Title
Pushing Stellar Archaeology Farther & Fainter with Low-Resolution Spectroscopy

Author
Sandford, Nathan Ross

Publication Date
2023

Peer reviewed|Thesis/dissertation

Pushing Stellar Archaeology Farther & Fainter with Low-Resolution Spectroscopy

by

Nathan Ross Sandford

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Astrophysics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor Daniel Weisz, Chair
Associate Professor Jessica Lu
Professor Joshua S. Bloom

Summer 2023

Pushing Stellar Archaeology Farther & Fainter with Low-Resolution Spectroscopy

Abstract

Pushing Stellar Archaeology Farther & Fainter with Low-Resolution Spectroscopy

by

Nathan Ross Sandford

Doctor of Philosophy in Astrophysics

University of California, Berkeley

Associate Professor Daniel Weisz, Chair

The chemical composition of individual stars provide a fossil record of chemical evolution in a galaxy over cosmic time, encoding important galactic enrichment mechanisms, timescales, and nucleosynthetic pathways. Within the Milky Way, large spectroscopic surveys on modest telescopes have enabled precise chemical abundance measurements for millions of stars and have transformed our understanding of the formation and evolution of the Galaxy. Extending this progress beyond all but the nearest Milky Way satellite galaxies, however, is severely limited by the faintness and crowding of stars at these distances, which require large (6–10+ meter) telescopes and low-resolution spectrographs to achieve sufficient signal/noise (S/N). Recent advances in spectroscopic techniques have dramatically improved our ability to recover accurate and precise chemical abundances from low-resolution spectra. However, these techniques have not yet been extended to the extragalactic regime despite the wealth of data that has been collected in Local Group dwarf galaxies over the past decade. Moreover, with increasingly powerful and multiplexed spectroscopic facilities on the horizon, high-quality (albeit low-resolution) spectra will soon be accessible for millions of stars throughout the Local Group and out to several Mpc. These new datasets have the potential to reveal in unprecedented detail the stellar chemistry of the smallest and faintest galaxies in the Universe, which can in turn be used to investigate the physics of low-mass galaxy evolution. In this dissertation, I lay critical groundwork necessary to fully capitalize on the wealth of chemical information embedded in existing and future extragalactic resolved stellar spectra.

As new spectroscopic facilities are designed, commissioned, and begin amassing large datasets of resolved stellar spectra outside the Milky Way, it is imperative that we understand what chemical abundance information these observations contain. In Chapter 2 of this dissertation, I employ the Cramér-Rao Lower Bound (CRLB) to forecast the theoretical precision to which 41 existing and future spectrographs can measure chemical abundances in metal-poor, low-mass stars in Local Group galaxies. I demonstrate that even at low- to moderate-resolution, blue-optical spectroscopy with modest S/N enables the recovery of a dozen or more elements

to a precision of $\lesssim$0.3 dex. Additionally, I find that high-resolution stellar spectra contain substantial chemical abundance information even at low S/N, which can be extracted via full-spectrum fitting techniques. Looking to the future, I show that with reasonable integration times JWST/NIRSpec and 30-m class telescopes can recover $\sim$10 and 30 elements, respectively, throughout the Local Group and bulk metallicities ([Fe/H] and [α/Fe]) for resolved stars out to several Mpc. This analysis is paired with the development and release of an open-source python package, `Chem-I-Calc`, that facilitates similar forecasts for additional spectrographs, stellar targets, and observing conditions relevant to the astronomical community at large.

In practice, achieving the precision forecasted by the CRLB is impeded by shortcomings of the stellar models used to analyze stellar spectra, which can introduce systematic biases and uncertainties into the measurement of chemical abundances, especially at low resolutions when absorption features are heavily blended. In Chapter 3, I perform a self-consistent analysis of archival Keck/HIRES spectra of low-metallicity stars in M15, which I convolve to lower resolutions in order to quantify the resolution dependence of systematics introduced by model-data mismatches. I demonstrate that systematic biases and uncertainties remain small ($\sim$0.1 dex) for 20 (9) elements down to $R \sim 10,000$ (2500). This analysis illustrates the great promise of low-resolution spectroscopy for stellar chemical abundance measurements in extragalactic systems. As stellar models and spectroscopic fitting techniques improve, the viability of low-resolution resolved stellar spectroscopy will further expand.

Ultra-faint dwarf galaxies (UFDs) represent some of the oldest, lowest mass, most metal-poor, and dark-matter dominated systems, which makes them excellent laboratories to study stellar and galactic physics in the high-redshift Universe and at the faintest end of the galaxy luminosity function. In Chapter 4, I apply a one-zone analytic chemical evolution model within a hierarchical Bayesian framework to CaHK-based photometric metallicity measurements in the reionization-era UFD Eridanus II (Eri II). I present novel constraints on the underlying galaxy formation physics at $z \gtrsim 7$, finding that Eri II is well characterized by a short ($\sim$400 Myr) and inefficient burst of star formation with large supernova-driven gas outflows. The inferred scenario is consistent with the prevailing notion that low-mass galaxies struggle to both convert their gas content into stars and retain their gas reservoirs. Spectroscopic follow-up of stars in Eri II with JWST/NIRSpec will greatly improve the presented constraints. The framework introduced in this chapter can readily be applied to all UFDs throughout the Local Group, providing new insights into the underlying physics governing the evolution of the faintest galaxies in the early Universe.

As a whole, this dissertation seeks to address a range of current and anticipated challenges in the field of extragalactic stellar archaeology, from the robust measurement of low-resolution stellar chemical abundance patterns to the interpretation of low-mass galaxy chemical evolution. Though considerable work remains in preparation for the next decade of extragalactic stellar spectroscopic observations and chemical evolution studies, this dissertation represents a substantial contribution to the field on which future work can build.

Contents

List of Figures

List of Tables

Acknowledgments

Financial support for the research presented in this thesis comes in part from the NSF GRFP under grants DGE 1752814 and DGE 2146752. Additional support was provided from HST-GO-15901, HST-AR-16159, HST-GO-16226, and JWST DD-ERS-1334 from the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555.

Much of the computational analysis included in this dissertation was run on the Savio computational cluster resources provided by the Berkeley Research Computing program at the University of California, Berkeley.

I thank the anonymous referees of the three papers that compose the primary sections of this dissertation. Their constructive comments substantially improved Chapters 2–4.

I thank Bob Kurucz for developing and maintaining programs and databases without which the work of Chapters 2 and 3 would not be possible.

I thank all those who contributed to the design and ongoing utility of the various exposure time calculators (ETCs) used throughout Chapter 2. Special thanks goes to Brad Holden, Luke Schmidt, Nicolas Flagey, and Kyle Westfall for providing additional information about the Keck, GMACS, MSE, and FOBOS/WFOS ETCs, respectively, as well as Maosheng Xiang for providing example S/N curves of LAMOST spectra. I would also like to thank Hans-Walter Rix, Kim Venn, Alex Ji, Douglas Finkbeiner, Julianne Delcanton, Josh Speagle, and Gregory Green for insightful discussions that improved the work included in Chapter 2.

Some of the data presented in Chapter 3 were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. Further, this data was made accessible by the Keck Observatory Archive (KOA), which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExScI), under contract with the National Aeronautics and Space Administration. I wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. I am most fortunate to have the opportunity to conduct observations and use data from this mountain.

The spectroscopic data analyzed in Chapter 3 was collected by Keck programs U09H (PI: R. Kraft), C316Hr (PI: E. Kirby), and C147Hr (PI: J. Cohen). Corresponding photometry was obtained from Gaia DR3 using the Vizier catalogue access tool and the SIMBAD database, operated by the CDS, Strasbourg France.

I thank Jennifer Sobeck, Chris Sneden, Anish Amarsi, Thomas Nordlander, and Mikhail Kobalëv for their helpful insight on NLTE effects in stellar spectra as they pertained to the interpretation of the results in Chapter 3.

The observational basis for the data analyzed in Chapter 4 was acquired as part of HST-GO-15901 (PI: D. Weisz).

I thank Ellie Abrahams, Ani Chiti, Alex Ji, James Johnson, Andrey Kravtsov, and Alessandro Savino for insightful discussion on the work presented in Chapter 4.

This thesis would never have existed without:

Dr. Dan Weisz, my primary advisor, for too many things to list but in particular for your guidance and your belief in my worth as a scientist.

Drs. Yuan-Sen Ting, David Weinberg, and Marla Geha, for being amazing external mentors, advisors, and collaborators without which I would have spent far more of my time stuck on seemingly impossible research problems.

Drs. Josh Bloom and Jessica Lu, for agreeing to serve on my dissertation committee and for their constructive comments in my qualifying exam and on this dissertation.

Dr. Phil Choi, for being an outstanding undergraduate advisor and mentor.

Yumi Choi, Nick Choksi, Kareem El-Badry, Sal Fu, Maude Gull, Anna McLeod, Alessandro Savino, Yong Zheng, and Tom Zick, past and present members of the Weisz Group, for their support, productive science discussions, and companionship.

Aliza Beverage, Wren Suess, James Sullivan, and Tom Zick, members of the Campbell 3rd floor "CubeZone," for delightful company, stimulating discussion, and a wonderful shared workspace.

Kiran Eiden, Christian Helum Bye, Emiko Gardiner, and all other contributors to the Astro Organizing Committee for dedicating your time to making sure workers' rights are respected. More broadly, all UAW 2865 and UAW 5810 organizers past, present, and future for fighting to make the University of California and academia generally a more equitable environment.

Steven Giacalone and Jordan Fleming, my closest friends, for your unending love and support over the last six years (and hopefully long into the future).

Mom and Dad, for providing a strong foundation on which I have built my life and for fostering my creativity and curiosity from a young age.

Flamingo "Gogo" Elizabeth Honeycat, for being the best cat.

Grace, the love of my life, for literally everything.

Chapter 1

Background: Historical Context and Scientific Motivation

Broadly, this dissertation is concerned with the end-to-end study of stellar chemistry in low-mass Local Group galaxies, from measurement to interpretation. Of this dissertation's three main chapters, the first two are primarily focused on the measurement of robust chemical abundance patterns from low-resolution extragalactic resolved star spectroscopy, while the third focuses on inferring the galactic physics of the ancient low-mass galaxy Eridanus from the metallicity of its stellar population. Because each of these primary chapters are dense with methodology already, the introductory text that follows focuses on providing broad historical context and scientific motivation. It is, in part, adapted from the introductions of the dissertation's constituting papers, supplemented as necessary.

I begin in Section 1.1 with a comprehensive overview of the current state of stellar spectroscopy with an emphasis on chemical abundance measurements. It highlights important theoretical, instrumental, and technical advancements that have enabled the work presented in Chapters 2 and 3, as well as the current challenges and future prospects that motivates this work. In Section 1.1.1, I expound on the "synthetic gap," a substantial barrier to robust low-resolution chemical abundance measurements, which is the focus of Chapter 3. In Section 1.2, I conclude this background chapter by discussing low-mass Local Group galaxies as laboratories to study a range of astrophysical phenomena, focusing primarily on their chemical evolution. This section emphasizes the underlying scientific motivation of this dissertation as a whole and illustrate the primary through line connecting Chapters 2 and 3 with Chapter 4.

1.1 Stellar Spectroscopy

Absorption features imprinted in the spectrum of a star encode its physical structure and chemical composition. In turn, the chemical composition of individual stars trace the chem-

istry of the interstellar medium at their birth,[1] providing a detailed fossil record of the chemical evolution of a galxy over cosmic time. Various enrichment processes (e.g., core-collapse and thermonuclear supernovae, stellar winds, neutron star mergers, and gas inflows) each leave a unique chemical signature on their environment, which are captured in the abundance patterns of stars observed today (Tinsley 1980). Accordingly, the spectra of resolved stars provide a wealth of information on everything from the formation histories of galaxies to detailed nuclear and quantum physics (e.g., Frebel and Norris 2015; Matteucci 2021, and references therein).

However, translating stellar spectra to stellar composition is a nontrivial undertaking that relies on ∼200 years of advancement in atomic and stellar physics, astronomical instrumentation, and computational methods. The field of stellar spectroscopy and chemical abundance measurements has had a rich history since the first recorded solar spectrum by Fraunhofer (1817) and the subsequent identification of specific elemental absorption features nearly 50 years later (e.g., Kirchhoff 1860; Kirchhoff and Bunsen 1860; Kirchhoff 1863; Huggins et al. 1864). As chronicled in Hearnshaw (2010), it was another ∼70 years until the first quantitative abundance measurements were made. Such measurements were only possible after breakthroughs in theoretical physics (e.g., atomic/ionization theory and stellar atmospheres), development of new instrumentation (e.g., blazed gratings, coudé spectrographs, and Schmidt cameras), and substantial investment in laboratory experiments (e.g., transition wavelengths, oscillator strengths, and opacities). Together, these advances enabled the pioneering abundance work of Payne (1925), Russell (1929), Unsöld (1938), Strömgren (1940), Aller (1942), Unsöld (1942), Aller (1946), and Greenstein (1948) and Wright (1948) upon which modern stellar spectroscopy is founded.

Since the first half of the 20$^{\text{th}}$ century, high-resolution ($R > 10,000$) spectroscopy with broad optical wavelength coverage and high signal-to-noise ratio (S/N; $>30\,\text{pixel}^{-1}$) has been the gold standard for measuring precise stellar atmospheric parameters and detailed chemical abundance patterns (Nissen et al. 2018). These spectra provide clean, un-blended absorption features that can typically be fit with equivalent widths (EW).[2] At the same time, such high-resolution studies are often limited to small numbers of bright stars due to high dispersion, low throughput, and poor multiplexing.

In comparison, low- and medium-resolution spectrographs provide the opportunity to observe more and fainter stars, but are burdened with the cost of having (sometimes heavily) blended features that prohibit the use of conventional EW techniques.

As a means around this challenge, a number of studies have employed spectral indices for low-resolution chemical abundance measurements. One especially common index is centered around the Ca II triplet at ∼9000 Å (e.g., Cenarro, Cardiel, et al. 2001; Cenarro, Gorgas, Cardiel, Pedraz, et al. 2001; Cenarro, Gorgas, Cardiel, Vazdekis, et al. 2002, and references therein). In this method, the strength of a blended spectral feature (e.g., the Ca II triplet)

[1] Modulo mixing and gravitational settling.

[2] The EW is a measure of the flux removed by an absorption feature in relation to the continuum of the spectrum. Formally, it is defined to be EW $\equiv \int(1 - F_s/F_c)d\lambda$, where F_s and F_c are the flux of the spectrum and the underlying continuum respectively. See Minnaert (1934) for an early discussion of equivalent widths.

is calibrated to (e.g., Olszewski et al. 1991; Rutledge et al. 1997; Carrera et al. 2013) or to theoretical (i.e., *ab initio*) spectra generated from stellar atmosphere and spectrum synthesis models (Baschek 1959; Fischel 1964; Bell 1970; Bell and Branch 1976).[3] However, spectral indices provide only bulk metal abundances (requiring assumptions of chemical abundance patterns) and are restricted to the parameter space of their calibrating stars or models (Battaglia, Helmi, et al. 2008; Koch, Grebel, Gilmore, et al. 2008; Starkenburg, Hill, Tolstoy, González Hernández, et al. 2010).

As computational resources and stellar models continued to improve, it became possible to directly compare *ab initio* spectra to observed spectra on a pixel-by-pixel basis (pioneering examples include Cayrel 1969; Sneden 1973, 1974; Suntzeff 1981; Carbon et al. 1982; Leep, Wallerstein, et al. 1986; Leep, Oke, et al. 1987; Wallerstein et al. 1987). This technique leverages the full statistical power of the many absorption lines in a spectrum, yielding precise abundance measurements without the use of EWs or spectral indices. These methods have proven powerful for the recovery of detailed abundance patterns from low- and medium-resolution spectra, which contain predominantly weak and blended absorption features.

