic thin disk stars ($|I| \ddot{Y} 30$ kpc km s^{-1}), as well as mean $)_e$ and $_+$ maps, for comparison. We highlight the apparent azimuthal structure in [Mn/Fe] and [O/Fe], which isn't easily explained by spurious correlation with $)_e$, $_+$, S/N, $|I|$, or $|I|$. These are two abundances for which our inferred values have lower relative uncertainty (Figure 2.9). Again, we note that a complete treatment of these data would involve explicit modelling of the LAMOST selection function. Note that the sensitivity of [Ba/Fe] to $)_e$ is clearly visible. Like the "cone" in the $I' - I^o$ plane, correlation of [Ba/Fe] with heliocentric distance disappears in maps including only red clump stars. The patterns in these maps along lines of sight are likely of an observational origin, but are not easily explained by a single confounding factor. They do not trace mean height, extinction, metallicity, or $)_e$, or $\log^1 6^o$.

Azimuthal trends in abundance are known to exist in Galactic gas, and are often attributed to spiral structure (e.g. Wenger et al. 2019). Variations in the height of the midplane in combination with LAMOST's selection function could give rise to azimuthal abundance gradients, but it is not

Figure 2.14 Mean abundances of LAMOST giants in the thin disk (30 kpc km s^{-1}), mapped in the [illegible] plane. [illegible] and [illegible] are also mapped for comparison.

Figure 2.15 Mean-abundance maps of thin disk (Ÿ 30 kpc km s^{1}) LAMOST giants in the ¹'– E$_q$º plane. The dashed box outlines the longest ridge of increased stellar number density. We plot only stars with log¹6º Ÿ 3 to minimize contamination from dwarfs, whose abundances are generally on a different scale and have -dependent systematic trends, but a vertical feature is still visible at the solar radius, = 8 kpc, presumably induced by the selection function.

clear why this would be manifest in some abundances and not others unless due to abundance-age correlations, which does not appear to be the case here. Additionally, there is no correlation between the strength of vertical (Figure 2.13) and azimuthal (Figure 2.14) gradients, which would be expected if the azimuthal trends were due to midplane variations.

Figure 2.15 shows mean-abundance maps of LAMOST giants for the thin disk (Ÿ 30 kpc km s^{1}) in the ¹'– E$_q$º plane. By using LAMOST radial velocities, we are able to probe further out into the disk than with the Gaia DR2 RVS sample. In this plane a particularly prominent feature are the `ridges' rst reported by Kawata et al. 2018. A number of interpretations have been given for the

Figure 2.16 The ([Ba/Fe]-[Fe/H]) plane, colored by azimuthal velocity, E_q, and plotted in spatial bins in the Galaxy, with (up to) 700 randomly selected stars plotted in each bin. Because of the T_{eff}-dependent systematics in our inferred [Ba/Fe] values, we have only plotted stars with 4800 K $\leq T_{eff} \leq$ 5000 K. With increasing $|J_\phi|$, halo stars (as distinguished by their lower E_q velocities) become more dominant.

origin of these ridges, including perturbations introduced by spirals, the bar, an external perturber or a combination of these (e.g. Antoja et al. 2018; Khanna et al. 2019; Bland-Hawthorn et al. 2019; Fragkoudi et al. 2018; Fragkoudi et al. 2019a; Laporte et al. 2020). Of particular interest is the longest ridge (outlined by a dashed line). Wheeler et al. 2019 (in prep) will discuss its dynamical origin. Of our abundances, the ridge is most visible in the [O/Fe] and [Mn/Fe]. These are two elements that display the clearest azimuthal abundance gradients, and which our inferred abundances have the lowest relative uncertainty (see Figure 2.9).

2.5.4 The disk-halo transition seen in chemistry

The distributions of the neutron capture elements, [Ba/Fe] and [Eu/Fe], vs [Fe/H], colored by E_q, are plotted in spatial bins in Figures 2.16 and 2.17. Because of the T_{eff}-dependent systematics in [Ba/Fe], we have only plotted stars with 4800 K $\leq T_{eff} \leq$ 5000 K in Figure 2.16. Using other temperature ranges doesn't qualitatively change the plot, but not restricting T_{eff} yields higher dispersion in [Ba/Fe]. The azimuthal velocity, E_q, allows us to clearly distinguish between the disk and halo populations at the scale heights (and vertical actions, J_z) where both are present

Figure 2.17 Same as gure 2.16, with [Eu/Fe] in place of [Ba/Fe] and with stars chosen without restriction on $)_e$.

(primarily the center row in Figures 2.16 and 2.17), illuminating the chemical differences between them. Disk stars are prograde across, and concentrated to low, with most disk stars having E_q & 100 km s^{1}. There are fewer metal-poor disk stars' as increases, with a narrower more metal rich distribution at larger, seen across the smallest range. Halo stars are seen at larger I and are characterized by their more isotropic, eccentric orbits 11 kpc Ÿ ' Ÿ 13 kpc, most halo stars have E_q . 200 km s^{1} with a distribution of E_q = 80 70 km s^{1}. The halo stars also appear to have increasingly negative velocities in the inner Galaxy. At our intermediate height from the plane, 2 kpc Ÿ I Ÿ 4 kpc, the metal-poor stars »Fe•H¼ Ÿ 1•0) are predominantly retrograde at our smallest range, 3 kpc Ÿ ' Ÿ 5 kpc. Cutoffs of 100 km s^{1} kpc and $_q$ 1500 km s^{1} kpc also clearly distinguish between disk and halo stars at large. These populations are not as dramatically distinguished at small because distance errors propagate to larger uncertainties in E_q (see Appendix A.4).

The distribution of (kinematic) halo stars in the ([Ba/Fe],[Fe/H]) plane has a transition at »Fe•H¼ 1 (most clearly seen in the middle row of Figure 2.16, 2 kpc Ÿ jIj Ÿ 4 kpc). This metallicity corresponds to the transition between the disk and halo, as well as the approximate boundary between the accreted and (at least one component of the) in-situ halo (Bonaca et al. 2017, called “the splash” in Belokurov et al. 2019 and the “heated thick disk” by Di Matteo et al.

2019). At least for barium, the abundance planes at $4\ \mathrm{kpc} \leq |z| \leq 6\ \mathrm{kpc}$, suggest an overlap in the chemical plane of different sequences, perhaps associated with the accreted halo. Both [Eu/Fe] and [Ba/Fe] have larger dispersion at high [Fe/H], but the sequence of europium abundances varies less across.

2.6 Discussion & Conclusions

We have trained a data-driven model (The Cannon) to estimate detailed abundances from low-resolution LAMOST spectra, delivering up to 7 abundances for 3.0×10^6 stars to a precision of 0.05 – 0.23 dex. 2.9×10^6 of these are dwarf stars, for which we infer labels with 0.05 – 0.23 dex precision. 8.8×10^5 are red giants for which we infer abundances with 0.07 – 0.22 dex precision. Our best-fit model spectra are easily reproducible using our catalog, implementation of The Cannon, and model coefficients,[3] which are available online. We used the red giants to examine the spatial distribution of abundances in the disk and halo and the dwarf stars to investigate the chemical similarity of wide binaries.

Our analysis of the chemical similarity of dwarf stars in wide binaries compared to field stars showed these stars are from a common birth site and enabled us to quantify the additional resolving or discriminating power in the vector of our derived abundances beyond an overall metallicity.

Using the red giants, we first mapped the profile of the disk in the $[\mathrm{X/H}]$–$[\mathrm{H/H}]$ plane in the elements O, Eu, Mg, Al, Mn, and Ba. These maps show the flaring of the disk and the distinction in abundances between the halo population at high latitudes and the disk. Second, we examined face-on projections across the disk in this set of abundances. These projections hint at some non-axisymmetric patterns in the abundances. Indeed, Gaia mission has revealed numerous dynamical deviations from axisymmetry in the disk and perturbations in the solar neighbourhood (Antoja et al. 2018; Sellwood et al. 2019; Trick 2020). Third, we constructed mean-abundance maps in L_z–E_z plane and discuss the chemical signature of the high density ridges in this plane. Finally, we investigated the abundance planes of [Ba/Fe]-[Fe/H] and [Eu/Fe]-[Fe/H] across, making similar

[3] See doi.org/10.7910/DVN/5VWKMC

maps to Hayden et al. 2015, but with the neutron capture abundances. These maps showed the disk and halo trends across [Fe/H] at all [α/Fe]. These different trends might be used to separate any in-situ halo from heated disk stars, from an accreted halo. As the set of abundances we deliver give higher discriminating power to identify chemically similar stars compared to [Fe/H] alone, we expect the multiple families of abundances will be useful for studies of the plethora of chemo-dynamical substructure in the Milky Way halo (see e.g. Belokurov et al. 2019; Helmi et al. 2018; Di Matteo et al. 2019; Myeong et al. 2019; Antoja et al. 2018).

We derive abundances with diverse nucleosynthetic channels and are demonstrably uncovering some of the breadth of chemical information in the Milky Way. However, a number of caveats are discussed in Section 2.5.1 and we further detail some of these here.

In contrast to Ho et al. 2017a; Ho et al. 2018, we find that our cross validation results do not vary strongly with S/N. This indicates that our precision is limited by that of reference labels themselves, and, if improved, we would obtain higher precision results for our test objects that scale as expected with SNR.

Because our model is in some cases not inferring abundances from the corresponding lines themselves, it may not be robust to stars with properties or enrichment histories not represented in the training set. This means that, while stars that are highly enriched or depleted in an element may not have their abundances accurately inferred, the best fit model should have large residuals in regions of its known lines. Casey et al. 2019a used this effect to identify stars that are unusually rich in lithium, but this approach could be extended to all elements with strong lines in the LAMOST wavelength range. In particular, it is challenging to measure r-process abundances from r-process absorption features even from extremely high quality high resolution spectra, and GALAH uses only 2 relatively unblended absorption regions for their [Eu/Fe] measurement (Table A.3). It raises questions about stellar spectra that we are inferring [Eu/Fe] (albeit noisily) from relatively low-(R ≈ 1800), low-resolution spectra, as confirmed by cross validation, and that the [Eu/Fe] distribution across [Fe/H] mirrors that of boutique studies (e.g. Bensby et al. 2020). The physical origin of the significant correlations between absorption features of nominally different nucleosynthetic families

(Feeney, Wandelt, and Ness 2019) is not clear, but is presumably caused by a combination of the inherent correlation induced by element-production mechanisms, shared chemical enrichment history, and stellar physics. In other words, the chemical manifold on which the majority of stars lie is not well-known. Of signi cant interest is most likely those stars where we can not well match the spectra with our data-driven model, which is by far the minority of stars in LAMOST.

The large number of low- and medium-resolution spectra available now (RAVE and SEGUE, in addition to LAMOST) and coming in the near future (e.g. WEAVE, Dalton et al. 2012, MOONs, Cirasuolo et al. 2014, and 4MOST, de Jong et al. 2019 in their lower-resolution modes; DESI, DESI Collaboration et al. 2016, Sloan V Kollmeier et al. 2017, and Gaia, Gaia Collaboration et al. 2016) makes honing our ability to learn from these data a fruitful endeavor. We also discussed future improvements to our methodology, especially the possibility of using open clusters to reduce the effect of systematic trends with stellar parameters in inferred abundances. Other promising methodological directions include using more robust inference for model parameters and labels, perhaps allowing more rigorous error estimation, and allowing missing labels in the training set, which would enable us to used training data from multiple surveys.

Chapter 3: An unsupervised method for identifying -enriched stars directly from spectra: Li in LAMOST

3.1 Introduction

Recent years have seen many resolved-star spectroscopic surveys, and corresponding growth in the use of data-driven spectral models, e.g. The Cannon (Ness et al. 2015), the DD-Payne (Xiang et al. 2019), kernel principal component analysis approaches (Xiang et al. 2017), and models using deep convolutional neural nets (Leung and Bovy 2019). These models have pushed astronomers into new regimes of precision (Jofré, Heiter, and Soubiran 2019), and allowed us to infer evolutionary state, mass, and detailed abundances from low-resolution spectra (Ho et al. 2017b; Ting, Hawkins, and Rix 2018; Xiang et al. 2019; Wheeler et al. 2020; Sandford, Weisz, and Ting 2020).

Given a complete enough understanding of stellar atmospheres, the interstellar medium, the earth's atmosphere, and our instrumentation, data-driven spectral models would be unnecessary, but we are far from such total knowledge. Most data-driven methods applied to stellar spectra (including all mentioned above) are concerned with supervised regression. They use a set of spectra labelled with a-priori atmospheric parameters, surface abundances, and reddening parameters, to calibrate a model which is then applied to unlabelled spectra. These approaches are limited by the quantity and precision of training data, restricting the datasets, labels, and regimes to which they can be applied. Furthermore, their reliance on labeled data means that they can be limited by the strong biases and systematic errors introduced by physical models.

In this work, we pursue an alternative approach. Rather than estimate abundances, we endeavour only to identify stars that are highly enriched[1] in a particular element. This relaxed goal permits

[1] Throughout this chapter, we use "enriched" or -enriched specifically to refer to stars with a high abundance relative to stars with similar spectra and hence parameters and abundances. A star which is -rich may not be -enriched, and vice versa.

approaches which require no labelled training data and minimal researcher input, but which retain high scienti c payoff, since -enriched stars often have undergone uncommon events. Unusual abundance patterns may also signal birth in an accreted galaxy with an enrichment history different from the Milky Way's (e.g. Horta et al. 2020; Hawkins, Lucey, and Curtis 2020; Molaro, Cescutti, and Fu 2020; Vincenzo et al. 2019).

We take advantage of the fact that stellar spectra, which are naturally expressed as high-dimensional ux vectors, are embeddable on (or near) a lower-dimensional manifold. We censor the relevant absorption region of a spectrum, then use its neighbors on a local (euclidean) patch of the spectral manifold to impute the masked pixels. This serves as a (possibly) couterfactual unenriched realization of the spectrum, against which we examine the residuals to identify "unexpectedly" strong absorption.

This work is related to ideas in manifold learning and nonlinear dimensionality reduction, especially local linear embedding (Roweis and Saul 2000) and isomap (Tenenbaum 2000). Hessian local linear embedding (Donoho and Grimes 2003) shares with our method the use of singular value decomposition to estimate the tangent space of the data-manifold. Unlike these methods, we never explicitly construct global nonlinear coordinates since all of our calculations can be performed in a small patch on the spectral manifold. While we leverage the embeddability of spectra in a low-dimensional space, we never construct a continuous low-dimensional representation.

This is not the rst unsupervised model deployed on a large spectral survey. Feeney, Wandelt, and Ness 2019 used a fully probabilistic nonparametric model to characterize the spectra of APOGEE red clump stars, denoising them by a factor of a few and demonstrating the mutual information present in features of elements belonging to the same nucleosynthetic family. Though their model could, in principle, be used to impute masked pixels, it is too expensive to deploy across all wavelengths of all observed stars. erjal et al. 2019 used a simple nearest-neighbor method to identify Li-enriched KM dwarfs in GALAH. Cotar et al. 2020 used autoencoders (a family of neural-net architectures) to identify emission stars.

We turn our attention to stars that are enriched in Li. Li-7 burns at a re 2.5×10^6 K (Bo-

denheimer 1965), and is depleted at all stages of stellar evolution, but thought to be replenished in myriad ways. Its low ionization potential means that in stellar atmospheres Li exists mostly as Li-II, which is not detectable. Li-I's strongest feature (the 6708 Å doublet) is thus very weak and $_e$ -sensitive. It is also sensitive to non LTE (local thermodynamic equilibrium) effects (Lind, Asplund, and Barklem 2009). All this means that Li can be challenging to study with physical models.

3.2 Data

We use the LAMOST DR 5 (Deng et al. 2012; Zhao et al. 2012) AFGK sample, which contains $5 \cdot 3 \quad 10^6$ low-resolution ($= 1800$) spectra of $4 \cdot 3 \quad 10^5$ unique stars. It covers wavelengths from 3800 Å to 9000 Å. We analyze repeat observations independently with TSP, but stack the resulting residuals before identifying strong absorption. We pre-treat all LAMOST spectra by interpolating to a common rest-frame wavelength grid, then applying the approximate continuum normalization rst used in Ho et al. 2017a, that is, dividing the spectrum by a itself smoothed with a 25 Å Gaussian kernel. While this transformation distorts broad features, it is applied homogeneously across all spectra and thus will not introduce biases. We impute any pixels with normalized ux, 5, outside of $0 \quad 5 \quad 2$ or with uncertainty greater than $\frac{1}{2}$ by setting $5 = 1$, and setting the associated uncertainty to of .

3.3 Methods

The inputs to our algorithm are the following:

- The data, assumed to be a set of uniform vectors. In this work these are rest-frame spectra, interpolated onto a common wavelength grid.
- The reference set, a subset of the data well-distributed throughout the underlying parameter space[2]

[2] In machine learning terminology, this might be called the training set. In the L-ISOMAP dimensionality reduction algorithm (de Silva and Tenenbaum 2002), these data are called “landmarks”.

- Integers: and @which specify the number of neighboring data points to use and the dimensionality of the manifold, respectively.

- The components of each data point that are of interest. For stellar spectra, these are the pixels containing the spectral feature(s) under investigation. We will refer to data as masked when these components are dropped. We refer to the components themselves as masked. We take = to be the number of unmasked pixels, and to be the number of masked pixels.

We take each spectrum to have ux uncertainty described exactly by a multivariate normal distribution with known covariance. The most naive version of this algorithm requires complete data. As discussed in section 3.2, we obtain complete data by imputing bad pixels with 1. Note, however, that TSP could in principle itself be used to impute bad pixels more robustly.

Ideally, the reference set would include every spectrum available (perhaps excluding those with strong absorption features, see discussion in section 3.5). Using a random subset of spectra instead speeds up computation.

We need to know the expected pro le of the spectral feature under investigation to identify enriched stars after imputing. For all but the strongest absorption lines in low-resolution spectra, knowing the instrument resolution is suf cient, since the line's observed pro le is dominated by the line spread function.

Figure 3.1 presents an overview of the algorithm, section 3.3.1 goes into detail, and 3.3.2 describes our approach to identifying enriched stars from imputed residuals.

3.3.1 Tangent Space Projection

Stated brie y, TSP is imputation of masked data via local principal component regression. Here we detail how to use TSP to predict masked spectral pixels of an arbitrary target spectrum, using a reference set of randomly-selected unlabeled spectra. The reference set and target spectra are assumed to be from the same instrument and interpolated to the same wavelength grid. Take the target spectrum's $\underline{s}^{th}$ unmasked pixel to be . First nd the : nearest neighbors in the reference

For each spectrum f;

1. Compare f to all spectra in the reference set (with the region of the absorption feature masked) to find its k nearest neighbors.
2. Compute a basis for the d-hyperplane that captures as much variance amongst the neighbors as possible, the approximate tangent space.
3. Impute the masked pixels by projecting f onto the approximate tangent space.
4. Determine if the residuals corresponds to excess absorption.

Figure 3.1 Algorithm summary.

set (leaving out the masked portion of the spectra), i.e. those that minimize the euclidean distance (L2 norm),

$$3 = \sqrt{\sum_{_=1}^{1} \left(5_{_} - 5_{_}^{refo2} \right)} \qquad (3.1)$$

for a reference spectrum whose $_^{th}$ pixel is $5_{_}^{ref}$. We exclude those reference spectra with missing data in the masked region.

Take F to be the $: \times 1 = ¸ < º$ matrix whose rows are the uncensored neighboring spectra in the zero-mean basis (with the neighborhood mean subtracted from each row). Next, calculate the d-hyperplane that captures as much variance as possible amongst the neighbors (using their full, uncensored spectra), that is the d principal components of F. Let E be the $d \times =$ matrix whose rows are the first d uncensored $=$-pixel eigenspectra, and let E^{c} be the $d \times <$ matrix whose rows are the $<$ masked pixels of the first d eigenspectra. Together, E and E^{c} are an estimate of the tangent space of the spectral manifold. In the context of Equation 3.2, below, each is a design matrix whose rows can be thought of as features.

Let f be the $=$-pixel (censored) target spectrum in the zero-mean basis, and let Σ be the $= \times =$ covariance matrix describing the uncertainty in f. We can predict the masked pixels, f^{pred}, by

projecting f onto the row space of E:

$$\hat{f}^{pred} = \hat{E}\,\underbrace{\left(E^{T}\hat{E}\right)^{-1}\left(E^{T}\hat{E}\right)^{-1}E^{T}\hat{f}}_{\text{linear least-squares coeffs}} \quad . \tag{3.2}$$

The marked part of this equation is uncertainty-weighted linear least-squares regression (Hastie, Tibshirani, and Friedman 2009 chapter 3 and Hogg, Bovy, and Lang 2010 provide comprehensive reviews); it is a representation of f in $\mathcal{O}$-dimensional local coordinates. That representation is then mapped to the masked pixels through multiplication by $\hat{E}$. $\hat{f}^{pred}$ is the maximum-likelihood estimate for the $<$ masked pixel values given f and , assuming that the ground truth spectrum lies on the hyperplane spanned by the $\mathcal{O}$ eigenspectra. Treating the eigenspectra $\hat{E}$ and E, as xed, the covariance describing the uncertainty of $\hat{f}^{pred}$ is $\hat{e}_{pred} = \hat{E}\left(E^{T}\hat{E}\right)^{-1}E^{T}\left(E^{T}\hat{E}\right)^{-1}$. In practice, we neglect this uncertainty since when $=$ $<$ (i.e. when a small fraction of all pixels are masked) it is subdominant to the measurement uncertainty in the masked pixels,

3.3.2 Identifying Strong Absorption in the Residuals

Having used the unmasked pixels to predict the masked ones, we can examine the residuals between the measured and predicted masked pixels to identify unusually strong absorption features and eliminate contaminants. We use a simple likelihood ratio test, i.e. matched ltering.