In the last two decades, massively multiplexed stellar spectroscopic surveys including the RAdial Velocity Experiment (RAVE; Steinmetz et al. 2006), the Sloan Extension for Galactic Understanding and Exploration (SEGUE; Yanny et al. 2009), the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST; Luo et al. 2015) survey, the Galactic Archaeology with HERMES (GALAH; De Silva et al. 2015) survey, the Apache Point Observatory Galactic Evolution Experiment (APOGEE; Majewski et al. 2017), and the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration, Aghamousa, et al. 2016; DESI Collaboration, Aguilar, et al. 2016) survey have collected millions of spectra of Milky Way (MW) stars. Coupled with steady progress in theoretical and laboratory astrophysics, these surveys have revolutionized our ability to collect and interpret the spectra of stars (see reviews by Allende Prieto 2016; Nissen et al. 2018; Jofré, Heiter, and Soubiran 2019). Importantly, they have motivated the development of novel fitting techniques designed to efficiently fit the full spectrum of many stars. Some techniques are data driven (e.g., The Cannon; Ness, Hogg, et al. 2015), some are trained on *ab initio* spectra (e.g., The Payne; Ting, Conroy, Rix, and Cargile 2019), and others adopt hybrid methods (e.g., The DD-Payne; Xiang, Ting, et al. 2019). All employ sophisticated statistical techniques (e.g., neural networks, Bayesian inference, and/or machine learning), enabling the precise recovery of dozens of elemental abundances from both low- and high-resolution spectra in modest compute times.

Full-spectrum fitting techniques will become increasingly crucial over the next decade as massively-multiplexed spectroscopic surveys, including WEAVE (Dalton et al. 2016), SDSS-V (Kollmeier et al. 2017), PFS (Tamura et al. 2018), MOONS (Taylor et al. 2018), 4MOST (de Jong et al. 2019), and FOBOS (Bundy et al. 2019), deliver an order-of-magnitude increase in the number of high-quality, but low-resolution, stellar spectra from which detailed stellar chemistry can be measured (see Figure 3.1 in Chapter 3).

However, extragalactic stellar spectroscopy has yet to experience the same tremendous

[3]Similar to how EWs are calibrated for high-resolution studies.

gains in quantity and quality of abundance measurements as seen for spectroscopy of stars in the MW. This is primarily due to the large distances to external galaxies, which result in their stars being fainter and thus more challenging to observe. Generally, only the few brightest stars ($m_V \lesssim 19.5$) in extragalactic systems can be observed at high resolution, even when using 10 meter class telescopes (e.g., Shetrone, Bolte, et al. 1998; Shetrone, Côté, et al. 2001; Shetrone, Venn, et al. 2003; Tolstoy, Venn, et al. 2003; Fulbright et al. 2004; Venn, Irwin, et al. 2004; Walker, Mateo, Olszewski, Gnedin, et al. 2007; Koch, Grebel, Gilmore, et al. 2008; Koch, McWilliam, et al. 2008; Aoki et al. 2009; Cohen et al. 2009; Walker, Belokurov, et al. 2009; Walker, Mateo, and Olszewski 2009; Frebel, Simon, Geha, et al. 2010; Starkenburg, Hill, Tolstoy, François, et al. 2013; Frebel, Simon, and Kirby 2014; Koch and Rich 2014; Walker, Mateo, Olszewski, Bailey, et al. 2015; Walker, Olszewski, et al. 2015; Frebel, Norris, et al. 2016; Ji, Frebel, Ezzeddine, et al. 2016; Ji, Frebel, Simon, and Chiti 2016; Ji, Frebel, Simon, and Geha 2016; Spencer et al. 2017; Venn, Starkenburg, et al. 2017; Spite et al. 2018; Hill et al. 2019; Theler et al. 2019).[4]

Instead, highly multiplexed low- and moderate-resolution ($R < 10000$) spectrographs on large-aperture telescopes have become the workhorse instruments of extragalactic stellar spectroscopy (e.g., the DEep Imaging Multi-Object Spectrograph (DEIMOS); Faber et al. 2003). Over the past 20 years, tens of thousands of low- and medium-resolution spectra have been acquired for extragalactic stars. Because detailed abundance measurements were typically viewed as the purview of high-resolution spectroscopy, most of the spectra were taken for the purpose of measuring radial velocities and bulk metallicities with spectral indices (e.g., Suntzeff et al. 1993; Pont et al. 2004; Tolstoy, Irwin, et al. 2004; Battaglia, Tolstoy, Helmi, Irwin, Letarte, et al. 2006; Muñoz et al. 2006; Koch, Grebel, Kleyna, et al. 2007; Koch, Wilkinson, Kleyna, Gilmore, et al. 2007; Simon and Geha 2007; Battaglia, Helmi, et al. 2008; Norris et al. 2008; Koch, Wilkinson, Kleyna, Irwin, et al. 2009; Leaman et al. 2009; Shetrone, Siegel, et al. 2009; Kalirai et al. 2010; Battaglia, Tolstoy, Helmi, Irwin, Parisi, et al. 2011; Hendricks et al. 2014; Ho et al. 2015; Simon, Drlica-Wagner, et al. 2015; Slater et al. 2015; Martin, Geha, et al. 2016; Martin, Ibata, Collins, et al. 2016; Swan et al. 2016; Li et al. 2017; Simon, Li, et al. 2017; Longeard et al. 2020).

The groundbreaking work of Kirby, Guhathakurta, Bolte, et al. (2009) was the first to demonstrate that precise abundances could be recovered from moderate-resolution spectra in external galaxies. Since then, the method has been further refined and applied to thousands of stars in Local Group (LG) galaxies, measuring up to ∼10 abundances in MW satellites and ∼5 abundances at the distance of M31 (e.g., Kirby, Guhathakurta, Simon, et al. 2010; Vargas, Geha, et al. 2013; Vargas, Gilbert, et al. 2014; Kirby, Cohen, Simon, and Guhathakurta 2015; Kirby, Guo, et al. 2015; Kirby, Simon, et al. 2015; Kirby, Cohen, Simon, Guhathakurta, et al. 2017; Kirby, Rizzi, et al. 2017; Duggan et al. 2018; Kirby, Xie, et al. 2018; Escala, Kirby, et al. 2019; Gilbert et al. 2019; Escala, Gilbert, et al. 2020; Kirby, Gilbert, et al. 2020).

As such, the field of extragalactic stellar archaeology is poised for enormous growth.

[4]To date, $\sim 10^4$ stars outside the MW have measured [Fe/H] from $R > 10000$ spectroscopy, though most have only been observed over a small (∼100 Å) range in wavelength.

Upcoming wide-area imaging surveys conducted by the *Nancy Grace Roman Space Telescope* (Spergel et al. 2015), the Vera C. Rubin Observatory (Ivezić et al. 2019), and *Euclid* (Laureijs et al. 2011) have the potential to discover and resolve the stellar populations of hundreds of low-mass galaxies, stellar streams, and stellar substructure throughout the Local Group and out to 5 Mpc (Mutlu-Pakdil et al. 2021; Pearson et al. 2022). These discoveries will require deep spectroscopic follow-up to confirm their existence, characterize their stellar populations, and investigate their origin and evolution. Fortunately, current and future spectroscopic facilities on large-aperture telescopes promise to increase the number of stars outside the MW with observed spectra by at least an order of magnitude. Already, existing spectrographs on 6+ meter telescopes have been used to measure abundances of over $\sim 10^4$ stars in LG dwarf galaxies and the halo of M31 (see Suda et al. 2017, and references therein) and are capable of measuring thousands more.

In the next decade, dedicated spectroscopic surveys on large telescopes including the Prime Focus Spectrograph (PFS; Takada et al. 2014), the Maunakea Spectroscopic Explorer (MSE; MSE Science Team et al. 2019), and the Fiber-Optic Broadband Optical Spectrograph (FOBOS; Bundy et al. 2019) will homogeneously collect hundreds of thousands of resolved star spectra in external galaxies. The next decade will also bring the *James Webb Space Telescope* (JWST) and Extremely Large Telescopes (ELTs; e.g., the Giant Magellan Telescope (GMT), the European ELT (E-ELT), and the Thirty Meter Telescope (TMT)), which will make possible the spectroscopy of stars in the most distant, faint, and crowded environments in the LG and beyond that are inaccessible to current ground-based facilities.

Chapters 2 and 3 of this dissertation lay critical groundwork necessary to fully capitalize on the wealth of chemical information that will be embedded in these extragalactic stellar spectra.

1.1.1 The "Synthetic Gap" in Stellar Spectroscopy

While stellar spectroscopy has seen enormous gains in the last decade, the field remains largely dominated by systematic uncertainties, especially in the low-resolution regime required for extragalactic stellar spectroscopy. A particular challenge of low-resolution spectroscopy is the need to forward-model the full stellar spectrum, which relies heavily on the fidelity of the chosen stellar models and their physical assumptions. For computational efficiency, many stellar models adopt simplifying assumptions regarding the nature of the stellar atmospheres. This includes utilizing 1D plane-parallel atmospheric geometries (instead of 3D spherical) and assuming radiative transfer occurs in local thermodynamic equilibrium (LTE; instead of non-local thermodynamic equilibrium). These assumptions hold valid for some stars and elemental abundances but become increasingly tenuous when modeling the low-metallicity giant stars typically targeted by extragalactic observations. This can result in poor matches between models and observations and introduce systematic trends in the recovery of elemental abundances, especially for low-resolution spectra with heavily blended absorption features (Nissen et al. 2018). The inability of model-generated spectra to perfectly reproduce observations is commonly referred to as the "synthetic gap."

The synthetic gap is exacerbated by two additional sources of systematics. First, synthetic stellar models are reliant on atomic and molecular line data from which transition rates can be calculated. Unfortunately, these line lists are frequently plagued with missing or extraneous transitions and imperfectly calibrated oscillator strengths, which can further inhibit accurate reproduction of absorption line depths (e.g., Martins et al. 2007; Shetrone, Bizyaev, et al. 2015). Second, the observed spectrum may also include features of observational or instrumental origin not captured in the theoretical model. For example, bad pixels, imperfect sky subtraction, and variations to the line spread function can all contribute to mismatches with the model spectrum.

When analyzing high-resolution stellar spectra, the synthetic gap can be minimized by ignoring problematic spectral features. But at low resolution, many absorption features blend together and the problematic spectral features become inseparable from their neighboring lines, contaminating a much larger fraction of the spectrum. Understanding and bridging the synthetic gap is thus crucial to unifying the high-resolution spectroscopic abundances of Milky Way surveys (e.g., APOGEE, GALAH) and the low-resolution spectroscopic abundances required for large-scale extragalactic surveys (e.g., PFS, MSE, FOBOS).

The need to bridge the synthetic gap has motivated considerable investment in improving theoretical models through the inclusion of 3D hydrodynamic effects and non-LTE (NLTE) radiative transfer calculations (e.g., Amarsi, Asplund, et al. 2015; Amarsi, Lind, et al. 2016; Kovalev et al. 2019). While the results are promising and yield marked improvement over 1D-LTE models, their large computational costs have thus far limited their widespread application. Only sparse model grids featuring a subset of elemental abundances are currently available. Similarly, despite serious sustained efforts to improve atomic and molecular line data (e.g., Lawler et al. 2013; Ryabchikova et al. 2015; Den Hartog et al. 2019), there remain many missing and un-calibrated atomic and molecular transitions.

Many novel spectroscopic fitting techniques have been developed in response to the synthetic gap. Data-driven methods, like `the Cannon` (Ness, Hogg, et al. 2015), attempt to sidestep the shortcomings of theoretical models by training empirical models on existing high-quality data with robust abundance measurements. With large stellar spectroscopic datasets from high(er)-resolution surveys, like APOGEE and GALAH, now readily accessible, data-driven methods have grown quite popular for fitting low-resolution spectra (e.g., Guiglion et al. 2020; Wang, Luo, Chen, et al. 2020; Wheeler et al. 2020; Zhang et al. 2020). Data-driven methods, however, are not without their shortcomings. Specifically, they cannot be extrapolated beyond the training dataset and are prone to learning astrophysical correlations beyond what chemical information is contained within the spectrum (e.g., that stars enhanced in Mg are also enhanced in O; Buder et al. 2021). Moreover, data-driven models do not truly avoid the need for theoretical models—the model-dependency is simply inherited from the training dataset and propagated into the new measurements.

One promising alternative to purely theoretical or purely data-driven methods is the use of a machine learning framework called domain adaptation, which forces two domains—here the synthetic spectra domain and observed spectra domain—to share an abstract representation (see Liu et al. 2017; Zhu et al. 2017, for details). In the last few years, this technique has

been incorporated into a spectral fitting framework, `CYCLE-STARNET`, and demonstrated to partially mitigate the synthetic gap in analyses of both APOGEE and LAMOST spectra (O'Briain et al. 2021; Wang, Luo, Zhang, et al. 2023). Domain adaptation methods have the benefit of combining the physical interpretability of theoretical models with the data consistency of empirical models. While certainly encouraging, these methods are still very much in their infancy. For example, in their analysis of LAMOST spectra, Wang, Luo, Zhang, et al. (2023), found non-negligible corrections are still required to bring their low-resolution abundances in line with previous high-resolution studies. Additionally, while `CYCLE-STARNET` does not need the pre-analyzed training dataset that data-driven techniques require, it does require a large observational dataset in order to learn the transformation from theoretical to observed spectra. As such, it has yet to be applied to the smaller datasets that currently exist for extragalactic systems.

As laid out above, the synthetic gap remains a major barrier to robust low-resolution abundance measurements despite enormous effort. Given the crucial role that low-resolution observations play in extragalactic chemical abundance measurements, identifying, understanding, and ultimately correcting for the synthetic gap is of the utmost importance if we are to draw robust conclusion connecting the chemical evolution of the Milky Way to that of our neighboring galaxies. It is within this conversation and ongoing work in which the material of Chapter 3, an investigation of the synthetic gap as a function of resolving power, is situated.

1.2 Chemical Evolution in Local Group Dwarf Galaxies

The chemical composition of individual stars trace a galaxy's evolution over cosmic time. Supernovae, stellar winds, neutron star mergers, and gas in/outflows each leave a unique chemical signature in the abundance patterns of stars observed today. Accordingly, studying the detailed chemical abundances of large numbers of resolved stars yields valuable insight into everything from galactic winds to the detailed nuclear and quantum physics that underpin stellar evolution and the production of heavy elements (see Pagel 1997; Käppeler et al. 2011; Nomoto et al. 2013; Arcones et al. 2023, and references therein). Within the Milky Way, large samples of stars with detailed chemical abundance measurements have revolutionized our understanding of the formation and evolution of the Galaxy's disk, bulge, and halo (e.g., Tinsley 1980; McWilliam 1997; Tumlinson 2010; Ness, Freeman, et al. 2013; Rix et al. 2013; Hayden et al. 2015; Bensby et al. 2017).

The stellar chemistry of lower-mass dwarf galaxies provides an especially powerful lens through which to study stellar and galactic physics, as their shallow gravitational potential wells make them particularly sensitive to the physics of internal and external feedback processes, including stellar feedback, gas accretion, and cosmic reionization (e.g., Boylan-Kolchin et al. 2016). Additionally, their (predominantly) ancient and metal-poor stellar

populations are among the best proxies for the low-metallicity star-forming conditions of the early Universe (e.g., Frebel and Norris 2015). At high redshift, direct observation of stellar populations in faint, low-mass galaxies is a Herculean task—even in the era of JWST. But within the Local Group, there exists an ever-growing population of low-mass galaxies in which individual stars can be resolved and studied as in the Milky Way.

For these reasons, Local Group dwarf galaxies have been the subject of large observing campaigns on large ground-based 6+ meter telescopes, which have acquired spectroscopy and chemical abundance measurements for over 10^4 resolved stars in Local Group dwarf galaxies (see Suda et al. 2017, and references therein). Meanwhile, space-based photometric observations with HST have produced deep color-magnitude diagrams (CMDs) and high-fidelity measurements of the star formation histories for these systems (e.g., Weisz et al. 2014; Savino et al. 2023). Despite the technical challenges associated with extragalactic stellar archaeology (e.g., faintness and crowding), these large observational investments have contributed substantially to our understanding of dwarf galaxy evolution (see Tolstoy, Hill, et al. 2009, and references therein).