Let r be the residuals of the prediction of the uxes of the masked pixels, $r = \hat{f} - \hat{f}^{pred}$, and let C be the covariance matrix describing its uncertainty. Consider a family of models of the form Um, where U is a scalar amplitude and m is a xed pro le. Since the uncertainty of $\hat{f}^{pred}$ is small, C is the diagonal matrix whose entries are the squares of the per-pixel measurement uncertainty of $\hat{f}$. The value of U which maximizes the likelihood is

$$U = \frac{m^{T} C^{-1} r}{m^{T} C^{-1} m} \quad . \tag{3.3}$$

Figure 3.2 Our line model (black) and contaminant models (grey, dashed). Only one single-pixel contaminant model is shown for clarity.

with uncertainty

$$f_U = {}^{1}m^{)} C^{1} m^{o\ 1\bullet 2} \quad \bullet \tag{3.4}$$

For a given m, the optimal-amplitude likelihood is

$$\log ?^{1} e\, jU\, m^{o} = \frac{1}{2} \frac{{}^{1}m^{)} C^{1} e^{o2}}{m^{)} C^{1} m}, \ \text{const} \quad \bullet \tag{3.5}$$

If m is a line model with equivalent width F, U F can be interpreted as best- t excess equivalent width (EEW), equivalent width in excess of that found in a counterfactually unenriched spectrum.

We calculate the likelihood, amplitude (EEW), and amplitude uncertainty for a Gaussian (in wavelength) absorption feature with width given by the instrument resolution, as well as for two contaminant models: uniform residuals, as might arise from a poorly modelled continuum, and residuals with a single non-zero pixel. Our line model and contaminant models are plotted in Figure 3.2. We require the contaminant models to be less likely than the line model to identify a star at Li-enriched (see section 3.4).

3.3.3 Hyperparameter Selection

To inform the choice of hyperparameters s and r, we evaluate TSP's predictive accuracy for arbitrary spectral regions. We take 1000 random LAMOST spectra and predict a different randomly-selected contiguous region of 100 pixels (roughly an order of magnitude larger than a mask for a single line) for each. Figure 3.3 shows the per-pixel χ^2 of the masked pixels as a function of subspace dimension r for different neighborhood sizes, Large values of each ($s = 250$, $r = 25$) give better predictions of the masked pixels, which is consistent with the large number latent parameters found to meaningfully describe LAMOST spectra by (Xiang et al. 2017). The nominal measurment uncertainty is saturated for nearly all values of s and r, indicating that it may be overestimated, although our continuum normalization procedure introduces correlations between nearby pixels that may contribute to this. Using the mean of the k nearest neighbor spectra as a prediction, while simpler than TSP, is not as effective. The horizontal dashed line in Figure 3.3 shows the predictive accuracy for $k = 3$, the neighborhood size for which a local average performs best. TSP achieves better predictive accuracy for nearly all hyperparameter values considered.

For our case-study, we choose $s = 1000$ and $r = 50$. We remark, however, that predictive accuracy in the Li absorption region is not especially sensitive to this choice, as demonstrated by Figure 3.4, which shows the masked portion of a spectrum exhibiting strong Li absorption with TSP's predictions for $r = 4$ and $r = 50$. TSP imputes these pixels very similarly regardless of our choice of r. Note that TSP predicts no absorption, even though some neighbors have absorption similar to that of the spectrum being imputed. This is because both Li absorption features have been masked, and the unmasked pixels do not contain correlated spectral features.

We found that increasing the size of the reference set decreased χ^2 per pixel, but changed the preferred values of s, as expected. We use a reference set of 30,000 randomly chosen spectra, but remark that a larger reference set would presumably marginally decrease predictive χ^2 per pixel at a greater computational cost. Using a reference set composed exclusively of spectra with a high signal to noise ratio ($S/N > 100$) results in poor predictive accuracy. Examination of stellar parameters for repeat observations of the same star did not reveal any S/N-dependent bias in stellar

Figure 3.3 Mean squared error of the predicted spectral region as a function of subspace dimension, @ for different neighborhood sizes, . The dashed grey line marks the best pixel (: = 3) achieved by using the local mean spectrum as a prediction. We use 1000, @= 50 for our analysis.

parameters However, both continuum-normalized and un-normalized spectra of repeat observations have small S/N-dependent systematics, indicated that both the LAMOST reduction pipeline and our continuum-normalization introduce biases which may prevent a reference set of all high-S/N spectra from performing well on low-S/N spectra.

3.4 Results

We apply our method to the 6708 Å feature, masking wavelengths in the range 6708Å – 6717Å (vacuum). Table 3.4 shows the output quantities for each star. We mask, but do not use, the 6106 Å Li feature. To consider stars to be Li-enriched, we require the detection of an absorption feature with an EEW of at least 0.15 Å with an uncertainty excluding null EEW at the 3f level (f EEW Ÿ 0•05Å), and that the absorption line model t the data better than any contaminant model. Of these, 147 stars have Li abundances available in GALAH DR3 (Buder et al. 2020). They have a mean [Li/Fe] of 1.7, with a standard deviation of 0.5, and 74% have [Li/Fe] 1•5, the traditional threshold. When evaluating the population properties of Li-enriched stars, we throw out those stars

Figure 3.4 The prediction of the Li absorption region in a spectrum of LAMOST J062739.73+463634.4, a Li-enriched giant. The imputed spectrum for 30 – @ = 4 is very similar to that for: = 1000 – @ = 50. Neither predict the absorption seen in the data. The 30 nearest spectral neighbors are shown in gray.

column name	type	description
designation	string	LAMOST unique star identi er
source_id	string	Gaia identi er
diff	oat array	ux residuals for each masked pixel
ivar	oat array	inverse variance in the ux residuals
max_best_fit_chi2	Float	largest whole-spectrum m^2 for any observation of this star
isline	bool array	true when the line model is more likely than any contaminant model
likelihoods	oat array	the likelihood values (up to an additive constant) of the line model and each contaminant model
EEW	oat	excess equivalent width [Å]
EEW_err	oat	uncertainty in the EEW, f_{EEW} [Å]
enriched	bool	true if the star is Li-enriched per de nition in 3.4
ra	oat	right ascention
dec	oat	LAMOST unique star identi er
teff	oat	$)_e$ [K] from LAMOST (mean of observations)
logg	oat	log¹6º from LAMOST (mean of observations)
feh	oat	[Fe/H] from LAMOST (mean of observations)
snrr	oat	LAMOST A-band¹(• # º (largest of all observations)

Table 3.1 Catalog schema. Data available at https://doi.org/10.7916/d8-3ap9-qe35 .

Figure 3.5left: Portions of 21 spectra of Li-enriched stars spaced roughly evenly on the Kiel diagram and sorted by y_U (6565 Å) amplitude.right: their positions on the Kiel diagram, with mass tracks for 2 , 1.5 " , 1 " , 0.8 " , and 0.5" solar metallicity stars (details in text).

whose uncertainties make such a detection impossible (those for with EEW uncertainty greater than 0.05 Å, about 40% of all LAMOST stars). Hereafter, we refer to stars with W_{EEW} Ÿ 0•05 as candidates (and those that are both candidates and have W_{EEW} 0•15 simply as Li-enriched). Within the set of candidates, there is no dependence of EEW on $(V-I)_0$. Figure 3.5 shows portions of 21 Li-enriched spectra on different parts of the Kiel diagram.

Figure 3.6 shows a histogram of the EEW values for all candidate stars, and all Li-enriched stars. Nearly all stars with EEW 0•15 Å have their masked-pixel residuals best matched by the line model (see section 3.3.2), except at EEW 0•7 , where a signi cant fraction of spectra have continuum offsets. The width of the EEW distribution is narrower than naively expected given the nominal LAMOST measurement uncertainty, again indicating that it is overestimated. There is an uptick of in the number of stars with EEW around 0.2 – 0.6 Å, formed by the excess of Li-enriched stars. There are 8,428 Li-enriched stars, and 2,147 candidate stars best t with the Li line model with EEWŸ 0•15 Å, suggesting a contamination rate of up to 25% (the presence of any Li-depleted stars will in ate this gure).

There are no obvious spatial trends in the fraction of Li-enriched stars; there are only trends in stellar parameters and evolutionary state. Figure 3.7 shows the 8,428 Li-enriched stars in the

Figure 3.6 The distribution of excess equivalent widths for all candidates (those with EW_{Li} > 0.05 Å) and for Li-enriched stars (candidates with EW_{Li} > 0.15 Å).

log(L/L$_\odot$) – T_{eff} plane (the Kiel diagram), along with the occurrence fraction of Li enrichment, and the number of candidates. Plotted for comparison are solar-metallicity mass tracks from MESA isochrones and stellar tracks[3] (MIST; Dotter 2016; Choi et al. 2016; Paxton et al. 2011; Paxton et al. 2013; Paxton et al. 2015), Li enrichment is especially prevalent among pre-main sequence stars, stars near the zero-age main sequence (ZAMS), especially at T_{eff} > 5000 K, subgiant branch stars, and red giants at and above the red clump. We go into detail in the sections below.

3.4.1 Planet Accretion on the Subgiant Branch

Presumably, stars occasionally consume their satellites. These events may enrich the star's photosphere, especially when the satellites are large. Accretion of sub-stellar companions has been identified many times as a Li-enrichment mechanism (Alexander 1967; Siess and Livio 1999; Villaver and Livio 2009; Adamów et al. 2012; Carlberg et al. 2012; Oh et al. 2018), and detection of cometary material accreted by white dwarfs has a long history (van Maanen 1917; Weidemann 1960; Zuckerman et al. 2003; Kepler et al. 2016, e.g.). Recently, Soares-Furtado et al. 2020 found that, while photospheric Li will become unobservable within roughly 10 Gyr for red giants, which

[3] MIST version 1.2. Mass tracks generated with initial v/v_{crit} set to 0.4.

Figure 3.7 top: The 8,428 Li-enriched stars plotted on the Kiel diagram with MIST solar-metallicity mass tracks. bottom left: fraction all stars with f_{EEW} Ÿ 0•05 Å with are Li-enriched. bottom right: the number of of stars with f_{EEW} Ÿ 0•05 Å. Dashed vertical lines are the same as in Figure 3.9. They mark the range over which young stars are Li-enriched.

have deep convective envelopes, it has a lifetime of up to 1 Gyr for 1.4 ¨ 1.6 stars on the subgiant branch and near the main sequence turn-off. Compounded with the fact that stars on the subgiant branch are expanding rapidly, this suggests that planet accretion could account for a signi cant fraction of Li-enriched subgiants.

Figure 3.7 shows a moderate enrichement fraction (0•1%) stars near the 1.5 subgiant branch. Neither the data-driven abundances of Xiang et al. 2019 (Na, Mg, Al, Si, Ca, Ti, Cr, Mn, Co, Ni, Cu, Ba) nor Wheeler et al. 2020 (O, Sc, Eu, Mg, Al, Mn, Ba) show a signi cant difference in mean abundances between Li-enriched and normal subgiants. Stars that have recently accreted planets are speculated to be especially enhanced in refractory elements, those with a high condensation temperature (e.g. Ramírez, Meléndez, and Asplund 2009; Melendez et al. 2009; Melendez et al. 2017). We nd no evidence for a positive trend between condensation temperature and comparative enrichment of Li-enriched subgiants over Li-normal subgiants. In fact, a weak negative trend is present—Li-enriched stars are preferentially depleted in refractory material, with the exception of Ba.

3.4.2 Li-enriched Red Giants

More than half the stars we identify as Li-enriched are giants. Of the candidate stars, 4459 have log¹6º Ÿ 2•7, consistent with the roughly 1% of red giants thought to be Li-enriched (traditionally de ned as ¹Liº ¡ 1•5; e.g. Gao et al. 2019, Casey et al. 2019b). As seen in Figure 3.7, they span temperatures 4000K .)e . 5600K (3•6 . $\log_{10}$)e »K¼ 3•75).

The rst Li-enriched red giant is often considered to be that identi ed by Wallerstein and Sneden 1982, although a Li-enriched asymptotic giant branch (AGB) star was reported four decades earlier by McKellar 1940. As mentioned in section 3.4.1, sub-stellar companion accretion is unlikely to explain a signi cant fraction of Li enrichment for red giants. Cosmic ray spallation can account for some Li-enrichment (Burbidge et al. 1957; Mitler 1964), but only a small fraction, and does not produce the observed isotope ratio (Reeves 1994). Mass-transfer from an AGB star companion can account for Li-enrichment in those that have one, but not in isolated giants. Li-

enriched giants are not preferentially found in binaries (Adamów et al. 2018). Li-enriched material is produced by classical novae (Starr eld, Gehrz, and Truran 1997; Molaro et al. 2016) and Type II supernovae (Dearborn et al. 1989), but it is unclear whether, and in what quantity, this material can be accreted onto giant stars.

Cameron 1955 and Cameron and Fowler 1971 rst suggested that surface Li could be enhanced if Be-7 is transported via convection from a depth at which PP-II fusion is occurring and decays via electron capture to Li-7. For the Be to be transported to a cool layer of the star before it decays into Li, the convection timescale must be faster than its decay timescale, thought to be 50 – 100 yr (Cameron 1955). While this process is thought to occur as originally suggested in AGB stars (Deepak and Reddy 2019; Singh et al. 2019), driven solely by convection (Smith and Lambert 1989) or thermal pulsations (e.g. Habing and Olofsson 2011 chapter 1), it must be augmented with an additional mixing mechanism for RGB stars. Two possibilities are internal rotation (Sweigart and Mengel 1979; Fekel and Balachandran 1993; Charbonnel 1995), and thermohaline mixing (Sackmann and Boothroyd 1999; Charbonnel and Balachandran 2000; Denissenkov and VandenBerg 2003; Lattanzio et al. 2015). Casey et al. 2019b found that tidal interactions were the likely culprit, while Martell et al. 2020 found that at least two mechanisms are likely in effect. Recently, Kumar et al. 2020 found that a ubiquitous process is in operation for all low-mass stars, on the basis of the high Li abundances in the red clump.

Extremely few of our Li-enriched giants are Ba-enriched (see section 3.5.1), which indicates that thermally-pulsating AGB are not contributing signi cantly to the Li-enriched sample. Early AGB stars (E-AGBs), which may not experience signi cant dredge-up are more dif cult to observationally distinguish from red giants. That we see a much higher Li-enrichment rate at and above the red clump than below is suggestive that E-AGBs may make up a signi cant fraction of the Li-enriched sample. On the other hand, our enrichment rate is similar to past studies, suggesting that we are identifying red giants.

We see no evidence on the basis of the $\log g$ and $T_{\rm eff}$ distributions of enriched stars that Li-enrichment is more common in the red clump (RC) than on the RGB, in sharp contrast with

Deepak and Reddy 2019; Martell et al. 2020; Casey et al. 2019b. This may be partially attributable to our direct use of absorption feature, rather than calculating the abundance from the equivalent width. The more likely possibility is that TSP is able to correctly predict Li-enrichment when it is due to whatever mechanism preferentially enriches red clump stars. If, as Kumar et al. 2020 posit, this mechanism is very common, it's especially likely that TSP would capture its effects. Understanding which spectral features contain joint information with the Li doublet might indicate the mixing mechanism driving photospheric Li-enrichment. Notably, we identify few Li-enriched giants below the red clump.

In agreement with other recent work (Casey et al. 2019b; Martell et al. 2020; Deepak and Reddy 2019), we see that the occurrence rate of Li-enriched giants increases strongly with metallicity (Figure 3.8). However, we see no evidence for a sharp increase in the occurrence fraction as [Fe/H] increases past the solar value, as noted by Martell et al. 2020 and hinted at in the data of Casey et al. 2019b. Curiously, Martell et al. 2020 found that a smooth increase on Li-enrichment was associated with RC stars, and that RGB stars only had enrichment “turning on” at super-solar metallicties. We see a similar smooth increase in Li-enrichment with metallicity when plotting giants only above or only below the red clump. Those distinct trends are likely attributable to a multiplicity of enrichment mechanisms, further suggesting that our analysis is sensitive some mechanisms but not others.

3.4.3 Young and pre-Main Sequence Stars

Some Li-enriched stars are enriched by virtue of having not yet depleted their birth Li. We are more likely to identify pre-main-sequence stars as Li-enriched than stars at any other evolutionary stage, with the enriched fraction exceeding 50% for those with $\log g$ ¡ 4. These stars are exhibit some of the largest EEWs (see Figure 3.7), with some larger than 0.5 Å, and roughly half have emission, indicating magnetic activity.

We know from studies of open clusters that on and near the zero-age main sequence (ZAMS), the Li abundance is only mildly $T_{\rm eff}$ -dependent for stars with $T_{\rm eff}$ & 5000 K ($\log_{10}$) $T_{\rm eff}$ »K¼& 3•7),

Figure 3.8 The raw count and occurrence fraction of Li-enriched giants as a function of metallicity. Errorbars are determined assuming Poisson uncertainty in raw counts. The fraction of Li-enriched giants increases smoothly with [Fe/H].

but decreased quickly with decreasing $T_{\rm eff}$ below that value (e.g. Sestito and Randich 2005). Together with the fact that the 6708 Å Li doublet is $T_{\rm eff}$-sensitive, this means that we primarily identify your main sequence stars of moderate effective temperature. These effects can be seen clearly in stars in the Pleiades (as identified with Gaia astrometry by Gaia Collaboration et al. 2018b), the open cluster in which LAMOST has observed the most stars. Figure 3.9 shows the EEW for each. Dashed lines mark EEW = 0.15 Å, our Li-enrichment threshold, and the temperatures within which Pleiades stars exceed this threshold ($T_{\rm eff} = 4850 - 5750(3.67 - \log_{10} T_{\rm eff})$ K = 3.76). Reliably identifying young main sequence stars, even in a limited temperature range, can help to dissect the Galaxy's dynamical perturbations (e.g. Laporte et al. 2020), identify and validate young exoplanetary systems (e.g. Mann et al. 2020), and potentially improve our understanding of star formation by identifying recently dispersed clusters (e.g. Foster et al. 2015).

3.5 Discussion

We have presented an unsupervised method, tangent space projection (TSP), for identifying stars which are high-n-enriched, for any element. It uses the fact that most stellar spectra lie on

Figure 3.9 Li EEWs as a function of T_{eff} for the Pleiades stars observed by LAMOST. Note that the x-axis is reversed. The grey dashed line marks the threshold at which we consider a star Li-enriched. The decrease in EEW with increasing effective temperature is due to Li's low ionization potential. The decrease with decreasing effective temperature is because those stars have depleted more of their Li.

a low-dimensional manifold, but those of chemically aberrant stars often don't. We applied TSP to the 6709 Å Li doublet in LAMOST DR5, identifying 8,428 Li-enriched stars.

TSP is applicable to any homogeneous catalog of spectra. Its most useful application is to blended features and those for which physical modelling remains a challenge. For low-resolution surveys like LAMOST, essentially all spectral features are blended to some extent, but medium- and high- resolution surveys, such as Gaia DR4, RAVE (Steinmetz et al. 2020; Casey et al. 2017), APOGEE (Majewski et al. 2017; García Pérez et al. 2016), and Sloan V (Kollmeier et al. 2017) also contain many such features, e.g. the recently noted Ce and Nd lines (Hasselquist et al. 2016; Cunha et al. 2017) in APOGEE. In the disk, we expect to find few stars with unusual enrichment patterns, but they have to potential to give us unique insight into nucleosynthetic processes (Weinberg et al. 2019). In the halo, the chemical signatures are strongly linked to the dynamical history of the Milky Way (e.g. Ji et al. 2020; Das, Hawkins, and Jofré 2020; Naidu et al. 2020).

Figure 3.10 A portion of a spectrum of an Ba-rich star (top; LAMOST J120744.31+172831.6) and an Fe-rich star (bottom; LAMOST J093835.58-064404.6) as observed by LAMOST. TSP correctly imputes the masked region (marked in grey) for the Fe-rich star, but not the Ba-rich star. Note that the continuum is underestimated to different degrees in each, so the y-axes are not directly comparable.

3.5.1 Blended lines

The 6708 Å Li doublet is blended with a weak Fe I line, which does not pose an analysis problem except in very metal-rich stars. To demonstrate that TSP can still effectively identify chemical outliers in the presence of stronger blended features, we masked and imputed the region surrounding the 4555 Å Ba line (with all other Ba lines masked), which is blended with two Fe lines. Figure 3.10 shows a portion of the spectrum for a Ba-enriched (and Ba-rich) star, and for a Fe-rich, Ba-normal star. As expected, TSP correctly imputes the masked region when there is strong absorption due to the Fe line, but not when there is strong absorption from the Ba line. Estimates of abundances from supervised models are often unreliable for extreme abundance values. A complete analysis of blended features should also take blends into account when examining the imputed residuals. For example, were the 6708 Å Li doublet blended with a stronger line, we would have included absorption due to this feature as a contaminant model (section 3.3.2).

3.5.2 Algorithm

We see the strengths of TSP not in sophistication, but in simplicity and suitability to the problem addressed. Further simpli cations would come at the expense of predictive accuracy, as discussed in section 3.3.3.

There are several elaborations on TSP potentially appropriate for future work. We found that predictive accuracy increase with the number of stars in the reference set, but we truncated our parameter search at a reference set of 30,000 stars for the sake of speed. Identifying the nearest spectral neighbors, currently the most computationally expensive step, could be accelerated in a variety of ways, e.g. by making a pre-pass with the spectra in compressed form. Presumably, the presence of Li-enriched stars in our reference set prevented us from identifying a fraction of enriched stars. Removing Li-enriched reference stars from the reference set via iterative application of TSP could address this potential problem and give more complete results.

Equation (3.2) is a form of unregularized regression. Ideally, all projection weights will be small, since large weights correspond to a part of the tangent space that doesn't overlap with the

spectral manifold. We have found that this is true for our data, but a regularized form of Equation (3.2) could help ensure that is is more often the case.