However, our understanding of the chemical evolution for the Local Group's lowest-mass galaxies, ultra-faint dwarf galaxies (UFDs), remains much less certain. At the faintest end of the galaxy luminosity function ($L < 10^5$ $L_\odot$), UFDs are some of the oldest, most dark matter-dominated, and lowest metallicity galaxies in the Universe and thus present a promising opportunity to study galaxy evolution at very early times (see Simon 2019, and references therein). But owing to the paucity of bright stars in these galaxies, only a few hundred stars across ∼50 known UFDs have bulk metallicity measurements and only ∼60 have measured [α/Fe] abundances. Fewer still, roughly a dozen across a small set of the nearest UFDs, have more detailed chemical abundance measurements (e.g., Frebel and Norris 2015; Ji, Simon, et al. 2019). These measurements, though sparse, provide a starting point from which the chemical evolution of UFDs can be investigated.

The stellar [Fe/H] metallicity distribution functions (MDFs) of UFDs feature both low mean metallicities ([Fe/H] < -2.0) and large metallicity dispersion (∼0.5 dex), which distinguish their stellar populations from those of mono-metallic globular clusters (see Simon 2019, and references therein). These characteristics point to extended star formation histories and large galactic outflows, though for most UFDs, the MDFs are too sparse for statistical comparisons to galactic chemical evolution models. Meanwhile the distributions of [α/Fe] vs. [Fe/H], which encodes several key galactic baryonic processes, including relative chemical enrichment from prompt core-collapse SNe and delayed Type Ia SNe, the timescale of star formation, and the strength of galactic outflows (e.g., Weinberg et al. 2017), have yielded discrepant conclusions from galaxy to galaxy. Some studies have claimed a universal knee in the [α/Fe]-[Fe/H] plane suggesting a minimum duration for UFD star formation (e.g., Vargas, Geha, et al. 2013), while others have found no evidence of a knee suggesting that star formation ended rapidly before the onset of Type Ia SNe (e.g., Frebel, Simon, and Kirby 2014). Higher-order stellar chemistry similarly exhibits large star-to-star and galaxy-to-galaxy variation, which suggests that UFDs, despite being the most primitive galaxies in the Universe, underwent complex, likely bursty, star formation episodes and rare nucleosynthetic events

like neutron star mergers (Ji, Frebel, Ezzeddine, et al. 2016).

Clearly, additional measurement of stellar chemistry are required to solidify our understanding of stellar and galactic physics in UFDs. Fortunately, the number of UFD stars with measured chemistry is set to increase by an order of magnitude or more over the next decade. Deep narrowband photometry of the calcium H&K doublet acquired with HST has enabled measurements of stellar [Fe/H] abundances for stars 2–3 magnitudes fainter than currently possible with existing ground-based facilities (Fu, Weisz, Starkenburg, Martin, Ji, et al. 2022; Fu, Weisz, Starkenburg, Martin, Savino, et al. 2023). In the next year, JWST/NIRSpec will also demonstrate its ability to measure [Fe/H] and [α/Fe] in distant low-mass Local Group galaxies (JWST-GO-3788; PI D. Weisz). Moreover, powerful new spectrographs on 10-m class telescopes (e.g., FOBOS, MSE) and 30-m class telescopes (e.g., TMT, GMT, E-ELT) will soon enable efficient observation of stars in faint galaxies throughout the Local Group, including those soon-to-be-discovered by upcoming wide-are imaging surveys. With these facilities operational, a complete picture of galactic chemical evolution that extends to the very smallest and oldest galaxies is possible.

As discussed previously, Chapters 2 and 3 are concerned with facilitating the observations necessary for future chemical evolution studies. In Chapter 4, I present my own contribution to the study of chemical evolution in low-mass galaxies, an analysis of UFD Eridanus II's newly measured Calcium H&K-based MDF.

Chapter 2

Forecasting Chemical Abundance Precision for Extragalactic Stellar Archaeology

As outlined in the introductory chapter, the body of this dissertation contains three primary pieces of work, which can be coarsely summarized as predictions, validation, and application. In this first main chapter, I present predictions of the chemical abundances that current and future spectrographs on large ground- and space-based telescopes are capable of measuring. As with the remainder of the dissertation, the scope of these predictions is focused on the observation of metal-poor red giant branch stars in galaxies beyond our Milky Way. This chapter provides an exciting glimpse into the enormous gains that await to be made over the next decade in the field of extragalactic stellar archaeology.

This chapter has been adapted from the paper Sandford, Weisz, et al. (2020), which was published in the Astrophysical Journal Supplement Series. It was coauthored by the following individuals, and is included in this thesis with their express permission: Daniel Weisz, Yuan-Sen Ting.

2.1 Abstract

Increasingly powerful and multiplexed spectroscopic facilities promise detailed chemical abundance patterns for millions of resolved stars in galaxies beyond the Milky Way (MW). Here, we employ the Cramér-Rao Lower Bound (CRLB) to forecast the precision to which stellar abundances for metal-poor, low-mass stars outside the MW can be measured for 41 current (e.g., Keck, MMT, the Very Large Telescope, and the Dark Energy Spectroscopic Instrument) and planned (e.g., the Maunakea Spectroscopic Explorer, the James Webb Space Telescope (JWST), and Extremely Large Telescopes (ELTs)) spectrograph configurations. We show that moderate-resolution ($R \lesssim 5000$) spectroscopy at blue-optical

wavelengths ($\lambda \lesssim 4500\,\text{Å}$) (i) enables the recovery of 2–4 times as many elements as red-optical spectroscopy ($5000 \lesssim \lambda \lesssim 10000\,\text{Å}$) at similar or higher resolutions ($R \sim 10000$) and (ii) can constrain the abundances of several neutron capture elements to $\lesssim$0.3 dex. We further show that high-resolution ($R \gtrsim 20000$), low signal-to-noise ($\sim$10 pixel^{-1}) spectra contain rich abundance information when modeled with full spectral fitting techniques. We demonstrate that JWST/NIRSpec and ELTs can recover (i) $\sim$10 and 30 elements, respectively, for metal-poor red giants throughout the Local Group and (ii) [Fe/H] and [α/Fe] for resolved stars in galaxies out to several Mpc with modest integration times. We show that select literature abundances are within a factor of $\sim$2 or better of our CRLBs. We suggest that, like exposure time calculators, CRLBs should be used when planning stellar spectroscopic observations. We include an open-source python package, `Chem-I-Calc`, that allows users to compute CRLBs for spectrographs of their choosing.

2.2 Introduction

The chemical composition of a galaxy's individual stars provide a fossil record of its chemical evolution over cosmic time, encoding important galactic enrichment mechanisms, timescales, and nucleosynthetic pathways. Over the last decade, large spectroscopic surveys (e.g., APOGEE, GALAH) have enabled precise chemical abundance measurements from millions of high-quality, high-resolution spectra of Milky Way (MW) stars, transforming our understanding of the Galaxy and its formation. While similarly large-scale spectroscopic surveys of stars outside the MW have thus far been impractical, the next decade will bring increasingly powerful and highly-multiplexed spectroscopic facilities (e.g., PFS, MSE, FOBOS, ELTs), supplementing existing spectrographs on large 10-m class telescopes (e.g., Keck, Magellan, VLT) and making accessible the spectra and chemical compositions of millions of stars throughout the Local Group.

To fully realize the scientific potential of upcoming massive data sets and to plan for observational campaigns further in the future, it is imperative that we can quantify what we expect to be able to measure from these spectra, and to what precision. While there exist preferred spectral wavelength regions, absorption features, and minimum S/N for abundance measurements, best practices are frequently informally passed down in the community. Comprehensive and quantitative analyses of the chemical information content of spectra given their wavelength coverage, resolution, and S/N are important planning tools, but are sparse in the literature (e.g., Caffau, Koch, et al. 2013; Bedell et al. 2014; Hansen et al. 2015; Ruchti et al. 2016; Ting, Conroy, Rix, and Cargile 2017; Feeney et al. 2019).

In this paper, we employ *ab initio* stellar spectra and the Cramér–Rao lower bound (CRLB) to quantify the chemical information content of stellar spectra in terms of the precision (not accuracy[1]) to which elemental abundances can be measured. We apply this method to realistic observing conditions of metal-poor, low-mass stars outside the MW

[1]See Blanco-Cuaresma (2019) and Jofré, Heiter, and Soubiran (2019) for investigations of the systematics present in spectroscopically derived elemental abundances.

for > 40 instrument configurations on current (e.g., Keck, the Large Binocular Telescope (LBT), Magellan, MMT, and the Very Large Telescope (VLT)) and future (e.g., JWST, GMT, TMT, E-ELT, and MSE) spectroscopic facilities. For this exercise, we assume the use of full-spectrum-fitting techniques and adopt many of the assumptions commonly used at present in this field (e.g., 1D local thermodynamic equilibrium (LTE) models). We note, however, that the techniques we present can readily be adapted for other choices (e.g., when large grids of non-LTE and/or 3D atmospheres become available).

This paper is organized as follows. In Section 2.3 we provide a technical description of the information content of spectra and how it can be quantified using CRLBs. In Section 2.4 we summarize the scope of stars, instruments, and observing scenarios evaluated in this work, our method of stellar spectra generation, and the assumptions that went into our CRLB calculations. We report the forecasted stellar abundance precision for current and planned spectrographs in Sections 2.5 and 2.6, respectively. We discuss the highlights and caveats of our forecasts in Section 2.7. In Section 2.8 we present `Chem-I-Calc`, an open-source Python package for calculating CRLBs of spectroscopic chemical abundance measurements. We summarize our findings in Section 2.9 and present a number of technical details in the appendices.

2.3 Information Content of Spectra

In this section we introduce the notion of the information content of a stellar spectrum and its relation to the maximal precision to which stellar labels[2] can be measured. We begin in Section 2.3.1 with a qualitative description of the factors that play a role in the degree of information contained in a stellar spectrum. This is followed by a quantitative description of the information content as represented by the CRLB in Section 2.3.2.

2.3.1 A Qualitative Description of Spectral Information

The information content of a stellar spectrum determines the precision to which we can measure the stellar labels—or more technically, how broad the stellar label posteriors are. The amount of information and how constraining that information is depends on the following intrinsic and observed properties of the spectrum:

(i) *Wavelength Coverage*: How many (and which) spectral features are included in the spectrum.

(ii) *Wavelength Sampling*: How many wavelength pixels are measured per resolution element.

[2]In this work we use "stellar labels" to broadly encompass both atmospheric parameters (e.g., effective temperature, surface gravity and microturbulent velocity) and elemental abundances. We do not, however, include radial velocities in our analysis.

(iii) *Spectral Resolution*: How distinct the spectral features of one label are from those of another label.

(iv) *Flux Covariance*: How uncertain/covariant is the flux in each spectral pixel.

(v) *Gradient Spectra*: How strongly spectral features respond to changes in the stellar labels.

Aspects (i)–(iv) are determined by the instrument configuration and observing conditions. Generally speaking, they set the size and quality of the spectral data set in question, modulating the availability and accessibility of the information contained in the spectrum. Larger wavelength coverage and higher wavelength sampling both increase the amount of information-carrying pixels contained in a spectrum. Increased spectral resolution, or resolving power ($R = \lambda/\delta\lambda$), reduces the blending of spectral features and the covariance between stellar labels. Lower flux covariance (i.e., higher S/N) increases the constraining power of informative spectral features. These various characteristics can depend on one another as well (e.g., spectral resolution and wavelength sampling affect the S/N and pixel-to-pixel flux covariance), and there are often trade-offs between them for a fixed instrument configuration or observational strategy.

The gradient spectra, aspect (v), is the most important factor in determining the stellar spectral information content. Generally speaking, it is the stellar labels that result in the largest spectral gradients that have the highest information content and therefore can be recovered to the highest precision. In a χ^2 sense, the more strongly a spectral feature responds to a change in stellar labels, the less the labels need to be offset from the true value to result in a large χ^2 value. More technically phrased: the expectation of the negative second derivative of the spectrum with respect to the stellar labels gives the Fisher information matrix (FIM), which provides a lower bound on the covariance matrix of the stellar labels as discussed in Section 2.3.2.

Figure 2.1 helps build intuition for the importance of spectral gradients. Here, we consider a moderate-resolution ($R = 6500$) *ab initio* normalized spectrum of a metal-poor ($\log Z = -1.5$) red giant branch (RGB) star[3] and the partial derivative of that spectrum with respect to Fe, Mg, and Y.[4] This spectrum and its derivatives were generated using the `ATLAS12` and `synthe` models (Kurucz 1970; Kurucz and Avrett 1981; Kurucz 1993, 2005, 2013, 2017), which we describe in more detail in Section 2.4.2. The locations and strengths of certain features in the spectral gradient may depend on the adopted stellar atmosphere and radiative transfer models, an issue we discuss in Section 2.7.4.

[3]It is important to remember that the gradient spectrum of a star depends on the stellar labels. Cool stars, giant stars, and metal-rich stars all have stronger gradients than hot stars, dwarf stars, and metal-poor stars, meaning that it is easier to precisely recover their stellar labels.

[4]Unless otherwise stated, elemental abundances are assumed to be in the form of standard solar-scaled abundance ratios with respect to H i.e., $[\mathrm{X/H}] = \log_{10}(\mathrm{X/H}) - \log_{10}(\mathrm{X/H})_\odot$, where $(\mathrm{X/H})_\odot$ is the Solar abundance ratio.

As depicted in panel (b) of Figure 2.1, Fe contributes strongly to a large number of absorption features between 6500 and 9000 Å, including over 200 lines with changes of $>1\%\,\mathrm{dex}^{-1}$ and nearly 50 lines with changes of $>5\%\,\mathrm{dex}^{-1}$. The large number of information-rich lines is the reason why Fe is one of the most readily recovered elements for cool, low-mass stars.

Compared to Fe, Mg contributes to only 20 features at the >1%/dex level and only one that is $>5\%\ ^{-1}$ (at λ8809). As a result, it is not as well constrained as Fe. Finally, Y exhibits only three features with gradients larger than $1\%\ ^{-1}$, illustrating the challenge of recovering its abundance, even with favorable telescope (high spectral resolution) and observational (high S/N) configurations.

Visually exploring the gradients is a particularly informative exercise. For example, there are clear peaks (i.e., positive deviations in the gradient) in the gradient spectra of Fe and Mg at ∼8500 Å. These peaks are not due to Fe or Mg transitions, but rather to the Ca II triplet, which is sensitive to the number density of free electrons that Fe and Mg contribute. Y, unlike Fe and Mg, is not a key electron donor and thus does not yield a strong gradient at the location of the Ca II triplet. In this manner, elements that change the atmospheric structure of a star or otherwise indirectly affect the line formation of other elements may be measured—even in the absence of strong absorption features of the element in question (e.g., O can be recovered from spectra that contain few, or no, O lines due to its important role in the CNO molecular network; see Ting, Conroy, Rix, and Asplund 2018). Such measurements, however, require a high degree of trust in the stellar atmosphere and radiative transfer models being used.

2.3.2 Quantifying Information Content with CRLBs

A main goal of this paper is to quantify the information content encapsulated in the gradient spectrum, modulated by commonly used instrumental setups and realistic observational considerations. To do this, we employ the CRLB (Fréchet 1943; Darmois 1945; Rao 1945; Cramer 1946), a formal metric for quantifying information content, which we now describe mathematically.

Suppose that we wish to quantify the information content of a stellar spectra observed using a spectrograph with a wavelength coverage of $\lambda_0 \leq \lambda \leq \lambda_N$, a resolving power R, and a wavelength sampling of $\Delta\lambda = \lambda/nR$, where n is the number of pixels per resolution element. Let $f_{\mathrm{obs}}(\lambda)$ be the continuum-normalized stellar flux and Σ be the covariance matrix of the normalized flux.