In this work, we separate calculation of the imputed prediction, f^{pred}, from examining residuals to identify strong absorption. Jointly fitting for EEW simultaneously with projection onto the approximate tangent space would be more principled, and potentially more effective. Using a probabilistic or robust form of PCA (e.g. Bishop 2006 chapter 12) could give a more principled estimate of the tangent space. Other neighbor selection schemes are also possible, such as using all neighbors within a given distance hypersphere. An adaptive selection scheme could potentially adjust dynamically for the changing density of the reference set across the manifold.

In section 3.3.3, we showed that the imputed prediction in the region of the Li doublet is remarkably insensitive to the choice of k and d. In fact, we found this to be the case even when we didn't mask the Li doublet region, presumably because the number of pixels outside the doublet region dwarf the number of Li-sensitive pixels, and dominate equation 3.2. In addition, if the subspace dimension d is small, the model will have too few degrees of freedom to capture Li variation independently from other abundances.

One might ask what advantages TSP has over any established nonlinear dimensionality reduction algorithm to lossily compress and decompress spectra. While such approaches would likely work, our censor-and-predict scheme is not naturally supported by these tools. Even if censoring turns out to be unnecessary for real data, we avoid a large computational optimization problem because we don't need a global and continuous low-dimensional basis. We incorporate uncertainty during the prediction step, and the simplicity of TSP means that a justified uncertainty estimate is easy to compute.

3.5.3 What could possibly go wrong?

Here we list situations and ways in which TSP can fail to identify Li-enriched or Li-depleted stars. Visualizing all steps of the analysis is the best way to diagnose these kinds of problems.

- The manifold is not well-sampled. If there are not enough points in the reference set, the

: -nearest points may “jump across a wrinkle in the manifold”. This will result in a poor t to both the masked and unmasked pixels.

- The manifold is sampled very non-uniformly. In general, the reference set will not be uniformly distributed across the manifold, meaning that the above issue can arise in some regimes, but not others. Using an adaptive method to pick the neighborhood size may help in these situations.

- The mask is wrong, i.e. the abundance effects the spectrum in many places. If the abundance affects stellar spectra in regions outside of the mask, TSP may be able to predict the masked pixels even of chemically aberrant stars. This may result in identi ed outliers disappearing for large subspace dimension, when the model has enough degrees of freedom to capture ¹- º as an independent factor of variation.

- The line model is wrong. If the line location or pro le are wrong, some enriched stars may fail to be identi ed, and EEWs may be misestimated. This will typically be identi able by visualizing residuals, and best- t line models, U m.

- The contaminant models are incomplete or too eager. Similarly, an unaccounted-for or misspeci ed contaminant can cause stars to be misclassi ed. Again, visualization is the best way to identify this situation.

- The method underperforms relative to a supervised model. Finally, a supervised data-driven model, or a completely ab-initio physical model will be a more appropriate choice, in cases where the physical models are fast and accurate or training labels are abundant, precise, and accurate.

We believe that this work is not hampered signi cantly by any of these issues. Visualization indicates that the model ts the data closely and can accurately predict held-out data. The tests described in section 3.3.3 shows that the typical predictive accuracy is better than the nominal

measurement uncertainty. The Li doublet's pro le is dominated by instrumental dispersion, and its pro le is well-described by a Gaussian, and its wavelength precisely known.

3.6 Conclusions

Our chief ndings are the following:

- We introduce TSP, a method for imputing data using ideas from manifold learning which can be used to identify stars enriched in a given element from their spectra without a physical model.
- We apply TSP to the 6708 Å Li doublet in LAMOST, identifying 8,428 Li-enriched stars.
- We examine the abundances of Li-enriched stars near the 1st subgiant branch, the regime where Li-enrichment is thought to be most likely to be due to planet accretion. We nd that Li-enriched and Li-normal subgiants have nearly identical individual abundance distributions in the 10 elements examined and are thereby not distinguished by any signature of potential engulfment in other abundances.
- We do not see a sharp increase in the fraction of Li-enriched red giants at solar [Fe/H] and we identify few Li-enriched red giants with log g above that of the red clump. Furthermore, we see no surplus of Li-rich giants on the red clump, in contraction to prior studies. This suggests that TSP are sensitive to a different set of enrichment mechanisms than abundance-based searches.
- Using observations of stars in the Pleiades, we demonstrate that we reliably identify young main sequence stars with T_{eff} » K¼ 5750 as Li-enriched.

Chapter 4: Chemodynamical signatures of bar resonances in the Galactic disk: current data and future prospects

4.1 Introduction

Gaia (Gaia Collaboration et al. 2016), especially data release 2 (DR2; Gaia Collaboration et al. 2018c), which included radial velocities, has revealed the Galactic disk to be dynamically complex (Gaia Collaboration et al. 2018a; Antoja et al. 2018). Many perturbers are at play, and the speci c in uence of each is not yet clear. The Galaxy's bar (e.g. Hunt and Bovy 2018; Fragkoudi et al. 2019a), spiral structure (Quillen et al. 2018; Kawata et al. 2018; Hunt et al. 2018; Antoja et al. 2018; Sellwood et al. 2019; Khanna et al. 2019; Khoperskov et al. 2020), and satellite galaxies (Laporte et al. 2019; Khanna et al. 2019) all affect disk kinematics. Furthermore, Hunt et al. 2019 showed that the signatures of each are dif cult to disentangle and that many interpretations of the data are possible. Hilmi et al. 2020 showed that perturbers can interact signi cantly, further complicating matters.

In this work, we focus on the chemical signatures of the bar in action space, and demonstrate that chemical abundances can aide the identi cation of bar resonances in the disk and constrain the bar pattern speed, $_p$. Because most abundances do not change over a star's lifetime, they trace its birth location, providing a non-dynamical “memory”, that can help distinguish the effects of disk perturbations. However, stars on orbits in resonance with the bar are not uniformly distributed in azimuth. The spatial selection function and limited footprint of current data sets therefore strongly shapes the observed and expected abundance signatures, as we will discuss in this work. Past work on chemistry and dynamics has examined the distribution of metallicities and ages near the Sun or throughout the disk to constrain radial migration (e.g. Roškar et al. 2008b; Schönrich and Binney 2009; Loebman et al. 2011; Hayden et al. 2015; Martinez-Medina et al. 2016; Loebman et al.

2016; Khoperskov et al. 2019; Frankel et al. 2020). Recently, Chiba and Schönrich 2021 examined the chemical signature of a decelerating bar.

The existence of the Milky Way's bar was rst inferred from gas velocities (e.g. de Vaucouleurs 1964; Cohen and Few 1976; Liszt and Burton 1980), then with photometry (Blitz and Spergel 1991; Weiland et al. 1994) and star counts (e.g. Nakada et al. 1991; Whitelock and Catchpole 1992; Weinberg 1992). Over time, its pitch angle has been constrained, with the modern consensus being that the bar is ahead of the Sun by roughly 27 (Dwek et al. 1995; Binney, Gerhard, and Spergel 1997; Stanek et al. 1997; López-Corredoira, Cabrera-Lavers, and Gerhard 2005; Rattenbury et al. 2007; Cao et al. 2013; Wegg and Gerhard 2013). The bar's pattern speed, however, is still disputed. Identifying the Hercules moving group, a feature of the local velocity distribution, with the Outer Linblad Resonance yields a pattern speed of roughly 55 km s^{-1} kpc^{-1} (Dehnen 1999; Dehnen 2000, later Chakrabarty 2007; Minchev, Nordhaus, and Quillen 2007; Minchev et al. 2010; Antoja et al. 2014; Fragkoudi et al. 2019b). On the other hand, studies of the inner Galaxy suggest that the bar is slower, with a pattern speed of 30–40 km s^{-1} kpc^{-1} (Weiner and Sellwood 1999; Rodriguez-Fernandez and Combes 2008; Long et al. 2013; Wegg, Gerhard, and Portail 2015; Sormani, Binney, and Magorrian 2015), and that Hercules could be due to the CR resonance (Pérez-Villegas et al. 2017; Monari et al. 2019a; Monari et al. 2019b; Binney 2020). Recent studies suggest a gure around the upper end of this range, and all but rule out a faster pattern speed (Portail et al. 2017; Clarke et al. 2019; Sanders, Smith, and Evans 2019; Monari et al. 2019b; Bovy et al. 2019; Moreno et al. 2021; Kawata et al. 2021).

While the true potential of the Milky Way is complex, asymmetric, and time-dependent, we can use an axisymmetric approximate potential model to calculate "axisymmetric actions exti-mates", $_q - ' - |$, for each star, associated with azimuthal, radial, and vertical motion, respectively (e.g. Binney and Tremaine 2008, chapter 3). From here on, we refer to axisymmetric action es-timates simply as actions, which are always calculated using axisymmetric potentials, even when applied to non-axisymmetric simulations. Unlike true actions, which are conserved but not gener-ally known, these change as the star orbits in a non-axisymmetric potential (e.g. Lynden-Bell and

Kalnajs 1972; Sellwood and Binney 2002; Binney 2018; Trick et al. 2021). The fact that they are not true actions does not pose a problem in this context, nor does this generate a mismatch between simulations and data Our axisymmetric action-angles provide a valid coordinate system for the orbits, that is analogous to other reference frames, for example ¹'n?8– I– E'– E? 8–º. For ¹A0– 342– 3– ?<A0– ?<342º. Stars on orbits resonant with the bar and with small librate around negatively-sloped lines in the– q plane, with ' decreasing with increasing q along them (Sellwood 2010). We refer to these as axisymmetric resonance lines.

The primary way the bar impacts the disk is through resonant trapping, which affects stellar orbits with frequencies commensurate with the bar, i.e. for which

$$< {}^{1} \quad p \qquad q^{o} = ; \quad ' \quad - \tag{4.1}$$

where < and ; are small integers with < ¡ 0, and q and ' are the frequencies of azimuthal and radial oscillation (approximately equal to the frequencies following from the axisymmetric Hamiltonian), respectively. These orbits librate around their parent orbits, which are closed in the frame that rotates with the bar (Contopoulos and Grosbol 1989; Athanassoula 1992). Because of the conservation of the Jacobi integral, changes to the axisymmetric actions of stars obey

$$' = \frac{;}{<} \quad q \tag{4.2}$$

to rst order (Lynden-Bell and Kalnajs 1972; Sellwood and Binney 2002). Consequentially, corotating (0:1, i.e.; = 0, < = 1) stars can undergo changes to their angular momentum, L_q, without being radially heated. Most stars are on near-circular orbits, meaning that scattering at the outer 1:1, 1:4, and 1:2 Linblad resonances preferentially increase eccentricity, e_r, inducing a net outward migration and leading to under- and over-densities on either side of the axisymmetric resonance line. See Trick et al. 2021 for a pedagogical introduction and numerical investigation of these effects.

In sections 4.2 and 4.3 we describe the observational data and simulations we use, respectively.

Figure 4.1 The radial distribution of stars within 1 kpc from the Galactic midplane in each input catalog, as well as in the crossmatched joint catalog. LAMOST radial velocities can fruitfully augment Gaia in the outer disk.

In section 4.4, we show how bar resonances give rise to distinct chemical signatures and compare them to observations. In section 4.5, we discuss the impact of spatial selection functions and consider the capabilities of future surveys. We conclude in section 4.6.

4.2 Data

We use astrometry from Gaia EDR3 (Gaia Collaboration et al. 2020), which includes positions, parallaxes, and proper motions for 1.3×10^9 sources, as well as radial velocities for 7.2×10^7 stars (hereafter the Gaia radial velocity sample; these are the same as in DR2). We also use the LAMOST (Cui et al. 2012; Zhao et al. 2012; Deng et al. 2012) DR 5 FGAK sample, which provides radial velocities and metallicities for 4.2×10^6 stars. We use distance estimates from Bailer-Jones et al. 2018, which used a Galactic prior together with Gaia parallax to obtain distance posteriors for stars from both catalogs.

Figure 4.1 shows the radial distribution of LAMOST and Gaia radial velocity stars. Interior to the Sun, there are far more Gaia radial velocity than LAMOST stars, and future Gaia data releases will include more radial velocity measurements. But at present, LAMOST has observed a similar

number of stars at beyond the Sun, many of which where not observed by Gaia radial velocity. Beyond $' = 10$ kpc, LAMOST nearly doubles the number of stars with radial velocities available. Despite the fact that the LAMOST has a larger typical radial velocity uncertainty (5 km s^{-1}, as opposed to Gaia's 1 km s^{-1}), it usefully augments Gaia in the outer disk.

We used `galpy` (Bovy 2015) to calculate actions for each star with the Stäckel fudge (Binney 2012; Bovy and Rix 2013) in the axisymmetric `MWPotential2014` model (Bovy 2015). The `MWPotential2014` assumes the solar Galactocentric radius to be $_0 = 8$ kpc and the local standard of rest to be $E_{circ}{}^{1}' {}_0{}^{o} = 220$ km s^{-1}. While more recent MW models suggest slightly different values (e.g. Gravity Collaboration et al. 2018), they do not have a strong impact on our results. This is illustrated in Appendix B.2, where we show our central observational plot with actions calculated using the potential from Eilers et al. 2019. For the solar proper motion, we use $E_{\cdot} = 11{\bullet}1$ km s^{-1}, $E_{\cdot} = 232{\bullet}24$ km s^{-1}, and $E_{\cdot} = 7{\bullet}25$ km s^{-1} by Schönrich and Binney 2009. While kinematic uncertainties propagate to actions, they are small enough not to obscure the chemodynamical signatures we investigate in this paper. Wheeler et al. 2020, Appendix D show, for example, mean action uncertainties as a function of Galactocentric radius for Gaia DR2 and LAMOST DR 4, which have radial velocity uncertainties very similar to Gaia EDR3 and LAMOST DR 5. In addition, we demonstrate in Appendix B.1 that our results stay robust when applying simulated observational errors to one of our test-particle simulations.

In order to focus the investigation on near-circular in-plane disk orbits, we used only stars with $_{\cdot} \ddot{Y} 10$ kpc km s^{-1} and $|_{\cdot}| \ddot{Y} 1$ kpc. This selection is dominated by thin disk, low-alpha sequence stars. We also removed the roughly 1% of stars with $_{\cdot}$ outside $1000 - 3000$ kpc km s^{-1} to facilitate comparison with simulations. In order to mitigate magnitude-driven selection effects, we use only stars with $2{\bullet}2 \ddot{Y} \log g \ddot{Y} 2{\bullet}7$ and $4500 \text{ K} \ddot{Y} {}_{e} \ddot{Y} 5000$ K when dealing with abundances, which often have stellar-parameter-dependent systematics. This selection corresponds roughly to the red clump. These cuts leave roughly 2×10^{6} stars when abundances are not needed, and 9×10^{4} when they are. For stars in LAMOST and Gaia radial velocity, the radial velocity with the lower uncertainty is used (this is Gaia radial velocity for about 99% of stars). The mean radial velocity

uncertainty for the sample with all cuts, including those in and $\log(\sigma_e)$, is 3.2 km s^{-1}, with roughly half of radial velocities coming from *Gaia* and from LAMOST.

4.2.1 The MW disk as seen in this dataset

Figure 4.2 shows the *Gaia* + LAMOST $\Omega_b - E_q$ and $\Omega_q - \Omega_b$ planes in stellar density, fraction of inwards-moving stars, mean [Fe/H] and mean [Mg/Fe], which is taken from Wheeler et al. 2020, who estimated detailed abundances for LAMOST DR4. Note that [Fe/H] and [Mg/Fe] increase and decrease, respectively, with increasing Ω_b, Ω_q, and E_q. While we make use of [Fe/H] only in the remainder of this paper, we note that the results presented here can in principle be generalized to other abundances. The ARLs for the OLR of Ω_b = 33.5 km s^{-1} kpc^{-1}, 39 km s^{-1} kpc^{-1}, and 51 km s^{-1} kpc^{-1} are plotted for orientation. In the $\Omega_b - E_q$ plane, contours of constant angular momentum are overplotted, corresponding to the values of Ω_q on these OLR axisymmetric resonance lines at Ω_b = 0. These pattern speeds, used throughout the paper, correspond to a fast bar whose OLR is associated with Hercules (51 km s^{-1} kpc^{-1}), an intermediate bar (39 km s^{-1} kpc^{-1}), and a slow bar (33.5 km s^{-1} kpc^{-1}). The intermediate bar and slow bar OLR ARLs align with the "Sirius" and "hat" ridges, respectively, in the $\Omega_q - \Omega_b$ plane. The fast bar OLR axisymmetric resonance line is slightly to the left of the inward-moving ridge that contains the "Horn" stars, and to the right of the outward-moving "Hercules" ridge with which the OLR of a rapidly rotating bar is often associated. The exact values of the slow bar and fast bar pattern speeds are chosen so that their OLR axisymmetric resonance lines align with transitions between net inwards and outwards movement with increasing Ω_q, as proposed by Trick et al. 2021.

In the $\Omega_b - E_q$ plane, the ridges identified with *Gaia* (Kawata et al. 2018; Gaia Collaboration et al. 2021) and attributed variously to the bar, spiral arms, the Sagittarius dwarf galaxy, or a combination thereof (Antoja et al. 2018; Khanna et al. 2019; Bland-Hawthorn et al. 2019; Fragkoudi et al. 2019a; Fragkoudi et al. 2019b; Laporte et al. 2019) are clearly visible. The largest of these can be seen extending to a larger Galactocentric radius, R, than with *Gaia* data alone.

Figure 4.2 The LAMOST $_q$– $\cdot$ and $'$ –E_q planes in stellar density (top row; 10^6 stars), fraction of inwards-moving stars (second row; 10^6 stars), mean [Fe/H] (third row; 10^5 stars) and mean [Mg/Fe] (bottom row; 10^5 stars). The OLR axisymmetric resonance lines for the three pattern speeds we consider are overplotted on the $_q$– $\cdot$ plane, and contours of constant $_q$ corresponding the circular OLR orbits for the same pattern speeds are overplotted on $'$ –E_q plane. The top four panels contain 800–000 stars and the bottom four contain 95–000, because we made additional cuts to mitigate abundance systematics. We use fewer stars when plotting metallicities or abundances because we restrict our sample to a small part of $_e$ space to minimize systematics.

4.3 Simulations

To better understand the connection between bar resonance and chemistry in action space, we use test particle simulations with the same pattern speeds introduced in section 4.2.1 (Ω_b = 33.5 km s^{-1} kpc^{-1}, 39 km s^{-1} kpc^{-1}, and 51 km s^{-1} kpc^{-1}). The setup of the simulations is analogous to those used by Trick et al. 2021 and Trick 2022, starting with an axisymmetric disk (Binney and McMillan 2011) of massless particles, followed by integrating stellar orbits in an analytic MW potential model (Bovy 2015) that slowly introduces a bar perturbation (Dehnen 2000; Monari et al. 2016). The boxy bar model has $m = 2$ and $m = 4$ components (Hunt and Bovy 2018) with strengths $U_{m=2} = 0.01$ and weak $U_{m=4} = 0.0005$. Orbits are integrated for 50 bar periods and the simulation snapshots at 20, 30, 40, and 50 bar periods are stacked to increase the number of particles to 4×10^7. Axisymmetric approximate actions are calculated exactly as for the observational data (see Section 4.2 for details).

We assign metallicity values to test particles based on their positions in the intial disk, with a metallicity decreasing with both Galactocentric cylindrical radius R, and height from the midplane, $|z|$. Each particle was assigned an iron abundance according to

$$[Fe/H] = -0.09 \left(\frac{R}{\text{kpc}} - 6\right) - 0.2 - 0.21\frac{|z|}{\text{kpc}} + Y \tag{4.3}$$

where Y is a random Gaussian perturbation with standard distribution 0.1, roughly the mean uncertainty of the LAMOST [Fe/H] values. The R and z dependence of this relation are based on the trends observed by Hayden et al. 2014 in the outer low-alpha disk. We note that the qualitative results of this work do not depend on the details of how metallicities are assigned, so long as the dominant trend is decreasing metallicity with Galactocentric radius (in Appendix B.3 we show that assigning metallicities based on initial action does not change our results).

4.4 Results

4.4.1 Metallicity signatures of bar resonances in action space

While bar resonances can change J_q and J_ϕ, they can not change a star's atmospheric metallicity, which remains very-nearly constant over a star's lifetime. As discussed in Section 4.1, actions oscillate in speci c proportions at resonances, so kinematically distinguishing between resonances in observational data is possible in principle. Stellar metallicities, however, provide an orthogonal source of information that makes this task much easier. Figure 4.3 demonstrates how bar resonances lead to signatures in J_q–[Fe/H] trends at different J_ϕ. The left column shows test particles evolved in an axisymmetric potential and the right shows particles evolved in the barred potential discussed in Section 4.3. Panels A and B show the H–θ plane in metallicity, and panels C-F show the J_q–J_ϕ plane in both metallicity and density. The bottom two panels show the [Fe/H] trends as a function of J_q, in bins of J_ϕ (panel G shows absolute metallicities, and panel H shows differences with respect to the smallest J_ϕ bin). The error bars in the bottom two panels show the standard error, the scatter in metallicity over the square root of number of stars in each bin. Except at the largest values of J_ϕ and J_q, the standard errors are too small to be visible, indicating that the trends are not due to sampling errors and that the effect of Υ in (4.3) has been washed out. Unbiased measurement error in the observational data will be similarly washed out. The colored bands in panels G and H show the range of J_q values each axisymmetric resonance line takes on below 160 kpc km s^{-1}. The J_ϕ bin boundaries used in panels G and H are shown as horizontal lines in panels C-F.