To make any assessment about the information contained within this spectrum requires a model that relates the physical characteristics of a star (e.g., T_{eff}, $\log g$, [Fe/H], [X/Fe]) to the observed spectrum. Suppose we have such a model, $f(\lambda, \theta)$, that predicts the normalized flux of a star at each wavelength, λ, given a set of stellar labels, θ. The nature of this model, whether it be data driven (e.g., Ness, Hogg, et al. 2015), *ab initio* (e.g., Ting, Conroy, Rix, and Cargile 2019), or a combination of the two (e.g., Xiang, Ting, et al. 2019), is unimportant provided that it is generative (i.e., it predicts a normalized flux that mimics

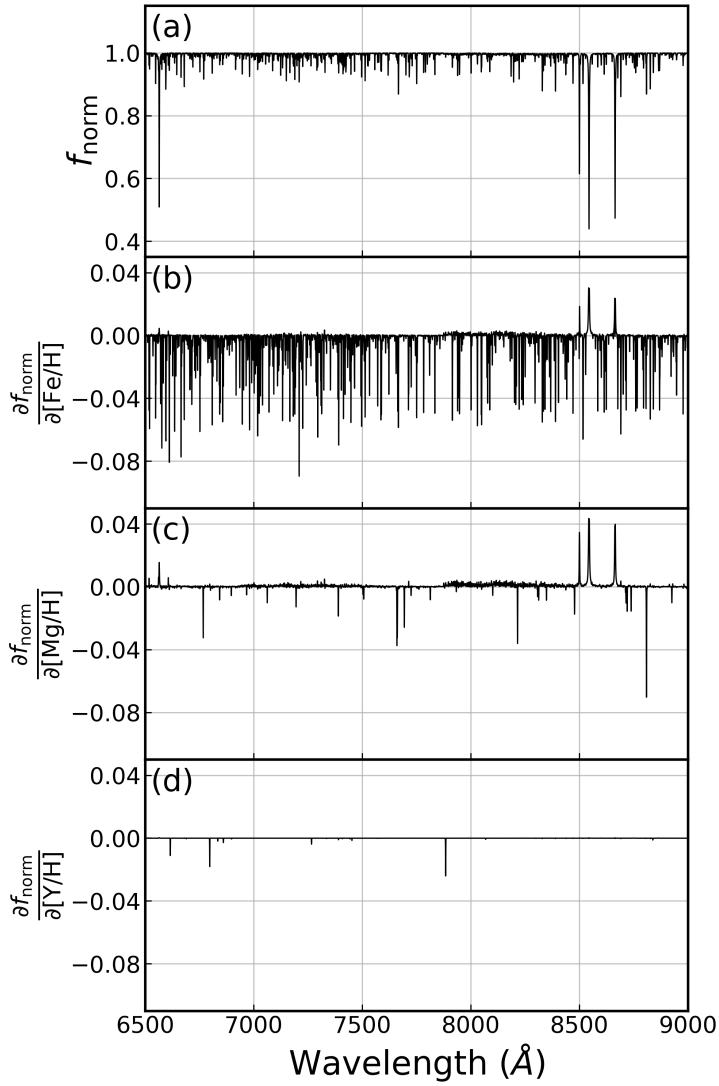


Figure 2.1: (a) Normalized flux of a synthetic $\log Z = -1.5$ RGB star at $R = 6500$ generated using `atlas12` and `synthe` (see Section 2.4.2 for model details). (b–d) Gradients of the normalized flux with respect to Fe, Mg, and Y respectively. Many features in the stellar spectrum respond strongly to changes in Fe, meaning that there is considerable information about the iron abundance contained in this spectrum. Changes in Y, on the other hand, cause very weak changes in only a few lines; as a result, the Y abundance would be difficult to recover precisely from this spectrum. Strong positive gradients for Fe and Mg can be seen at the location of the Ca II triplet, which is sensitive to the number of free electrons provided by Fe, Mg, and other electron donors.

the observed spectrum from a set of stellar labels) and differentiable in θ (i.e., the spectrum varies smoothly as the labels change).

We can then quantify the precision of our measurements by evaluating the log-likelihood of the data given our model,

$$\ln L(D|\boldsymbol{\theta}) = -\frac{1}{2}\sum_{i=0}^{N}\Big[\left(f_{\rm obs}(\lambda_i) - f(\lambda_i,\boldsymbol{\theta})\right)^T \Sigma^{-1}\left(f_{\rm obs}(\lambda_i) - f(\lambda_i,\boldsymbol{\theta})\right) + \ln\left(2\pi|\Sigma|\right)\Big]. \quad (2.1)$$

for all θ (i.e., over all stellar labels).

The precision to which these labels can be recovered is given by the width of this likelihood function. In practice, however, evaluating the likelihood over a sufficiently large region of parameter space is computationally expensive (and sometimes infeasible) given the high-dimensional nature of spectral fitting.[5] If one assumes priors on the stellar labels (uniform or otherwise), a Markov Chain Monte Carlo (MCMC) method, which enables more efficient sampling of the full posterior than evaluating the likelihood at a grid of labels, can be employed. However, it ultimately still succumbs to the curse of dimensionality when the simultaneous fitting of >20 elemental abundances is required. Because we require our model to be differentiable, this can be made more tractable with alternative sampling techniques like the Hamiltonian Monte Carlo (HMC) algorithm (Duane et al. 1987). Even so, this is still a very computationally expensive exercise to do for every instrument and observational combination.

A more efficient way to obtain the width of the distribution (and in turn the precision on each label) is with the CRLB. Within astrophysics, the CRLB has been used extensively in cosmological contexts (e.g., Albrecht et al. 2006; Adshead et al. 2008; Wang 2010; Becker et al. 2012; Betoule et al. 2014; Font-Ribera et al. 2014; King et al. 2014; Eriksen et al. 2015), but has only recently been applied to abundance measurements from full-spectrum stellar spectroscopy (Ting et al. 2016, 2017a).[6]

Formally, the CRLB is the highest possible precision achievable for a set of observations and can be derived from the FIM,

$$F_{\alpha\beta} = E\left[\frac{\partial^2\left[-\ln L(D|\theta)\right]}{\partial\theta_\alpha\partial\theta_\beta}\right]_{\hat{\theta}}, \quad (2.2)$$

where $E[.]$ denotes the expectation value, $\hat{\theta}$ is the maximum likelihood estimate, and α and β are each a specific label. In simpler terms, the FIM describes how fast the likelihood function declines for each label around the maximum likelihood point. The steeper the decline, the narrower the distribution, and the more precisely a label can be measured.

[5]Note that the number of grid points needed to fully sample the likelihood scales exponentially with the number of dimensions.

[6]Ireland (2005) first applied the CRLB formalism to stellar spectroscopy in their analysis of the limiting precision of Solar emission lines. Hansen et al. (2015) later used CRLBs to quantify the precision of EW measurements of blended stellar absorption lines.

Using the Cramér–Rao inequality, this curvature can be related directly to the width of the Gaussian likelihood. Specifically, the inverse of the FIM gives the lower bound on the covariance matrix of the labels,

$$K_{\alpha\beta} \geq (F^{-1})_{\alpha\beta} \tag{2.3}$$

or in terms of measurement uncertainty,

$$\sigma_\alpha \geq \sqrt{(F^{-1})_{\alpha\alpha}}. \tag{2.4}$$

This lower bound on the measurement uncertainty, σ_α, is the CRLB for the label α.

In order to apply CRLBs to the fitting of stellar spectra, we must make two fundamental assumptions:

(i) The observed spectra have Gaussian noise, and the likelihood of the spectra given our model is well described by a multivariate Gaussian.

(ii) The spectral models accurately reproduce the observed spectra (i.e., the fitting is free of systematic errors and $\hat{\theta}$ is an unbiased estimator of the true labels)[7].

Assuming Gaussianity (i) is standard practice in the fitting of stellar spectra with S/N $>$ 10 pixel^{-1} and enables substituting Equation (2.1) for the log-likelihood in Equation (2.2).

Though rarely strictly true, the assumption of accurate models (ii) is commonplace across all of astronomy and astrophysics. Model fidelity is a necessary assumption in all matters of parameter estimation, and so we too assume the stellar models to be correct though we know them to have flaws and oversimplifications (e.g., 1D LTE atmospheres, mixing length theory, incomplete line lists, miscalibrated oscillator strengths). It is important to remember that the CRLBs we calculate are predictions of precision, not accuracy. And while they may be challenging to achieve in practice due to various systematics (see Section 2.7.4 for further discussion), they nevertheless provide useful guidance for stellar abundance work (see Section 2.5.1.1 and Appendix 2.10.4 for a comparison of CRLBs with the abundance precision measured in practice).

Under the assumption of perfect models, we can replace $f_{\rm obs}(\lambda_i)$ in Equation (2.1) with $f(\lambda_i, \hat{\theta})$, noting that $\hat{\theta}$, as an unbiased estimator, corresponds to the true stellar labels. Combined with the assumption of a multivariate Gaussian log-likelihood, we can re-write Equation (2.2) in terms of the gradient spectra as

$$F_{\alpha\beta} = \left[\frac{\partial f(\lambda, \boldsymbol{\theta})}{\partial \theta_\alpha}\right]^T_{\hat{\theta}} \Sigma^{-1} \left[\frac{\partial f(\lambda, \boldsymbol{\theta})}{\partial \theta_\beta}\right]_{\hat{\theta}} + \frac{1}{2}\mathrm{tr}\left(\Sigma^{-1}\frac{\partial \Sigma}{\partial \theta_\alpha}\Sigma^{-1}\frac{\partial \Sigma}{\partial \theta_\beta}\right) \tag{2.5}$$

as worked out in Kay (1993). Because in the context of stellar spectra the covariance matrix of the normalized flux, Σ, is independent of the stellar labels, the second term in Equation

[7]The CRLB can be generalized to relax the assumption that $\hat{\theta}$ is an unbiased estimator (see Appendix 2.10.1), but this requires knowing the bias of $\hat{\theta}$ as a function of the stellar labels, which is beyond the scope of this paper.

(2.5) vanishes, leaving the FIM as the quadrature sum of the gradient spectra across all wavelength pixels, weighted by the uncertainty of the normalized flux:

$$F_{\alpha\beta} = \left[\frac{\partial f(\lambda,\theta)}{\partial\theta_\alpha}\right]_{\hat{\theta}}^{T} \Sigma^{-1} \left[\frac{\partial f(\lambda,\theta)}{\partial\theta_\beta}\right]_{\hat{\theta}}. \tag{2.6}$$

Using this form of the FIM, we can now write the CRLB in terms of the spectral gradients as

$$\sigma_\alpha = \left(\left[\frac{\partial f(\lambda,\theta)}{\partial\theta_\alpha}\right]_{\hat{\theta}}^{T} \Sigma^{-1} \left[\frac{\partial f(\lambda,\theta)}{\partial\theta_\alpha}\right]_{\hat{\theta}}\right)^{-1/2}. \tag{2.7}$$

Equation (2.7) shows that the CRLB is sensitive to the factors that affect the information content of spectra as discussed in Section 2.3.1. More specifically, if the gradient of the spectrum with respect to a given label is high ($\partial f(\lambda,\theta)/\partial\theta_\alpha$ is large), then σ_α is small and more precise measurements are possible.

Similarly, having high S/N (Σ^{-1} is large) in informative regions of the spectrum will also result in small σ_α and high possible precision. Larger wavelength coverage and higher wavelength sampling mean summing over more pixels and thus higher precision, provided that the pixels are informative and not highly correlated. The importance of instrumental resolution is embedded in the matrix multiplication, where higher resolution gradients lead to deeper spectral features and less blended features, resulting in smaller covariances between stellar labels.

An analytic description of the resolution dependence of the CRLBs is presented in Ting, Conroy, Rix, and Cargile (2017), which we summarize here:

(i) The rms depth per pixel (and information) of an absorption feature in the gradient spectrum scales as R.

(ii) For fixed exposure time and stellar flux, the S/N scales as $R^{-1/2}$ due to Poisson statistics.

(iii) For fixed number of detector pixels, the wavelength range scales as $1/R$. Assuming that absorption features are evenly distributed in wavelength space, the information content scales as $R^{-1/2}$ since information adds in quadrature.

(iv) Together, the simple arguments in (i)–(iii) show that to first order the stellar label precision is independent of spectral resolving power.

We add to this analytic description that, similar to (iii), the information content scales as $n^{-1/2}$, where n is the number of independent pixels per resolution element. In the extreme case where all n pixels in a resolution element are 100% correlated, the CRLB will be n larger than if the pixels were entirely uncorrelated. We present a more detailed exploration of the effects of sampling and pixel-to-pixel correlation on the CRLBs in Appendix 2.10.3.

For a given spectral model (i.e., 1D LTE, as we employ in this work, or 3D non-LTE when they become widely available), forecasting abundance precision is reduced to a matter of calculating derivatives and multiplying matrices. Furthermore, because most spectra have thousands, if not tens of thousands, of pixels, the central limit theorem can be used to show that the CRLB becomes theoretically attainable (i.e., Equation (2.3) becomes an equality if all assumptions hold). CRLBs are thus an incredibly valuable tool for efficiently exploring the possible precision of a large number of instrumental and observational scenarios when the high dimensionality of the problem makes more rigorous sampling techniques costly or unfeasible.

2.3.2.1 Incorporating Prior Information

In many cases, there may be additional knowledge of the stellar properties beyond the spectra in hand. For example, in an extragalactic context, we may know the distance to the host galaxy quite well and/or we may have photometry of the star. Such information can give external constraints on the luminosity, surface gravity, temperature, and even metallicity of a star, and can be used to improve the spectral fitting process. We now demonstrate how this information can be included in the CRLB calculation.

While the CRLB was initially derived in a frequentist context, a Bayesian equivalent of the CRLB can be formulated for application to scenarios in which prior information on the stellar labels is available. This is done by replacing the log-likelihood in Equation (2.1) with the full Bayesian probability:

$$\ln P(\theta|D) = \ln \Pi(\theta) + \ln L(D|\theta), \tag{2.8}$$

where $\Pi(\theta)$ is the prior on the stellar labels. This results in the following equation for the Bayesian FIM:

$$F_{\text{Bayes}} = F_{\text{spec}} + F_{\text{prior}}. \tag{2.9}$$

Appendix A of Echeverria et al. (2016) presents a detailed derivations of Equation (2.9).

The first term on the right-hand side of the equation is the standard spectral gradient FIM found previously (Equation (2.6)). The second term on the right-hand side of the equation is the FIM of the prior and encapsulates the additional information included in the prior. It can be shown that for Gaussian priors with standard deviation $\sigma_{\text{prior},\alpha}$ for each stellar label, the prior FIM is the diagonal matrix

$$F_{\text{prior},\alpha\alpha} = \left(\frac{1}{\sigma_{\text{prior},\alpha}}\right)^2. \tag{2.10}$$

As a result, we can write the Bayesian CRLB of a stellar label, α, with Gaussian priors as

$$\sigma_{\text{Bayes},\alpha} = \sqrt{(F_{\text{spec}} + F_{\text{prior}})^{-1}_{\alpha\alpha}} \tag{2.11}$$

$$= \left(\left[\frac{\partial f(\lambda,\theta)}{\partial \theta_\alpha}\right]^T_{\hat{\theta}} \Sigma^{-1} \left[\frac{\partial f(\lambda,\theta)}{\partial \theta_\alpha}\right]_{\hat{\theta}} + \frac{1}{\sigma^2_{\text{prior},\alpha}}\right)^{-1/2}. \tag{2.12}$$

As a check, we note that in the case of weak priors or strongly informative data, the CRLBs approach the value predicted by Equation (2.7), while in the case of strong priors or uninformative data, the CRLBs approach the standard deviation of the priors.

2.3.2.2 Combining Information From Multiple Spectra

The CRLB can also be applied to the context in which multiple disjoint spectra of the same star exist across different wavelength ranges and resolutions, but are to be fit together. Such cases commonly arise for multi-armed spectrographs (e.g., the Low Resolution Imaging Spectrometer (LRIS) on Keck, Multi-Object Double Spectrographs (MODS) on the LBT, and DESI) and for echelle spectrographs, which observe multiple discrete orders of the stellar spectrum (e.g., GIRAFFE on the VLT).

Replacing the log-likelihood in Equation (2.2) with the sum of the log-likelihoods for each spectra and following through the previous derivation (Equations 2.5–2.7) reveals that the relevant FIM for the joint fitting is simply the sum of the FIM of the individual sepctra. This is equivalent to concatenating the gradient spectra and covariance matrices of each observation together and using these combined quantities in Equation (2.7). This can be done for arbitrary combinations of stellar spectra provided that the covariance of overlapping wavelength ranges is properly accounted for (as done in Czekala et al. 2015), otherwise the number of independent information-carrying pixels is artificially inflated.