Because stars in corotation with the bar oscillate strongly in J_q without changing coherently in J_ϕ (Equation 4.2), and density changes slowly with J_q, the effect of the corotation resonance is dif cult to see in stellar density in the J_q–J_ϕ plane. Contrasting the mean [Fe/H] around the CR axisymmetric resonance line for the barred and unbarred simulations reveals that the metallicity gradient becomes shallower when the bar causes stars to librate symmetrically in J_q, washing out the initial metallicity gradient. This attening is easily visible in panel G, where the metallicity

Figure 4.3 The metallicity signatures at the bar resonances in the intermediate bar test particle simulation. A and B: mean metallicity in the meridional plane for the axisymmetric and barred test particle simulations, respectively. C and D: number density in the $_q - \ $ plane for the axisymmetric and barred test particle simulations, respectively. E and F: mean [Fe/H] in the $_q - \ $ plane for the axisymmetric and barred test particle simulations. G: mean [Fe/H], in bins of $_q$ and $\ $, for the axisymmetric (dashed) and intermediate bar (solid) test particle simulations. The CR resonance flattens the mean-metallicity trends, while the higher-order resonances (especially the OLR) create wave-crest signatures. H: differences between the above: the difference between mean [Fe/H] of each $\ $ bin and the lowest $\ $ bin. Error bars in the bottom two panels show the standard error in each bin. Except at the largest values of $\ $ and $_q$, the standard errors are too small to be visible, indicating that the trends are robust. The colored bands show the range of values each axisymmetric resonance line takes on between $_q$ = 160 km s^{-1} kpc^{-1}. See also Appendix B.1.

gradient is visibly attened.

At the other resonances (OLR, 1:1, and 1:4), changes in lockstep with f_q (equation 4.2). Because of the strong density gradient in f (and weaker gradient in f_q), more stars are on average displaced from near-circular, low- f orbits to eccentric, high- f orbits than vice versa, creating a density ridge to the right of the corresponding axisymmetric resonance line (panel D). Below approximately f = 50 kpc km s^{-1}, the metallicity contours to the right of the OLR axisymmetric resonance line are more vertical in the barred simulation because higher-metallicity stars on circular orbits have moved upward (and slightly to the right) in the f_q–f plane, shifting the mean metallicity up at the location of the overdensity ridge. In panels G and H, the chemical trends around the OLR are most easily identi able. For near-circular orbits (f » kpc km s^{-1} ¼ Ÿ 15) in the barred potential, the mean metallicity drops slightly below that in the unbarred potential. This is because of the small number of lower-metallicity stars scattered from larger f_q and f to more circular orbits. In contrast, among the high f stars, the deviation from the overall metallicity trend rises above that of the unbarred potential to form a “wave-crest” in the f_q–[Fe/H] plane across bins of f. Like the OLR, the 1:1 and 1:4 resonances create metallicity wave-crests, but they are weaker because of the weaker Fourier m = 4 bar component, and the 1:1 comprising stars far beyond the bar.

The qualitative wave-crest and attening signatures at OLR and CR, respectively, in Figure 4.3 do not change with pattern speed. Figure 4.4 shows the metallicity trends for particles evolved in all three simulations. While the value of f_q at which the resonances occur changes, the chemical signature of the resonances does not. They do change with bar strength, however. As the bar strength increases, the region of trapped orbits around the axisymmetric resonance line becomes larger (c.f. Binney 2018). Figure 4.5 shows the attening and wave-crest for the bar simulation, and for stars evolved in the same potential but with a bar of twice the strength ($U_{<=2}$ = 0•02, $U_{<=4}$ = 0•001). The attening in [Fe/H] at CR as a function of f_q is more prominent and extends across a larger range of f_q. Likewise, the wave-crest at the OLR and 1:1 resonance becomes more extreme, with the mean metallicity for high- f stars exceeding that of those on circular orbits in

Figure 4.4 Mean [Fe/H] as a function of in bins (same as Figure 4.3) of for particles evolved in all three simulations. Error bars show the standard error. The qualitative attening and wave-crest signatures are not sensitive to pattern speed. The vertical bands show the location of axisymmetric resonance lines, as in Figure 4.3.

Figure 4.5 Mean [Fe/H] as a function of η in bins (same as Figure 4.3) of for particles evolved with two different bar strengths, but the same pattern speed. Error bars show the standard error. Because the stronger bar traps more stars on resonant orbits, the metallicity signatures at CR and the OLR cover a larger range in η, and are therefore also stronger at larger

the resonance zone, and the effect extending across a wider range of $_q$ for the hottest bins.

4.4.2 Action-metallicity signatures in the LAMOST data

Figure 4.6 shows mean [Fe/H] as a function of $_q$ and $'$, the diagnostic introduced in section 4.4.1, drawn with the observational data introduced in Section 4.2. The three panels show axisymmetric resonance line zones corresponding to the three pattern speeds used in the test particle simulations inspired by values in the literature. The limitations of available data (the number of stars observed, their footprint coverage, and systematics in their metallicities) make it dif cult to unambiguously identify either the attening or wave-crest signatures. Only for $_p$ = 33•5 km s^{1} kpc^{1} are both the CR and the OLR region populated with data, although this is sensitive to the rotation curve of the approximate axisymmetric potential used to calculate actions. Increasing incompleteness with distance from the Sun yields mean-[Fe/H] uncertainties large enough to obscure potential resonance signatures for $_q$. 1600 kpc km s^{1} and $_q$ & 2700 kpc km s^{1}, even though some data is present. As we will discuss in section 4.5, the azimuthal coverage of the LAMOST and Gaia data also obscure a potential signature of bar resonances. Appendix B.4 brie y discusses the status of the wave-crest and attening signatures in APOGEE.

In the context of the above caveats, we note that a possible wave-crest is present in Figure 4.6 around $_q$ 2200 kpc km s^{1}, i.e. close to the Hat moving group. This corresponds to the OLR for $_?$ = 33•5 kpc km s^{1}. There is a attening in the mean [Fe/H] trend at 1400 kpc km s^{1} . $_q$. 1800 kpc km s^{1} (6•5 kpc . $'$. 8•55 kpc), which corresponds to CR for this pattern speed. Similar trends have been observed before in the Milky Way (Hayden et al. 2015; Haywood et al. 2018; Haywood et al. 2019) and in other galaxies (Stott et al. 2014; Leethochawalit et al. 2016), and have been largely attributed to well-mixed gas, although Hayden et al. 2015 states that stellar migration due to the bar is the likely cause. This atting is mildly attenuated by our cuts in $\log g$ and $)_e$, suggesting that a more thorough understanding of metallicity systematics may shed light on it. If the attening is caused by stellar migration, rather than the birth properties of stars in the inner disk, it is consistent with the CR resonance of a slow (38•5 kpc km s^{1}) bar. At

Figure 4.6 Mean [Fe/H] as a function of $_q$ in bins of $\cdot$ for LAMOST. The errorbars show the standard error on the mean. There are not enough stars with $_q$ 60 kpc km s^{1} to plot trends. Grey vertical bands mark the approximate $_q$ of known moving groups for orientation. Gray dashed lines mark the $_q$ values of circular orbits with $\dot{}$ at the radius where the solar azimuth intersects hypothesized spiral arms (taken from Reid et al. 2014). Observational data is not yet plentiful enough to clearly identify or rule out the resonance signatures described in section 4.4.1, but the attening around 1400 kpc km s^{1} Ÿ $_q$ Ÿ 1700 kpc km s^{1} and the elbow at $_q$ 2200 kpc km s^{1} are suggestive of a slow bar.

$_q$ 1850 kpc km s^{1}, there is a small bump in the lowest bin which can be interpreted as corresponding to the 1:4 resonance for this pattern speed.

While the resonance locations for $_? = 33.5$ km s^{1} kpc^{1} are fairly consistent with the observational data, those for the other two pattern speeds are much less so. For $_? = 39.5$ km s^{1} kpc^{1}, neither the OLR no 1:4 resonance correspond to signature in the [Fe/H] trends, and the error bars at the 1:1 resonance are large enough to mask the weak signature expected there. For $_? = 51$ km s^{1}, the 1:1 resonance corresponds to a possible wave-crest, but the detailed trends at the OLR are masked by the error bars, and the attening discussed above is not explained. We conclude that a slow bar, ($_? = 33.5$ km s^{1} kpc^{1} or slightly faster) is most consistent with the observational data.

The dependence of mean [Fe/H] on differs strongly across $_q$. For $_q$. 2000 kpc km s^{1} (at least until $_q$ 1600 kpc km s^{1}, when the uncertainties become large), stars on higher-eccentricity orbits have lower [Fe/H]. For $_q$ & 2000 kpc km s^{1} the dependence is much weaker. This effect is not a direct consequence of the dependence of [Fe/H] on height from the disk, surface gravity ($\log$ 6, or effective temperature $_e$). It also does not appear to be caused by "contamination" from the high- disk, since removing stars with large [Mg/Fe] (per Wheeler et al. 2020) does not have any effect. Additionally, the dependence of mean [Fe/H] exhibits the same behavior when Figure 4.6 is made with APOGEE (Majewski et al. 2017) DR16 (Ahumada et al. 2020), rather than LAMOST data (see Appendix B.4). We conclude that variation in correlation of mean [Fe/H] with is likely physical, rather than an artifact of the observational data or our analysis.

4.5 Discussion

The bar and spiral arms have been shown to lead to a nonaxisymmetric metallicity distribution (e.g. Di Matteo et al. 2013; Grand et al. 2016; Khoperskov et al. 2018; Fragkoudi et al. 2018), meaning that the spatial location of the observational data with respect to the bar has a signi cant impact on both the wave-crest and attening signatures. In contrast to previous Figures, which use

Figure 4.7 The location of observed stars (i.e. the spatial selection function) has a strong impact on the attening and wave-crest signatures. Here, resonance signatures as exhibited by stars in different azimuthal regions of the disk in the intermediate bar simulation are plotted. bins above 60 kpc km s^{1} are not plotted because of the lack of suf cient data. Each row shows stars within a cylinder centered at the solar Galactocentric radius at different azimuths: exactly along the minor axis of the bar (top), in the position of the Sun (middle rows), and exactly along the bar's major axis (bottom). The left column shows number density in the plane, i.e. the footprint of each mock data set in the disk. The white line marks the orientation of the bar, and the white star marks the location of the Sun. The right column shows metallicity trends in and for each region. In row 3, only stars with azimuth farther from the bar azimuth than the Sun are used. The CR attening and the OLR wave-crest are much more prominent in stars anti-aligned with the bar, but could be seen in a large volume of local data by plotting only stars farther from the bar's azimuth than the Sun (row 3).

stars at all azimuths, Figure 4.7 shows mean [Fe/H] trends in $_q$ and $\cdot$ using only stars within cylinders centered at different points in the disk for the intermediate bar test particle simulation, demonstrating that stars with azimuth nearer to the bar's minor axis show the resonance signature much more prominently.[1] All cylinders are centered on points 8 kpc from the center of the disk, the approximate Galactocentric radius of the Sun. The rst row, which plots stars centered on a point on the bar's minor axis (i.e. with $_q$ = q 75 , taking the bar's pitch angle to be 25), exhibits resonance signatures much stronger than the others. For these stars, the attening at CR turns sharply on and off for each bin, and the OLR wave-crest is very strong, with mean [Fe/H] dependence on fully inverting at $_q$ 2250 kpc km s^{1}. For survey volumes centered on the Sun or aligned with the bar's major axis (rows 2 and 4 in Figure 4.7), the CR attening and OLR wave-crest are attenuated and nearly absent, respectively. But if only the stars within 3 kpc of the Sun which are farther from the bar azimuth than the Sun are plotted, both signatures become much more prominent (Figure 4.7 row 3), demonstrating that a sample in the "extended solar neighborhood" can be used to uncover these features.