2.4 Methods

In this section, we outline our process of generating synthetic stellar spectral gradients and using them to compute CRLBs for a variety of stars, observing scenarios, and spectrographs. We begin by describing the non-exhaustive scope of instruments (Section 2.4.1.2) and stellar targets (Section 2.4.1.1) considered in this work. In Section 2.4.1.3, we describe the determination of realistic S/N estimates for each spectrograph and stellar target. Lastly, we walk through our methodology for generating gradient spectra in Section 2.4.2. The technical details of the matrix multiplication and inversion used to calculate the CRLBs can be found in Appendix 2.10.2.

2.4.1 Observational Scope

While the CRLB is broadly applicable to the entire field of resolved star spectroscopy, we choose to focus this work on forecasting the precision possible for spectroscopy of stars outside of the MW. In general, this limits the scope of this work to large-aperture ground- and space-based telescopes observing faint, metal-poor RGB stars at low and moderate resolution ($R < 10000$). In the rest of this section, we describe in detail our choice of targets, instruments, and observing conditions.

Phase	M_V	$T_{\rm eff}$ (K)	$\log g$	$v_{\rm turb}$ (km/s)	$\log Z$
RGB	-0.5	4200	1.5	2.0	-0.5
RGB	-0.5	4530	1.7	1.9	-1.0
RGB	**-0.5**	**4750**	**1.8**	**1.9**	**-1.5**
RGB	-0.5	4920	1.9	1.9	-2.0
RGB	-0.5	5050	1.9	1.9	-2.5
MSTO	3.5	6650	4.1	1.2	-1.5
TRGB	-2.5	4070	0.5	2.3	-1.5

Table 2.1: Stellar labels of the stars considered in this work. The bold line designates the fiducial stellar reference used throughout this study. All stars have solar abundance patterns. $T_{\rm eff}$ and $\log g$ are determined from MIST isochrones given the age (10 Gyr), metallicity, and absolute magnitude of the star. $v_{\rm turb}$ is found using the scaling relationship presented in Holtzman et al. (2015). For $\log Z = -1.5$, $M_V = -0.5$ corresponds to a star roughly halfway up the RGB; for more metal poor stars, the same magnitude corresponds to stars lower on the RGB closer to the main sequence turn-off (see Figure 2.2).

2.4.1.1 Properties of Reference Stars

In this work, we limit our analysis to the stars predominantly accessible to spectroscopic campaigns of extragalactic stellar populations: metal-poor RGB stars. We also consider how the CRLBs vary from this fiducial star along several axes, including apparent magnitude, metallicity, and evolutionary phase as described below. The stellar labels used for these reference stars can be found in Table 2.1. Their position on the Kiel and Hertzprung–Russell diagrams can be seen in Figure 2.2.

For each of the stellar targets considered in this work, we determine the effective temperature and surface gravity of the star using an isochrone from the MESA Isochrones and Stellar Tracks (MIST) project corresponding to the age, metallicity, and absolute magnitude of the star(Paxton, Bildsten, et al. 2011; Paxton, Cantiello, et al. 2013; Paxton, Marchant, et al. 2015; Choi et al. 2016; Dotter 2016; Paxton, Schwab, et al. 2018). As was done in Ting, Conroy, Rix, and Cargile (2017), we assume a microturbulent velocity for each star using the relationship between microturbulent velocity and surface gravity found by Holtzman et al. (2015):

$$v_{\rm turb} = 2.478 - 0.325 \log g \text{ km/s} \tag{2.13}$$

Fiducial Star—We adopt as our fiducial stellar reference a star that is roughly halfway up the RGB with a V-band absolute magnitude of $M_{V,\rm Vega} = -0.5$ ($M_{g,\rm AB} \sim -0.2$). This choice splits the difference between the brighter but rarer stars at the tip of the RGB (TRGB) and the more numerous but fainter main-sequence turn-off (MSTO) stars. Furthermore, we assume that this fiducial star is 10 Gyr old, has a metallicity of $\log(Z/Z_\odot) = -1.5$, and has solar abundance patterns.

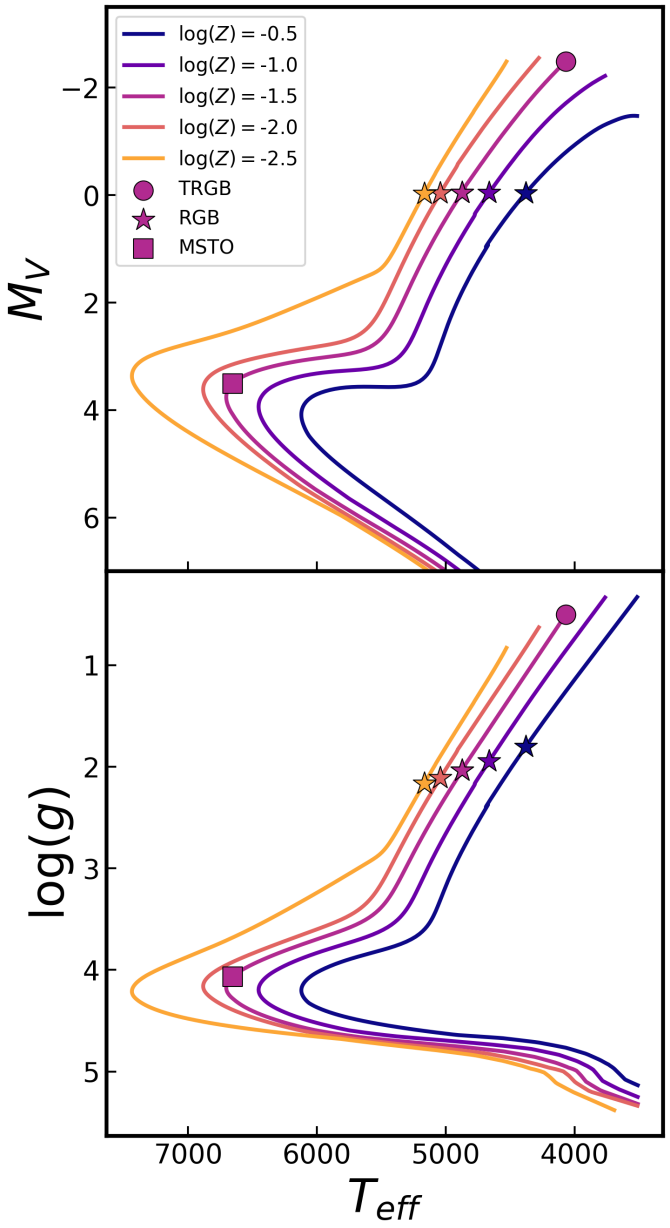


Figure 2.2: Hertzsprung-Russell (top) and Kiel (bottom) diagrams of the seven reference stars considered in this work (see Table 2.1). Shapes denote stellar evolutionary phase and colors denote metallicity. The five RGB stars of differing metallicity were chosen to have the same V-band absolute magnitude and thus lie on slightly different portions of the RGB. Solid lines are MIST isochrones of a 10 Gyr old main sequence and red giant branch.

Apparent Magnitude—As can be seen from Equation (2.7), the CRLB scales inversely proportional to the S/N of the spectrum. We consider our fiducial star with apparent magnitudes $m_V = 18$, 19.5, and 21, but at fixed stellar evolutionary phase, to avoid conflating the effects of S/N and the atmospheric parameters of the star. This amounts to observing an identical star at distances of ∼50, 100, and 200 kpc, which are typical distances to nearby MW satellites. When not evaluating the effects of S/N on the chemical abundance precision, we assume the star is located at a distance of 100 kpc ($m_V = 19.5$).

Metallicity—We also investigate how the the information content of an RGB star spectrum changes as its metallicity decreases from $\log \log Z = -0.5$ to -2.5. Because the shape of the RGB changes as a function of metallicity, we make this comparison at fixed M_V instead of at fixed evolutionary phase. As a result, the lower-metallicity stars considered in this work are located farther down the RGB (i.e., have higher effective temperature and surface gravity; see Figure 2.2).

Evolutionary Phase—To isolate the effect of stellar evolutionary phase on the chemical abundance precision, we compare the CRLBs of our fiducial RGB star to that of an MSTO or RGB star of the same metallicity and apparent brightness.

2.4.1.2 Instruments

Because the stars we consider in this work are so faint ($m_V = 19.5$), we limit our forecasts to instruments, both existing and planned, that can efficiently acquire spectra with modest S/N (>15 pixel^{-1}) in reasonable amounts of time (<1 night). In practice, this includes instruments on ground-based telescopes with >5 meter apertures and large-aperture space telescopes. This excludes most of the spectrographs responsible for large MW surveys (e.g., RAVE; Steinmetz et al. (2006), SEGUE; Yanny et al. (2009), LAMOST; Luo et al. (2015) GALAH; De Silva et al. (2015), and APOGEE; Majewski et al. (2017)) and most spectrographs with very high resolving powers ($R > 50000$). We do not include any instruments with very low resolving powers ($R < 1000$), though there is reason to believe that the information content accessible to very low-resolution grism spectroscopy is still considerable (Bailer-Jones 2000). Lastly, the line lists[8] we use to generate synthetic spectra are limited in extent to wavelengths between 3000 Å and 1.8 μm. As such, we exclude instruments observing in the ultraviolet (UV) and infrared (IR) despite the significant chemical information that these wavelength regimes contain (e.g., García Pérez et al. 2016; Roederer 2019; Ting, Conroy, Rix, and Cargile 2019).

Even with the aforementioned restrictions, the list of spectrographs already on sky suitable for extragalactic stellar spectroscopy is extensive. As shown in Table 2.2, we consider 12 existing spectrographs at five world-class observing facilities as well as 9 spectrographs that will be coming online within the next decade. Each of these instruments features numerous choices of observing modes, dispersive elements, and other specifications. This flexibility enables a broad range of science, but makes an exhaustive evaluation of each observing con-

[8] `http://kurucz.harvard.edu/`

figuration infeasible. Instead, we consider only the setups that we believe most relevant to acquiring precise chemical abundances in extragalactic stellar populations for a total of 41 configurations.[9] For each observational setup, we attempt to use realistic wavelength coverage, wavelength sampling, and resolving power as reported either in literature or in design documents.

Despite an extensive literature search, not all pertinent spectrograph details were readily available, and we had to make some assumptions. For example, for several instruments, the number of pixels per resolution element could not be found; in these cases we adopt a fiducial wavelength sampling of 3 pixels/FWHM as assumed in Ting, Conroy, Rix, and Cargile (2017). For multi-object spectrographs (MOSs), we assume the nominal wavelength coverage for a star observed in the center of the instrument field of view and ignore the variations in wavelength coverage incurred for off-center stars. Additionally, most instruments have wavelength-dependent resolving powers, usually decreasing toward the blue. The manner in which the resolving power changes across the spectrum, known as the line-spread function (LSF), depends on the position of the star in the slit and can vary from slit to slit. For simplicity, we assume all instruments have a fixed LSF with a resolution approximately equal to the average across the entire spectrum.

Lastly, while we do compare and contrast the forecasted precision of these instruments, we emphasize that the "best" instrument is largely of a science-dependent nature. There are numerous trade-offs between field of view and multiplexing (see Table 2.3), radial velocity precision, and detailed chemical abundance measurements. Balancing them is a matter of their relative importance to the science at hand.

Telescope/ Instrument	Spectroscopic Configuration	Wavelength Range (Å)	R ($\lambda/\Delta\lambda$)	Sampling (Pixels/FWHM)	Aperture (m)	Section	Refs
Existing Instruments							
Keck II/DEIMOS[a]	1200G	6500–9000	6500	4	10.0	2.5.1	[1]
	1200B	4000–6400	4000	4	10.0	2.5.2.1	[1]
	600ZD	4100–9000	2500	5	10.0	2.5.2.1	[1]
	900ZD	4000–7200	2500	5	10.0	2.5.2.1	[1]
Keck I/LRIS[a]	600/4000	3900–5500	1800	4	10.0	2.5.2.1	[2]
	1200/7500	7700–9000	4000	5	10.0	2.5.2.1	[2]
Keck I/HIRESr[b]	B5 Decker	3900–8350	49000	3	10.0	2.5.3.1	[3]
	C5 Decker	3900–8350	35000	3	10.0	2.5.3.1	[3]
LBT/MODS[a]	Blue Arm	3200–5500	1850	4	11.8	2.5.2.2	[4]
	Red Arm	5500–10500	2300	4	11.8	2.5.2.2	[4]
Magellan/MIKEr[b]	Blue (1″.0 slit)	3500–5000	28000	4	6.5	2.5.3.1	[5]
	Red (1″.0 slit)	5000–10000	22000	3	6.5	2.5.3.1	[5]
Magellan/M2FS[c]	HiRes	5130–5185	18000	3†	6.5	2.5.3.2	[6]
	MedRes	5100–5315	10000	3†	6.5	2.5.3.2	[6]
MMT/Hectochelle[c]	RV31	5160–5280	20000	6	6.5	2.5.3.2	[7]
MMT/Hectospec[a]	270 mm^{-1}	3900–9200	1500	5	6.5	2.5.2.2	[8]
	600 mm^{-1}	5300–7800	5000	5	6.5	2.5.2.2	[8]
MMT/Binospec[a]	270 mm^{-1}	3900–9200	1300	4	6.5	2.5.2.2	[9]
	600 mm^{-1}	4500–7000	2700	3	6.5	2.5.2.2	[9]
	1000 mm^{-1}	3900–5400	3900	3	6.5	2.5.2.2	[9]

[9]This list is extensive but far from complete. We encourage readers interested in spectrographs not listed in Table 2.2 to calculate their own chemical abundance precision using the `Chem-I-Calc` python package detailed in Section 2.8.

VLT/MUSE[d]	Nominal	4800–9300	2500	3†	8.2	2.5.2.2	[10]
VLT/X-SHOOTER[b]	UVB (0″.8 slits)	3000–5500	6700	5	8.2	2.5.3.1	[11]
	VIS (0″.7 slits)	5500–10200	11400	4	8.2	2.5.3.1	[11]
	NIR (0″.9 slits)	10200–18000	5600	4	8.2	2.5.3.1	[11]
VLT/FLAMES-UVES[e]	r580	4800–6800	40000	5	8.2	2.5.3.1	[12]
VLT/FLAMES-	LR8	4200–11000	6500	3†	8.2	2.5.3.2	[13]
GIRAFFE[c]	HR10	5340–5620	19800	3†	8.2	2.5.3.2	[13]
	HR13	6120–6400	22500	3†	8.2	2.5.3.2	[13]
	HR14A	6400–6620	28800	3†	8.2	2.5.3.2	[13]
	HR15	6620–6960	19300	3†	8.2	2.5.3.2	[13]
Future Instruments							
JWST/NIRSpec[a]	G140M/F070LP	7000–12700	1000	3†	6.5	2.6.1	[14]
	G140M/F100LP	9700–18400	1000	3†	6.5	2.6.1	[14]
	G140H/F070LP	8100–12700	2700	3†	6.5	2.6.1	[14]
	G140H/F100LP	9700–18200	2700	3†	6.5	2.6.1	[14]
GMT/GMACS[a]	Blue Arm (LR)	3200–5500	1000	3	24.5	2.6.2	[15]
	Blue Arm (MR)	3700–5500	2500	3	24.5	2.6.2	[15]
	Blue Arm (HR)	4200–5000	5000	3	24.5	2.6.2	[15]
	Red Arm (LR)	5500–10000	1000	3	24.5	2.6.2	[15]
	Red Arm (MR)	6100–8900	2500	3	24.5	2.6.2	[15]
	Red Arm (HR)	6700–8300	5000	3	24.5	2.6.2	[15]
GMT/G-CLEF[a]	Med Res	3000–9000	35000	3	24.5	2.6.2	[16]
TMT/WFOS[a]	B1210	3100–5500	1500	3†	30.0	2.6.2	[17]
	B2479	3300–4750	3200	3†	30.0	2.6.2	[17]
	B3600	3250–4100	5000	3†	30.0	2.6.2	[17]
	R680	5500–10000	1500	3†	30.0	2.6.2	[17]
	R1392	5850–8400	3200	3†	30.0	2.6.2	[17]
	R2052	5750-7250	5000	3†	30.0	2.6.2	[17]
E-ELT/MOSAIC[a]	HMM-Vis	4500–8000	5000	4	39.0	2.6.2	[18]
	HMM-NIR	8000–18000	5000	3	39.0	2.6.2	[18]
Subaru/PFS[a]	Blue Arm	3800–6300	2300	4	8.2	2.6.3	[19]
	Red Arm (LR)	6300–9400	3000	4	8.2	2.6.3	[19]
	Red Arm (MR)	7100–8850	5000	4	8.2	2.6.3	[19]
	NIR Arm	9400–12600	4300	4	8.2	2.6.3	[19]
MSE[a]	Blue Arm (MR)	3900–5000	5000	3	11.3	2.6.3	[20]
	Green Arm (MR)	5750–6900	5000	3	11.3	2.6.3	[20]
	Red Arm (MR)	7370–9000	5000	3	11.3	2.6.3	[20]
	All Arms (LR)	3600–13000	3000	3	11.3	2.6.3	[20]
Keck/FOBOS[a]	Proposed	3100–10000	3500	6	10.0	2.6.3	[21]
4-meter Instruments							
LAMOST[a]		3700–9000	1800	3†	4.0	App. 2.10.4	[22]
Mayall/DESI[a]	Blue Arm	3600–5550	2500	3	4.0	App. 2.10.6	[23]
	Red Arm	5550–6560	3500	3	4.0	App. 2.10.6	[23]
	Infrared Arm	6560–9800	4500	3	4.0	App. 2.10.6	[23]