Figure 4.8 shows the degree of CR attening as a function of position in the disk with two CR orbits plotted, demonstrating again the importance of the azimuth of observed stars and helping to reveal its cause. It is colored by the slope of the [Fe/H] relation for stars roughly in the inter-mediate bar CR region, with 1200 kpc km s^{1} Ÿ $_q$ Ÿ 1300 kpc km s^{1} and $\cdot$ Ÿ 60 kpc km s^{1}, minus the slope outside the CR zone (calculated with the stars 1500 kpc km s^{1} Ÿ $_q$ Ÿ 1800 kpc km s^{1}). This gure shows, similarly to Figure 4.7, that the CR attening is much stronger in stars around the bar's minor axis.[2] The two grey lines show the paths of two stars trapped by the CR resonance as they orbit over 3000 time steps. The azimuthal dependence is a consequence of the "banana" shape of these orbits, which oscillate around Lagrange points 4 and 5 on the bar's minor axis.

We use a simple rejection sampling scheme to simulate the LAMOST spatial selection function

[1] Particles with positive are rotated 180 to the Sun's side of the Galaxy to increase the number available to plot.

[2] There is a weak asymmetry in the metallicity slope at CR around the bar's minor axis, at about 4 kpc, slightly ahead of the minor axis in Figure 4.8. This is a phase-mixing relic in the simulation which depends on how stars with different initial phases and metallicities get scattered at CR.

Figure 4.8 The “banana-shaped” orbits of stars trapped at CR spend little time near the bar's major axis, and many dwell exclusively near its minor axis, which explains the azimuthal sensitivity demonstrated in Figure 4.7. The relative slope of the [Fe/H] relation around CR with respect to the overall slope for stars with $\dot{Y}$ 60 kpc km s^{1} as a function of position in the disk for the intermediate bar simulation. The black rectangle indicates the orientation of the bar. The orbits of two stars trapped by the CR resonance are overplotted (in the rotating frame).

Figure 4.9 Mean [Fe/H] as a function of R_0 and ϕ for the intermediate bar simulation with the LAMOST's spatial selection function imposed. top: full LAMOST footprint. bottom: partial LAMOST footprint, with only stars with azimuth less than the Sun's. The flattening and wave-crest signatures at CR and the OLR, respectively, are extremely attenuated because stars are concentrated near the Sun.

in detail and understand why current data is inadequate to reveal the signatures explored in this work. We approximate the spatial distribution of the test particles and LAMOST stars, $?_{sim}(r)$ and $?_{LAMOST}(r)$, where r is 3D position, by counting the fraction of stars in each bin in a $n$ne grid[3], then interpolating with a cubic spline. We then include each test particle in the subset for comparison with probability proportional to the ratio of $?_{sim}(r) \cdot ?_{LAMOST}(r)$. This reduces the number of particles from 4×10^7 to about 1.6×10^5, but allows a more "apples to apples" comparison between simulation and data. Figure 4.9 shows the mean [Fe/H] trends in action space for the intermediate bar simulation with the observational spatial footprint imposed. (The results are similar for the slow bar and fast bar simulations.) Both the wave-crest and $f$attening are attenuated and distorted by the limited azimuthal coverage. The $f$rst row uses the full observational footprint, and the second row uses the same cut as Figure 4.7 row 3, keeping only stars with azimuth closer to the bar's minor axis than the Sun. Both selections are too concentrated near the Sun to show the signatures clearly. While LAMOST and other current datasets don't have the spatial coverage to look for resonance signatures in stars further away from the bar's major axis, in the near-future metallicities from the Gaia radial velocity spectrograph or Sloan-V (Kollmeier et al. 2017) will likely make such an analysis possible.

4.6 Conclusions

We investigate the effect of bar resonances on the metallicity distribution in the Milky Way's thin disk as a function of the stars' orbital actions J_r, J_ϕ, and J_z. We use test-particle simulations of a stellar disk in a Milky-Way-like potential with a simple rotating bar to identify distinct signatures: the corotation resonance leads to a $f$attening of the metallicity gradient, and higher-order resonances, especially the Outer Lindblad resonance; this creates a "wave-crest" in the mean [Fe/H] as a function of J_r binned by J_ϕ. We con$r$m that these signatures do not change qualitatively with pattern speed or strength, although the strength of the signal varies with bar strength. We demonstrate that a survey's selection function can strongly affect the appearance of these signatures.

[3] Bin sizes are 0.5 kpc in the x and y dimensions, and 0.1 kpc in the z dimension

We search for these resonant metallicity-vs.-action features in stars using metallicities and radial velocities from LAMOST DR 5 and parallaxes, proper motions, and supplementary radial velocities from *Gaia* EDR3. While distinguishing signatures of bar and spiral resonances remains dif cult, the diagnostics discussed here can identify both the OLR and CR resonance at once. We nd weak evidence for a slow bar with an OLR associated with the “hat” moving group, but conclude that present data does not allow us yet to unambiguously identify these signatures. We do not know how strong the real bar in the MW is, and also do not know if and by how much the quoted measurement uncertainties are underestimated. However, even by exploring different combinations, we did not manage to produce an exact match of the model to the data. This could indicate that additional mechanisms might be at play in the Galactic disk (pattern speed variations, spiral arms, other mixing mechanisms. etc.) or that the relationship between metallicity and birth radius is not captured will by either of the models we used. But the investigation of their exact in uence on the bar's action-metallicity pattern is beyond the scope of this paper.

Interpreting observational data is made more dif cult, not only by the complexities of the Galaxy not present in our simpli ed test particle model, but also because current surveys provide 6-D kinematics and metallicities only for stars in a limited region of the Galactic disk. In particular, we show that the strength of the attening and wave-crest signatures depends on the Galactocentric azimuth. The strongest signature is found at an azimuth along the bar's minor axis, while along the bar's major axis the resonant signature in the metallicity almost vanishes. Consequently, our observing location within the Galactic disk,25 behind the bar, is not ideal, but observations of more stars with Galactocentric azimuth far from the major-axis of the bar, which near-future data will deliver, can remedy this.

There are several promising directions for future progress on chemo-dynamically identifying bar resonances. In this paper, we have used a fairly simple galaxy model as basis for the comparison to the data. More complex models may yield modi ed predictions. For example, a bar with signi cant deceleration will trap stars with a different distribution of orbits (Chiba, Friske, and Schönrich 2019), and Chiba and Schönrich 2021 show that metallicity should decrease from the

center of the corotation resonance to its boundary. N-body models with self-consistent bars could also provide a richer picture, including the possibly confounding effects spiral arms, which impact metallicity distributions and can overlap with bar resonances (e.g. Khoperskov and Gerhard 2021; Asano et al. 2022). Stellar ages and chemical abundances beyond [Fe/H] provide additional clues to a star's birth location and, given a better understanding of the Galaxy's chemical evolution, can doubtlessly further constrain birth radius. On the other hand, if the bar resonances can be conclusively identi ed by other means, their chemical abundances inform us on the star formation history of the disk. Sloan-V and Gaia's future data releases, which will increase the amount of suitable data and its azimuthal coverage, will likely allow us to pinpoint the pattern speed of the bar and shed light on the Galaxy's dynamical structure.

Chapter 5: Korg: a modern 1D LTE spectral synthesis package

5.1 Introduction

Improvements in instrumentation have yielded exponential growth in the amount of spectral data to analyse. Creating pipelines that can keep up with analysis is nontrivial. There are several extant codes for 1D LTE spectral synthesis, including TURBOSPECTRUM (Plez, Smith, and Lambert 1993; Plez 2012), MOOG (Sneden 1973a; Sneden et al. 2012), SYNTHE (Kurucz 1993; Sbordone et al. 2004), SME, (Valenti and Piskunov 1996; Valenti and Piskunov 2012; Piskunov and Valenti 2017), and SPECTRUM (Gray and Corbally 1994). While they have enabled a huge volume of research, these codes can be difficult to use for the uninitiated, and require input and output through custom file formats, impeding integration into analysis code. Here we present Korg, a new 1D LTE synthesis package, written in JULIA and suitable for easy integration with scripts and use in an interactive environment. As the first such code in more than two decades, Korg benefits from numerical libraries not available at the time earlier packages were authored.

Like most existing spectral synthesis packages, Korg simplifies analysis by assuming that atmospheres are in LTE—that baryonic matter is described by thermal distributions, and radiation is only very slightly out of equilibrium. The source function (the ratio of the per-volume emission and absorption coefficients) is dominated by collisions of baryonic matter and is Planckian. While non-LTE (NLTE) calculations are important for producing the most unbiased possible model spectra, they are prohibitively slow, and not yet suitable for application that require computing many spectra. Fortunately, corrections to LTE level populations (e.g. Amarsi et al. 2020), equivalent widths, or abundances (e.g. Lind et al. 2011; Amarsi et al. 2015; Bergemann et al. 2012; Osorio and Barklem 2016) can be calculated from NLTE simulations. These can be applied to LTE codes to produce approximate NLTE spectra at relatively little computational cost. Additionally, biases

from the assumption of LTE will roughly cancel for similar stars, yielding differential abundance estimates with high precision.

The performance of spectral synthesis codes is most important when tting observational data. Because synthesis must be embedded in an inference loop, the analysis of of a single spectrum may trigger tens or hundreds of syntheses. Even when many parameters may be estimated by interpolating over (or otherwise comparing to) a precomputed grid of spectra (e.g. Recio-Blanco, Bijaoui, and De Laverny 2006; Smiljanic et al. 2014; Holtzman et al. 2018; Boeche, Vallenari, and Lucatello 2021; Buder et al. 2021), individual abundances are best done with targeted syntheses of individual lines. Furthermore, generating a grid from which to interpolate can be computationally expensive. Inference and optimization can also be sped up by fast and accurate derivatives of the function being sampled or minimized, most easily produced via automatic differentiation. KORG is designed to be compatible with automatic differentiation packages (e.g. FORWARDDIFF), which can provide derivative spectra in roughly the same amount of time required for a single synthesis (as discussed in Section 5.4; see Figure 5.1).

5.2 Description of code

To synthesize a spectrum, KORG calculates the number density of each species (e.g. H I, C II, CO) at each layer in the atmosphere (Section 5.2.2), then computes the absorption coef cient at each wavelength and atmospheric layer due to continuum (Section 5.2.3) and line absorption (Section 5.2.4). Given the total absorption coef cient at each wavelength and atmospheric layer, it then solves the radiative transfer equation to produce the ux at the stellar surface (Section 5.2.5).

5.2.1 Inputs

KORG takes as inputs, a model atmosphere, a linelist, and abundances for each element in [1] form,[1] assumed to be constant throughout the atmosphere. By default, abundances follow the solar ratios from Asplund Chemical Composition Sun 2009a, and the abundances of elements

[1] [1] = log$_{10}$[1] = • = H , 12 where = is the total number density of element and = H that of hydrogen.