Table 2.2: This table lists the spectroscopic configurations we adopt for computing the chemical abundance precision as well as the section in which those precisions are presented. For each instrument, we adopt a constant resolution and number of pixels per resolution element across the wavelength range indicated. The instruments listed here span a large range in wavelength coverage (3200 Å–1.8 μm), resolving powers ($1000 < R < 49000$), and instrument designs. **References.** [1] Faber et al. (2003), [2] Oke et al. (1995), [3] Vogt et al. (1994), [4] Pogge et al. (2010), [5] Bernstein et al. (2003), [6] Mateo et al. (2012), [7] Szentgyorgyi, Furesz, et al. (2011), [8] Fabricant, Fata, Roll, et al. (2005), [9] Fabricant, Fata, Epps, et al. (2019), [10] Bacon et al. (2010), [11] Vernet et al. (2011), [12] Dekker et al. (2000), [13] Pasquini et al. (2002), [14] Bagnasco et al. (2007), [15] DePoy et al. (2012), [16] Szentgyorgyi, Baldwin, et al. (2016), [17] Pazder et al. (2006), [18] Jagourel et al. (2018), [19] Tamura et al. (2018), [20] MSE Science Team et al. (2019), [21] Bundy et al. (2019), [22] Cui et al. (2012), [23] DESI Collaboration, Aghamousa, et al. (2016)
[a]Low-/Medium-Resolution Multi-Object Spectrograph
[b]Single-Slit Multi-Order Echelle Spectrograph
[c]Multi-Object Single-Order Echelle Spectrograph
[d]Integral Field Unit Spectrograph
[e]Multi-Object Multi-Order Echelle Spectrograph
[†]Sampling information was not found so a nominal value of 3 pixels/FWHM is assumed.

2.4.1.3 Observing Conditions and Integration Time

We assume the the flux covariance, Σ, is due entirely to photon noise and thus is a function solely of exposure time, instrument throughput, observing conditions, and the brightness of the star, ignoring any uncertainty introduced by imperfect data reduction or continuum normalization.[10] Whenever possible, we use the exposure time calculator (ETC) specific to each instrument listed in Table 2.4. This allows us to adopt a flux covariance as specific as possible to each facility and accordingly compute realistic CRLBs. For instruments that do not have public ETCs, we scale the S/N from a similar instrument according to

$$\mathrm{S/N} \propto D(nR)^{-1/2}, \tag{2.14}$$

where D is the effective aperture of the telescope, R is the instrumental resolving power, and n is the instrumental wavelength sampling.

For our S/N calculations we assume an airmass of 1.1 and a seeing of $0''.75$ (or as close to these values as possible with each ETC). We assume read-noise is negligible such that the

[10]Reliably determining the (pseudo-)continuum in practice is challenging and is a potential source of systematic errors (see Section 2.7.4). However, self-consistently normalizing both the observed and model spectra can mitigate these systematics. Evaluating these effects is beyond the scope of this paper.

Telescope/Instrument	Field of View	$N_{\rm slits}$ or $N_{\rm fibers}$
Keck II/DEIMOS	$16' \times 4'.0$	100
Keck I/LRIS	$6'.0 \times 7'.8$	40
Keck I/HIRESr	...	1
Magellan/MIKEr	...	1
Magellan/M2FS	$30''.0$	250
MMT/Hectochelle	$1°.0$	240
MMT/Hectospec	$1°.0$	300
MMT/Binospec	$16'.0 \times 15'.0$	150
VLT/MUSE	$1'0 \times 1'.0$	...
VLT/X-SHOOTER	...	1
VLT/FLAMES-UVES	$25'.0$	8
VLT/FLAMES-GIRAFFE	$25'.0$	130
LBT/MODS	$6'.0 \times 6'.0$	50
JWST/NIRSpec	$3'.0 \times 3'.0$	100
Mayall/DESI	$2°.8$	5000
Subaru/PFS	$1°.3$	2400
MSE	$9'.5$	3250
Keck/FOBOS	$20'.0$	1800
GMT/GMACS	$7'.4$	100
GMT/GMACS+MANIFEST	$20'$	100s
GMT/G-CLEF+MANIFEST	$20'$	40
TMT/WFOS	$4'.2 \times 9'.6$	600
E-ELT/MOSAIC (HMM-Vis)	$6'.0$	200
E-ELT/MOSAIC (HMM-NIR)	$6'.0$	100

Table 2.3: Field of view and multiplexing of instruments. $N_{\rm slits}$ ($N_{\rm fibers}$) is the approximate number of slits (fibers) that an instrument can use in a single pointing. This provides a rough estimate for the number of stars a spectrograph can observe simultaneously. In practice, of course, not all slits/fibers can be placed on stars because some may be required for guiding, alignment, or sky-subtraction, while others may go unused simply due to the distribution of stars in the field. Single numbers for the field of view (FoV) indicate the diameter of the FoV, while pairs of number indicate the approximate rectangular dimensions of the FoV. For single-slit spectrographs, the field of view is irrelevant for resolved star spectroscopy. As an IFU, the multiplexing of MUSE depends on the density of stars in the field and the source extraction method employed.

S/N of a single one-hour exposure is the same as that of four 15 minute exposures stacked together.

Because not all ETCs provide the same stellar spectral energy distribution (SED), we use a K0I, K2V, or K0V spectral template (in preferential order when provided) to best match the SED of our fiducial RGB star. Additionally, we use a K0V spectral template for the RGB reference stars with $\log Z \leq -1.5$ and a K5V spectral template for the RGB stars with $\log Z > -1.5$. For the $\log Z = -1.5$ MSTO and TRGB reference stars we use G5V and K5III/K5V stellar templates respectively.

Once calculated by the ETC, the S/N is interpolated onto the same wavelength grid as the stellar spectra corresponding to the resolving power, spectral sampling, and wavelength range of that instrument.

Because most spectrographs are designed to slightly oversample the spectrum ($\geq$3 pixels/FWHM), adjacent pixels are not completely uncorrelated, though most stellar abundance studies treat them as such (see however Czekala et al. 2015). For simplicity, we also assume no correlations between adjacent wavelength pixels so that we can write the covariance matrix of the normalized flux, Σ, as the diagonal matrix

$$\Sigma = \begin{bmatrix} \sigma^2(\lambda_1) & & \\ & \ddots & \\ & & \sigma^2(\lambda_N) \end{bmatrix}, \tag{2.15}$$

where $\sigma^2(\lambda_i) = (\mathrm{S/N})^{-2}$ is the variance in each pixel. A more accurate treatment of the pixel-to-pixel covariance would effectively reduce the number of independent information-carrying pixels in the spectrum, increasing the CRLB slightly—recall that the CRLB is proportional to $n^{-1/2}$, where n is the number of independent pixels per resolution element. A more in-depth analysis of pixel correlation and wavelength sampling is presented in Appendix 2.10.3.

The large variety of resolving powers included in this work means that a universal "observing strategy" cannot be applied to all instruments. Instead, we consider separate observing setups for a fiducial spectrograph, low- and medium-resolution spectrographs ($R < 10000$), high-resolution spectrographs ($R > 10000$), and JWST/NIRSpec, which we describe below. A summary of all of the relevant assumptions used in the S/N calculation of each instrument is contained in Table 2.4.

Fiducial Spectrograph—To investigate the effects of exposure time, object brightness, and stellar evolutionary phase and metallicity, we adopt the 1200G grating on Keck/DEIMOS as our fiducial spectroscopic setup. We consider 1, 3, and 6 hr integration times and stars with $m_V = 18$, 19.5, and 21. For comparisons of metallicity and stellar evolutionary phase, we hold the integration time and apparent magnitude fixed at 1 hr and $m_V = 19.5$, respectively.

Low- and Medium-resolution Spectrographs—For spectrographs with $R < 10000$, we consider the baseline observing strategy to be 1 hr of integration of our fiducial $m_V = 19.5$ RGB star. This is generally sufficient for spectrographs on 6+ meter telescopes to achieve $\mathrm{S/N} > 15\,\mathrm{pixel}^{-1}$ across the optical spectrum. In this category, we include DEIMOS, LRIS, and FOBOS on Keck; Hectospec and Binospec on the MMT; the Multi Unit Spectroscopic

Explorer (MUSE) on the VLT; LBT/ MODS on the LBT; PFS on Subaru; MSE; the GMT Multi-object Astronomical and Cosmological Spectrograph (GMACS) on the GMT; the Wide Field Optical Spectrometer (WFOS) on the TMT; and the Multi-Object Spectrograph (MOSAIC) on the E-ELT.

The GMACS ETC provides two sample settings, each of which assume a constant $\delta\lambda$ across both the blue and red channels, resulting in wavelength-dependent resolutions. We choose the higher resolution setting ($\Delta\lambda = 1.4$) and scale the S/N at each pixel according to S/N $\propto R^{-1/2}$ to match the constant resolving power we are attempting to emulate. Because ETCs do not yet exist for MOSAIC, we scale the S/N from GMACS for the MOSAIC visual high multiplex mode (HMM-Vis) and from NIRSpec for the MOSAIC near-infrared high multiplex mode (HMM-NIR) according to Equation (14).[11]

High-resolution Spectrographs—Due to the higher dispersion and generally lower throughput of high-resolution spectrographs, a single hour of integration is insufficient to achieve adequate S/N ($>15\,\mathrm{pixel}^{-1}$) for a $m_V = 19.5$ RGB star. Instead we consider an integration of 6 hr ($\sim$1 night of observing). Instruments in this category include the High Resolution Echelle Spectrometer (HIRES) on Keck; the Magellan Inamori Kyocera Echelle (MIKE) on Magellan and the Michigan/Magellan Fiber System (M2FS) on Magellan; Hectochelle on MMT; X-SHOOTER,[12] GIRAFFE, and the Ultraviolet and Visual Echelle Spectrograph (UVES) on the VLT; and the GMT Consortium Large Earth Finder (G-CLEF) on GMT. M2FS and Hectochelle do not have public ETCs so we scale the average S/N from the GIRAFFE HR10 ETC according to Equation (2.14) and assume the S/N is roughly constant over the short wavelength range observed by these instruments.

JWST/NIRSpec—The strength of JWST/NIRSpec is its high sensitivity and high angular resolution. The most likely use case will be to acquire spectra in distant and/or crowded environments, which may require longer integration times than our fiducial 1 hr setup for ground-based, low-resolution instruments. Thus, for JWST only, we adopt a 6 hr of integration on an $m_V = 21$ TRGB star.[13] This scenario is chosen to mimic the observation of bright stars in the disk of M31 or in a galaxy at the edge of the Local Group.

Beyond 1 Mpc—To investigate the distance to which JWST/NIRSpec and GMT/GMACS (as a representative ELT) can provide useful chemical measurements, we additionally hold the exposure time constant at 6 hr and systematically decrease the apparent magnitude of our target TRGB star from mV = 21–26. This corresponds to observing a TRGB star at distances between 0.5 and 5 Mpc

[11]By using the ETC of space-based NIRSpec for MOSAIC (HMM-NIR), we ignore a number of telluric features that affect observations in the NIR.

[12]Despite the more moderate resolution of the X-SHOOTER UVB and NIR arms, we include X-SHOOTER with the other high-resolution spectrographs due to its higher resolution VIS arm and single-slit echelle design.

[13]Specifically we assume three exposures each of which includes one integration of 170 groups (subintegrations) for a total exposure time of 6 hr, 5 minutes, and 35 s.

Instrument	m_V	t_{exp} (hours)	Airmass	Seeing	Slitwidth/ Fiber Diameter	Spatial×Spectral Binning	Stellar Template	ETC
DEIMOS	18.0, 19.5, 21.0	1, 3, 6	1.1	0″.75	0″.75	1×1	G5V, K0V, K5V	[1]
LRIS‡	19.5	1	1.1	0″.75	0″.70	1×1	K0V	[2]
HIRESr (B5/C5)	19.5	6	1.1	0″.75	0″.86/1″.10	2×2	K0V	[3]
MIKE	19.5	6	1.1	0″.75	1″.00	3×1	K0V	[4]
M2FS	19.5	6	1.1	0″.75	1″.20	2×2	K2V	[5]†
Hectochelle	19.5	6	1.1	0″.75	1″.00	3×2	K2V	[5]†
Hectospec	19.5	1	1.1	0″.75	1″.5	1×1	K0V	[6]
Binospec	19.5	1	1.1	0″.75	1″.0	1×1	K0V	[6]
MUSE	19.5	1	1.1	0″.80	...*	$(3 \times 3) \times 1$	K2V	[7]
X-SHOOTER (UVB/VIS/NIR)	19.5	6	1.1	0″.75	0″.80/0″.70/ 0″.90	1×1	K2V	[8]
UVES	19.5	6	1.1	0″.80	1″.00	1×1	K2V	[9]
GIRAFFE	19.5	6	1.1	0″.75	1″.20	1×1	K2V	[5]
MODS	19.5	1	1.1	0″.75	0″.70	1×1	K2V	[10]
NIRSpec	21.0–26.0	6	...	...	0″.2	1×1	K5III	[11]
PFS	19.5	1	1.1	0″.75	1″.05	1×1	K2V	[12]
MSE	19.5	1	1.0	0″.75	0″.80	1×1	K2V	[13]
FOBOS	19.5	1	1.1	0″.75	0″.80	1×1	K2V	[14]
GMACS	19.5, 21.0–26.0	1, 6	1.1	0″.75	0″.70	4×4	K0V, K5V	[15]
WFOS	19.5	1	1.1	0″.75	0″.75	1×1	K0V	[14]
MOSAIC (NIR/Vis)	19.5	1	1.1	0″.75	0″.80/0″.60	1×1	K0I/V	[10]†/[14]†
G-CLEF	19.5	6	1.0	0″.79	0″.70	6×9	K2V	[16]

Table 2.4: ETC configurations used in this work. Exposure times are chosen to mimic realistic observing strategies for each instrument. Multiple apparent magnitudes, exposure times, and stellar templates are used with the fiducial 1200G grating on the Keck/DEIMOS spectrograph to investigate their effects on chemical abundance precision. Stellar templates are chosen to best match the stellar energy distribution of the relevant reference star. URLs for the referenced ETCs are provided in Table 2.5.

†S/N adapted from ETC of similar instrument according to Equation (2.14).

‡The LRIS ETC does not include the 1200/7500 grating throughput so the 1200/9000 grating throughput is used in its place.

*As an IFU, MUSE does not have a definite fiber or slit size on the sky.

#	Name	URL
[1]	DEIMOS ETC	http://etc.ucolick.org/web_s2n/deimos
[2]	LRIS ETC	http://etc.ucolick.org/web_s2n/lris
[3]	HIRES ETC	http://etc.ucolick.org/web_s2n/hires
[4]	LCO ETC	http://www.lco.cl/scripts/lcoetc/lcoetc_sspec.html
[5]	GIRAFFE ETC	https://www.eso.org/observing/etc/bin/gen/form?INS.NAME=GIRAFFE+INS.MODE=spectro
[6]	SAO ETC	http://hopper.si.edu/etc-cgi/TEST/sao-etc
[7]	MUSE ETC	http://eso.org/observing/etc/bin/gen/form?INS.NAME=MUSE+INS.MODE=swspectr
[8]	X-SHOOTER ETC	https://www.eso.org/observing/etc/bin/gen/form?INS.NAME=X-SHOOTER+INS.MODE=spectro
[9]	UVES ETC	https://www.eso.org/observing/etc/bin/gen/form?INS.NAME=UVES+INS.MODE=FLAMES
[10]	MODS Instrumental Sensitivity	http://www.astronomy.ohio-state.edu/MODS/ObsTools/Docs/MODS1_InstSens.pdf
[11]	JWST ETC	https://jwst.etc.stsci.edu/
[12]	PFS ETC and Spectrum Simulator	https://github.com/Subaru-PFS/spt_ExposureTimeCalculator
[13]	MSE ETC	http://etc-dev.cfht.hawaii.edu/mse/
[14]	FOBOS/WFOS ETC	https://github.com/Keck-FOBOS/enyo
[15]	GMACS ETC v2.0	http://instrumentation.tamu.edu/etc_gmacs/
[16]	G-CLEG ETC	http://gclef.cfa.harvard.edu/etc/

Table 2.5: URLs for ETCs referenced in Table 2.4.

2.4.2 Gradient Spectra

ab initio spectra are generated using the same method as described in Ting, Conroy, Rix, and Cargile (2017). Briefly, we first compute 1D LTE model atmospheres using the `atlas12` code maintained by R. Kurucz (Kurucz 1970; Kurucz and Avrett 1981; Kurucz 1993, 2005, 2013, 2017). We adopt solar abundances from Asplund et al. (2009) and assume the standard mixing length theory with a mixing length of 1.25 and no overshooting for convection.[14] We then evaluate spectra for these atmospheres at a nominal resolution of $R = 300,000$ using the `synthe` radiative transfer code (also maintained by R. Kurucz). The spectrum is then continuum normalized using the theoretical continuum from synthe.[15] These high-resolution, normalized spectra are then subsequently convolved down to the average resolution of the relevant instrument (assuming a uniform Gaussian LSF) and finally subsampled onto a wavelength grid with $\Delta\lambda/nR$, where n is the number of pixels per resolution element.

To calculate stellar spectral gradients for each label, we generate a grid of 200 mock spectra, each with one of 100 stellar labels offset from the reference labels of the star (see Table 2.1) by

$$\begin{aligned} \Delta T_{\rm eff} &= \pm 50 \text{ K}, \\ \Delta \log g &= \pm 0.1, \\ \Delta v_{\rm turb} &= \pm 0.1 \text{ km/s, or} \\ \Delta [X/H] &= \pm 0.05, \end{aligned}$$

where X refers to elements with atomic numbers between 3 and 99. These step sizes are chosen to be small enough such that the spectral response to each label change is approximately linear, but large enough that the spectral responses remain dominant over numerical noise (>0.1%). For each spectrum in which the abundance of an element is changed, the hydrogen mass fraction is renormalized to compensate, while the helium mass fraction remains constant.[16]

As in Ting, Conroy, Rix, and Cargile (2017), we reevaluate the atmospheric structure whenever a stellar label is varied. While more computationally expensive, this is not only essential to capture the response of the spectrum with respect to the atmospheric parameters (i.e., $T_{\rm eff}$, $\log g$, and v_{micro}), but is also important for certain elemental abundances that have substantial impact on the atmospheric structure of the star (see Ting, Conroy, and Rix 2016, for details). For example, Mg and Fe are both major electron donors in the atmospheres of cool stars and affect the absorption features of many other elements (Figure 2.1). While

[14] We note that these are not identical assumptions to those made in the MIST isochrones used in Section 2.4.1.1. This may have a small impact on the consistency of the bolometric magnitudes of the reference stars but should not otherwise affect the results presented in this paper.

[15] Again, the use of imperfectly continuum-normalized spectra here should not dramatically change the results of this work as long as all spectra are self-consistently normalized.

[16] We opt not to calculate gradients with respect to the helium fraction, but recognize that this may be of relevance to abundance measurements of hot ($T_{\rm eff} > 8500$ K) stars in globular clusters or other environments where light element variations are common (see review by Bastian et al. 2018, and references therein).

not necessary for all elemental abundances (e.g., Y, which contributes negligibly to the structure of the atmosphere), we nevertheless recompute the stellar atmosphere in all cases for consistency.

The final step is to calculate the gradients via the finite difference method. In past work, Ting, Conroy, Rix, and Cargile (2017) calculated an asymmetric approximation of the gradient of the spectrum with respect to each stellar label by considering the difference of the reference spectrum and the spectra with offsets in that label. In this work, we use a symmetric approximation of the gradient, using the two spectra offset positively and negatively from the reference spectra as we find it yields a more accurate instantaneous derivative at the location of the reference labels. Thus, the gradient of the spectrum with respect to each stellar label, α, evaluated at the reference point θ is

$$\frac{\partial f(\lambda,\theta)}{\partial \theta_\alpha} = \frac{f(\lambda,\theta+\Delta\theta_\alpha) - f(\lambda,\theta-\Delta\theta_\alpha)}{2\Delta\theta_\alpha}. \tag{2.16}$$

2.4.3 Summary of Assumptions

For reference, we provide a list of the simplifying assumptions employed throughout our methods. This does not include any assumptions inherent to the derivation of the CRLBs in Section 2.3.2.

Stellar Model Assumptions:

- `atlas12` stellar atmosphere model (1D LTE; mixing length of 1.25; no overshoot for convection).
- `synthe` radiative transfer code.
- Perfectly normalized spectra.
- MIST stellar isochrones.
- Solar abundance patterns.
- Holtzman et al. (2015) empirical relationship between surface gravity and microturbulent velocity.
- SED approximated by a K0I, K2V, or K0V spectral template.

Instrument Assumptions:

- Gaussian LSF constant with wavelength.
- Nominal wavelength sampling of 3 pixels/FWHM adopted when unknown.
- No correlations between adjacent pixels.
- Negligible read noise.
- Same instrument throughput when scaling the S/N using Equation (2.14).

2.5 Forecasted Precision of Existing Instruments

Having established how to calculate CRLBs, we are adequately positioned to forecast the chemical abundance precision of existing instruments. With an emphasis on extragalactic stellar spectroscopy, we begin with a thorough analysis of our fiducial instrument setup: the 1200G grating on Keck/DEIMOS. We then proceed to forecast the precision of other low- and moderate-resolution MOSs on large ground-based telescopes, emphasizing those with wavelength coverage bluer than 5000 Å. Finally, we investigate the capability of low-S/N, high-resolution spectroscopy for precise abundance measurements. With the exception of the analysis in Section 2.5.1.4, we assume uniform priors on all stellar labels throughout this section.

2.5.1 D1200G: A Fiducial Example

Though designed with galaxy spectra in mind, the DEIMOS spectrograph on the 10 meter Keck telescope has been critical to our understanding of the resolved stellar populations and chemical evolution of dwarf galaxies. Over the past two decades, observational campaigns with DEIMOS have measured spectra of nearly 10,000 stars in roughly 60 Local Group dwarf galaxies and the halo of M31 (e.g., Chapman et al. 2005; Martin, Ibata, Chapman, et al. 2007; Simon and Geha 2007; Kirby, Guhathakurta, Simon, et al. 2010; Collins et al. 2013; Vargas, Gilbert, et al. 2014; Martin, Geha, et al. 2016; Martin, Ibata, Collins, et al. 2016; Kirby, Xie, et al. 2018). The majority of these observations have been made with the 1200G grating centered at 7000 Å (see Table 2.2 for details). We will refer to this observational setup as D1200G throughout this work.

In the years immediately following the commissioning of DEIMOS, its primary scientific application was the measurement of radial velocities (e.g., Chapman et al. 2005; Martin, Ibata, Chapman, et al. 2007; Simon and Geha 2007). Stellar chemistry was often a secondary goal, particularly because high-resolution spectroscopy was often assumed to be necessary for any reliable abundance determinations (see Tolstoy, Hill, et al. 2009, and references therin). Kirby, Guhathakurta, Bolte, et al. (2009) demonstrated that the D1200G setup on Keck (and medium-resolution spectroscopy more generally) could be used to recover accurate abundances. Since then, D1200G has become a predominant observing mode for resolved star abundance measurements in dwarf galaxies, making it an excellent fiducial setup for our CRLB calculations.

For this exercise we consider 1, 3, and 6 hr of integration on our fiducial [Fe/H] $= -1.5$ RGB with apparent magnitudes of $m_V = 18$, 19.5, and 21.0 (or equivalently at 50, 100, and 200 kpc). The S/N in each case is calculated using the public ETC according to the configurations in Table 2.4.

The CRLBs for D1200G are displayed in Figure 2.3. Throughout this work, we report precisions for solar-scaled relative abundances with respect to hydrogen (i.e., σ[X/H]).[17] We

[17]The precision of abundances with respect to Fe (i.e., σ[X/Fe]) can be found by adding σ[X/H] and

consider $\sigma_{CRLB} = 0.3$ dex to be the worst precision that still enables useful science and thus restrict our analysis to those that can be recovered to this precision or better. We forecast that one hour on D1200G is sufficient to measure 13 elements to better than 0.3 dex in RGB stars out to 50 kpc, 10 elements out to 100 kpc, and 3 elements out to 200 kpc.

As expected from the many features seen in the gradient spectrum (Figure 2.1(b)), the Fe abundance is recovered to the highest precision. The many strong (and weak) Fe lines included in the D1200G spectrum lead to a precision of 0.02 dex at 50 kpc and to better than 0.2 dex at 200 kpc in only 1 hr of integration. Ni and Si are also precisely recovered due to their numerous features ($\sim$40 lines with gradients $>1\%\,\mathrm{dex}^{-1}$) in the red optical. The high precision possible for Ca, however, is predominantly a result of the very strong Ca II triplet[18] at $\lambda\lambda$8498, 8542, and 8662. Meanwhile, elements like Y have only a few weak lines within the D1200G wavelength range (Figure 2.1(d)) and are thus only recoverable in nearby stars.

Longer exposures provide better S/N, allowing for more precise measurements of more abundances. For a 3 hr observation, the number of elements measured to $<$0.3 dex increases to 20, 11, and 7 for RGB stars at 50, 100, and 200 kpc, respectively. For a nearby 18$^{\mathrm{th}}$ mag star, the S/N is sufficient ($\sim$150 pixel^{-1}) to measure elements with only weak signatures in the spectrum. For example, C and N can be recovered from broad, weak CN molecular features between 7000 and 9000 Å. Cu can be measured from two weak ($\sim1\%\,\mathrm{dex}^{-1}$) absorption lines at $\lambda\lambda$7935, 8095. Similarly, elements like La, Mn, O, and Eu have no more than 10 absorption lines with gradients $>0.5\%\,\mathrm{dex}^{-1}$ and only 1 or 2 lines with gradients $>1\%\,\mathrm{dex}^{-1}$. However, given the high S/N of these observations, they can nevertheless be recovered to a precision of $<$0.3 dex.

At six hours of integration, the S/N is approximately 200, 75, and 30 pixel^{-1} for RGB stars at 50, 100, and 200 kpc respectively. This enables the recovery of 22, 13, and 9 elements to better than 0.3 dex for these stars. Only after 6 hr of exposures are the weak Y lines enough to measure its abundance out to 100 kpc. These extra three hours of integration are necessary to measure Nd and V in the 18$^{\mathrm{th}}$ mag RGB from roughly a dozen very weak lines with gradients $<0.5\%\,\mathrm{dex}^{-1}$.

In Figure 2.3, we also include the spectroscopic precision on the atmospheric parameters T_{eff}, $\log g$, and v_{micro}. With the continuum shape removed from our spectrum, the effective temperature can only be constrained by its impact on atomic and molecular transitions as seen in absorption features. Compared to changes in abundance, the effect of T_{eff} on absorption lines is quite weak ($\sim$2% per 100 K for Hα and $<$1% per 100 K for most other lines), but because it manifests in thousands of lines across the D1200G wavelength coverage, it nonetheless allows for T_{eff} to be recovered to better than 100 K in most of the scenarios considered here. In contrast to T_{eff}, changes in $\log g$ affect fewer lines, but much more strongly. Hα and the Ca II triplet are notable lines sensitive to the surface gravity in the

σ[Fe/H] in quadrature.

[18] We note that the Ca II triplet is produced in the chromosphere of stars and is subject to substantial non-LTE effects, especially at low metallicities and so must be treated with caution in practice (Jorgensen et al. 1992; Mashonkina, Korn, et al. 2007; Starkenburg, Hill, Tolstoy, González Hernández, et al. 2010).

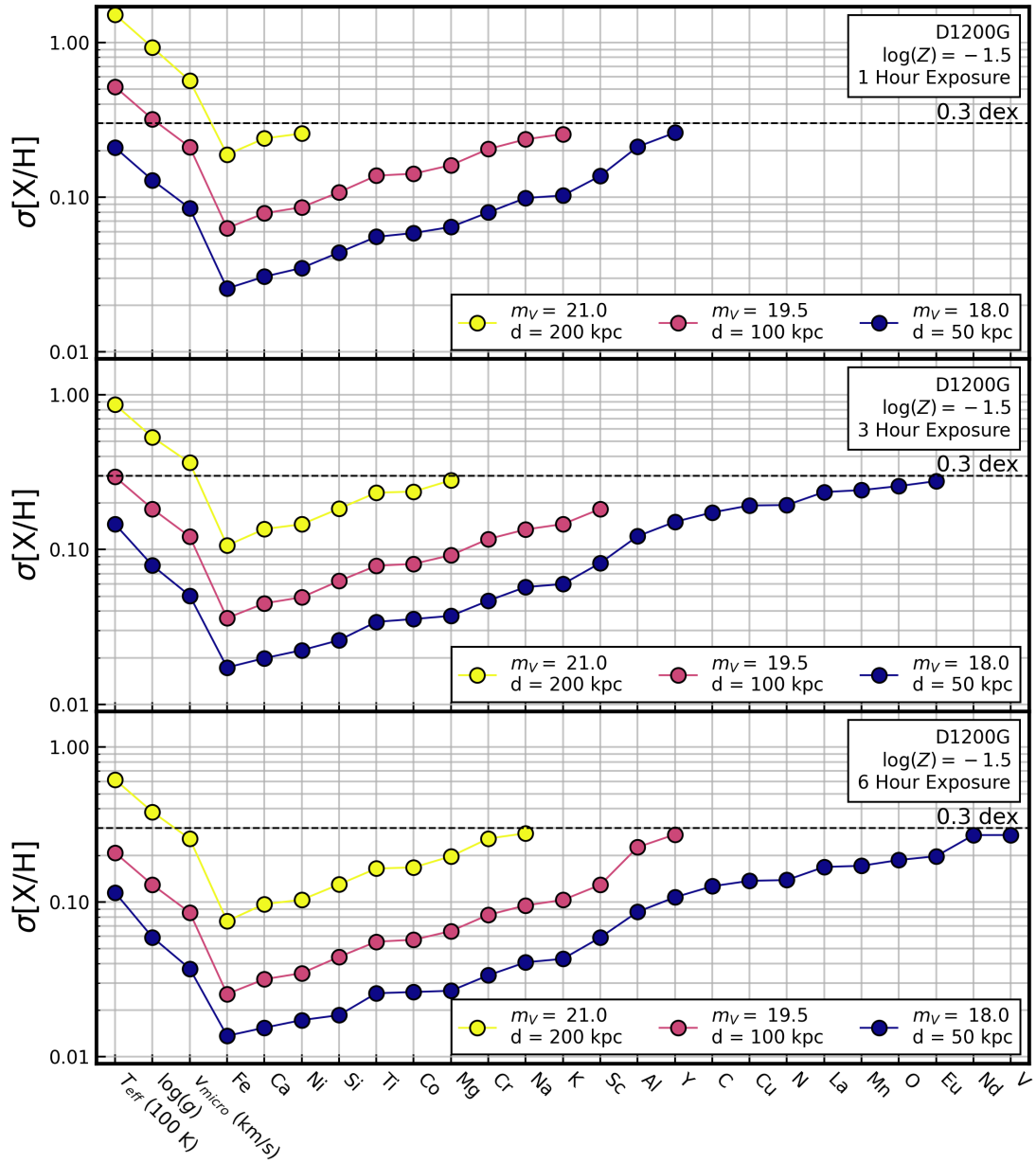


Figure 2.3: CRLBs for 1, 3, and 6 hr exposures (top, middle, and bottom respectively) of a $\log Z = -1.5$, $M_V = -0.5$ RGB star (see Table 2.1) using the 1200G grating on Keck/DEIMOS (see Table 2.2). Each panel includes the CRLBs for the RGB star located at a distance of 50, 100, and 200 kpc. The elements are ordered by decreasing precision up to 0.3 dex.

red optical. The microturbulent velocity lies somewhere between $T_{\rm eff}$ and $\log g$, moderately impacting (1%–4% per $\rm km\,s^{-1}$) $\sim$50 absorption features across the spectrum.

2.5.1.1 Comparison to Literature Precision

Our CRLBs formally represent the best achievable abundance precision via full spectral fitting, not necessarily what is obtained in practice (due to imperfect models, variable LSFs, masked or obscured features, etc.). It is therefore useful to compare our CRLB estimates to published abundance precisions from full spectral fitting to get a sense of how close current abundance measurements get to our predictions.

For an illustrative comparison, we select abundances measured by Kirby, Xie, et al. (2018), who use a full spectral fitting technique (as opposed to EWs) for RGB stars in Local Group galaxies (Kirby, Guhathakurta, Bolte, et al. 2009). Because of the large variety in stellar targets and spectral quality, we make several cuts to the Kirby, Xie, et al. (2018) sample in order to fairly compare the reported precision and our CRLBs. First, we consider only stars with $T_{\rm eff}$ between 4500 and 5000 K, $\log g$ between 1.7 and 1.9, and [Fe/H] between -2.0 and -1.0. Second, we consider only stars that were observed to 35 Å^{-1} $<$ S/N $<$ 65 Å^{-1}, which corresponds to roughly the mean S/N of a 1 hr exposure of a 19.5 mag star. These cuts leave the reported abundance precision of 33 stars.

Before we make a direct comparison, we modify our CRLB calculation to closely adhere to the choices made by Kirby, Xie, et al. (2018). For example, $\log g$ and $v_{\rm micro}$ are not fit via spectroscopy, but held fixed at values determined by the photometry of the star. This can lead to more precise recovery of abundances by removing their covariances with these labels. Similarly, only Fe, Ca, Ni, Si, Ti, Co, Mg, and Cr are fit, while all other abundances are fixed at the solar abundance value. These are not unreasonable assumptions because the information content of the spectra is dominated by these elements, and $\log g$ is typically better constrained with photometry than spectroscopy in extragalactic contexts where the distance is well constrained. We mimic this analysis by adopting a delta function prior on all stellar labels that are not fit for by Kirby, Xie, et al. (2018).

In addition, Kirby, Xie, et al. (2018) mask a handful of specific spectral regions that are contaminated by poorly modeled lines or strong telluric absorption features. Following Kirby, Guhathakurta, and Sneden (2008) we mask 13 spectral regions including notable spectral features such as the Ca II triplet ($\lambda\lambda$6498, 8542, 8662) and the Mg I λ8807 line.

It is worth noting that there are several aspects of the method used by Kirby, Xie, et al. (2018) that we cannot account for. First, they adopt a different set of stellar models and line lists than we do, albeit with similar 1D LTE assumptions (e.g., `ATLAS9` vs. `ATLAS12`; see Kirby, Guhathakurta, Simon, et al. 2010). Second, they fit stellar labels iteratively by looping through the labels and fitting each individually while holding the rest constant until convergence is achieved. It is possible that this approach may ignore some covariances between labels that are expected when all labels are fit simultaneously as assumed by the CRLB. Third, the specific wavelength coverage of each spectrum varies from the nominal depending on the location of the star on the DEIMOS detector.

Lastly, we note that the chemical abundance uncertainties reported by Kirby, Xie, et al. (2018) include both a statistical and systematic uncertainty component added in quadrature. Because CRLBs are purely a measure of statistical precision and not accuracy, we subtract out in quadrature the systematic component (of order 0.2 dex for Co and 0.1 dex for all other elements) to make a better one-to-one comparison with the literature uncertainties.

Figure 2.4 shows the reported precision of the 33 stars from Kirby, Xie, et al. (2018) plotted with our D1200G CRLBs—both with and without adjustments to match their specific analysis We find that the abundances reported by Kirby, Xie, et al. (2018) are within a factor of ∼2 of our corresponding CRLBs. The precisions reported for Fe (0.05 dex), Co (0.12 dex), and Cr (0.22 dex) are slightly less than our predicted precisions (0.06, 0.14, and 0.20 dex respectively). This may be due to a slight overestimation of the systematic uncertainty on these labels or the underestimation of label degeneracies as a result of the iterative fitting. The reported precision for Co, Mg, and Cr, are likely skewed to higher precision because only abundances recovered to better than 0.3 dex are reported, leaving only eight stars with Co abundances, one star with Mg abundances, and six stars with Cr abundances.

The biggest difference between the CRLBs calculated previously and those calculated to mimic the analysis of Kirby, Xie, et al. (2018) is in the forecasted uncertainty of Ca and Mg, which increased from 0.07 and 0.16 dex to 0.14 and 0.22 dex, respectively. This is the result of masking strong lines for these elements, which are both highly informative but challenging to model correctly. Fixing $\log g$ would have considerably improved the precision for Ca had the Ca I triplet not been masked due its strong dependence on surface gravity. Instead, it only very slightly increases the precision of Fe and Ni from 0.06 and 0.09 dex to 0.05 and 0.08 dex, respectively, but otherwise does not change the CRLB substantially. From this comparison, we can see the importance of folding in these effects to our ability to estimate the expected precision.

While the reported uncertainty for most elements is slightly higher than the CRLB, it is encouraging to see them within a factor of ∼2. There are several reasons why poorer precision in practice could be expected. Examples include poor model fidelity, imperfect calibrations, and masked or lost spectral regions (see Section 2.7.4 for further discussion). While future comparisons with abundance precisions from full-spectrum fitting are necessary to more completely understand the prospects of achieving the CRLB in practice, this comparison with D1200G illustrates that the CRLBs at least provide a realistic benchmark for spectroscopic abundance precision. In Appendix 2.10.4, we perform an analogous comparison with LAMOST and find similar agreement between our CRLBs and the literature abundance precision.

2.5.1.2 CRLBs vs. [Fe/H]

We now consider how the CRLB changes as a function of metallicity. To do this we compare the CRLBs for RGB stars with $\log Z = -0.5$, -1.0, -1.5, -2.0, and -2.5. In order to achieve similar observing conditions for each star, we make comparisons at fixed m_V instead of at fixed stellar phase (or fixed location on the RGB; see Figure 2.2). As a result of the

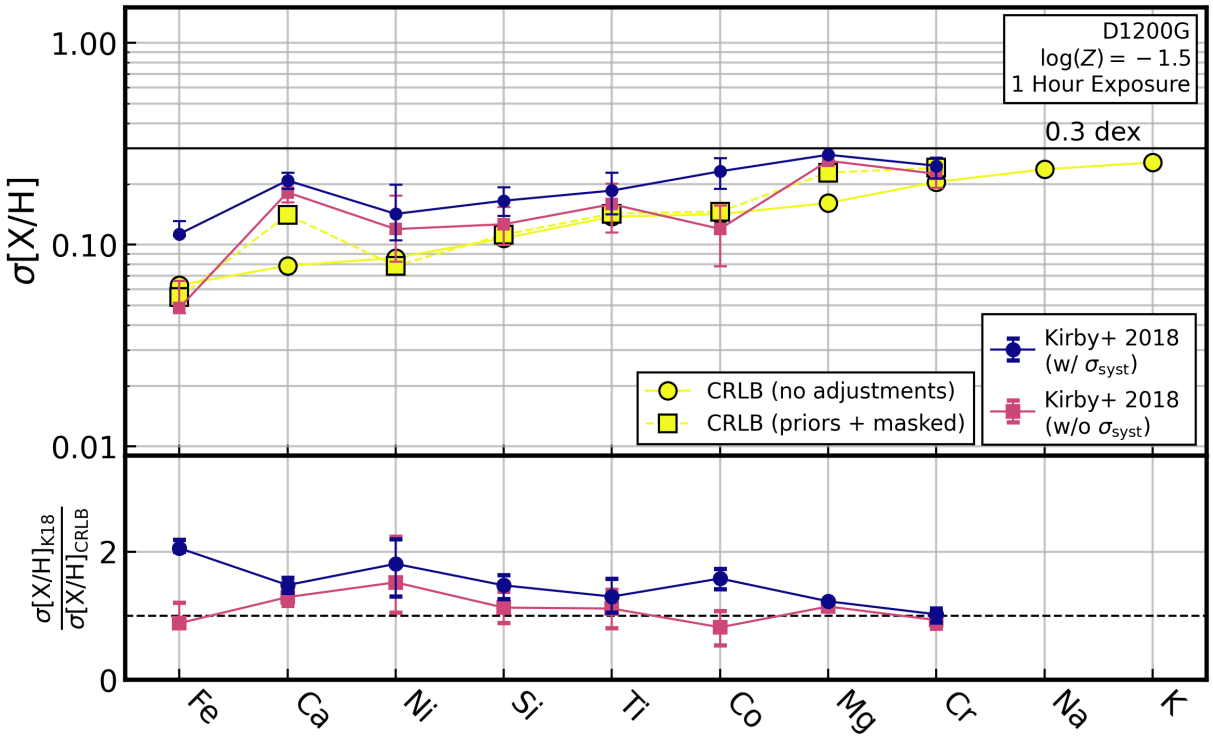


Figure 2.4: (Top) D1200G CRLBs for a 1 hour exposure of a 19.5 mag $\log Z = -1.5$ RGB star over-plotted with the uncertainties of abundances for 35 comparable RGB stars reported by Kirby, Xie, et al. (2018). The CRLBs represented by squares and dashed lines are calculated by fixing the same stellar labels and masking the same spectral features as Kirby, Xie, et al. (2018), while the CRLBs represented by circles and solid lines are the same as those presented in Figure 2.3. Literature uncertainties include a systematic uncertainty and are only provided for stars with uncertainties less than 0.3 dex. Uncertainties for atmospheric parameters $T_{\rm eff}$, $\log g$, and $v_{\rm micro}$ are not provided. Kirby, Xie, et al. (2018) did not measure [Na/Fe] or [K/Fe] abundances and therefore have no uncertainties to report for those elements. (Bottom) The ratio of the reported precision to the CRLBs that mimic the analysis techniques of Kirby, Xie, et al. (2018). Measurement precisions for most elements are within a factor of 2 larger than the CRLBs.

metallicity-dependent morphology of the RGB isochrone, $T_{\rm eff}$ and $\log g$ for these stars are all slightly different with more metal-poor stars having higher $T_{\rm eff}$ and $\log g$ (Table 2.1). The S/Ns for these stars are calculated for our fiducial observation of a 1 hr exposure of a star at 100 kpc ($m_V = 19.5$) and the configurations summarized in Table 2.4.

The CRLBs for the stars of various metallicity are plotted in Figure 2.5. As expected, the achievable abundance precision decreases toward lower metallicity as there are fewer

and weaker absorption features. However, the dependence of precision with metallicity is not uniform across all elements. For example, the precision of Fe steadily decreases from $\sim$0.03 dex to $\sim$0.1 dex as the metallicity decreases from $\log Z = -0.5$ to -2.5. The precision of V, however, decreases dramatically from $\sim$0.05 dex to $\sim$0.2 dex between $\log Z = -0.5$ to -1.0 as a result of its absorption features being strongly temperature dependent. At even lower metallicities (and slightly higher $T_{\rm eff}$), V features are nearly entirely absent.

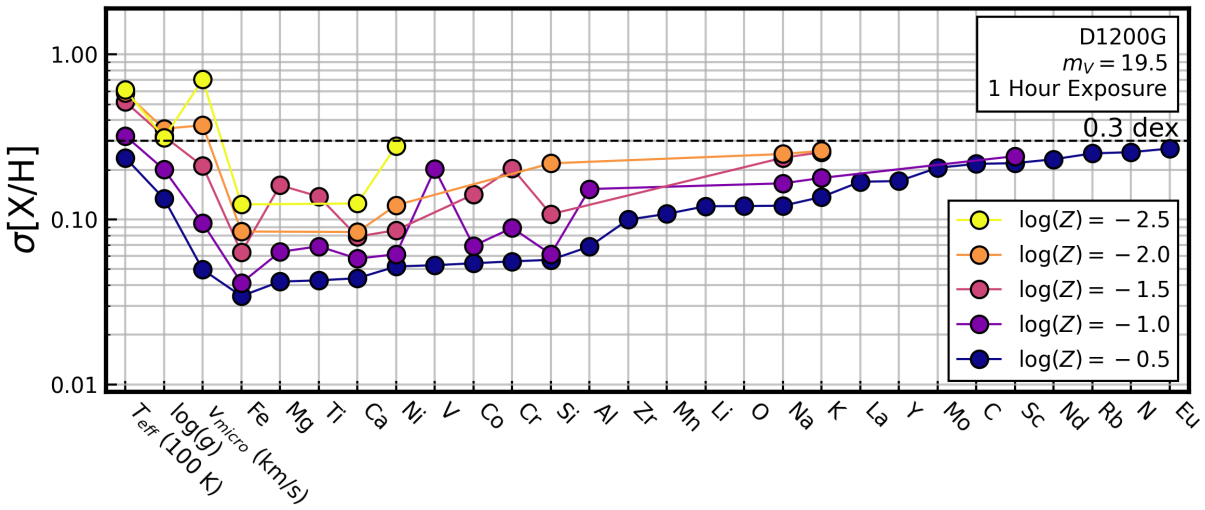


Figure 2.5: D1200G CRLBs for a 1 hour exposure of RGB stars with metallicities of $\log Z = -0.5$, -1.0, -1.5, -2.0, and -2.5 at a distance of 100 kpc ($m_V = 19.5$). Table 2.1 lists the atmosphere parameters for each star. In general, abundance recovery is less precise for lower-metallicity stars due to weaker absorption features.

Below $\log Z = -0.5$, the CRLBs for $T_{\rm eff}$ and $\log g$ remain constant, or even improve. This seemingly counterintuitive result is due to increasingly prominent Paschen lines redward of 8200 Å with increasing temperature. These lines are very sensitive to $T_{\rm eff}$ and $\log g$, allowing for precise measurements of these atmospheric parameters despite the lower metallicities.

2.5.1.3 CRLBs vs. Stellar Phase

Just as the spectral gradients of a star vary as a function of metallicity, it also varies as a function of atmospheric structure (i.e., $\log g$, $T_{\rm eff}$, and v_{micro}). As a result, we expect the achievable abundance precision at varying stellar phases to be different even at fixed metallicity and apparent magnitude. While we focus our analysis on a typical RGB star, stars from the MSTO to the tip of the red giant branch (TRGB) are also targets of extragalactic studies.

Here, we consider the CRLBs for the $\log Z = -0.5$ RGB star considered previously with that of an MSTO and TRGB star at the same metallicity (see Table 2.1). We once more

consider a 1 hr integration of a $m_V = 19.5$ star with the relevant ETC configuration in Table 2.4.

The CRLBs of each of these stellar phases are plotted in Figure 2.6, illustrating that the chemical abundance precision is best for TRGB stars and worst for MSTO stars (all other things being equal). While only 3 elements can be measured to better than 0.3 dex from the spectrum of the MSTO star, 10 elements can be measured to this precision in the RGB star, and 19 in the TRGB star. For a fixed element the precision is roughly two times better for the TRGB star than the RGB star and another two times better than the MSTO star.

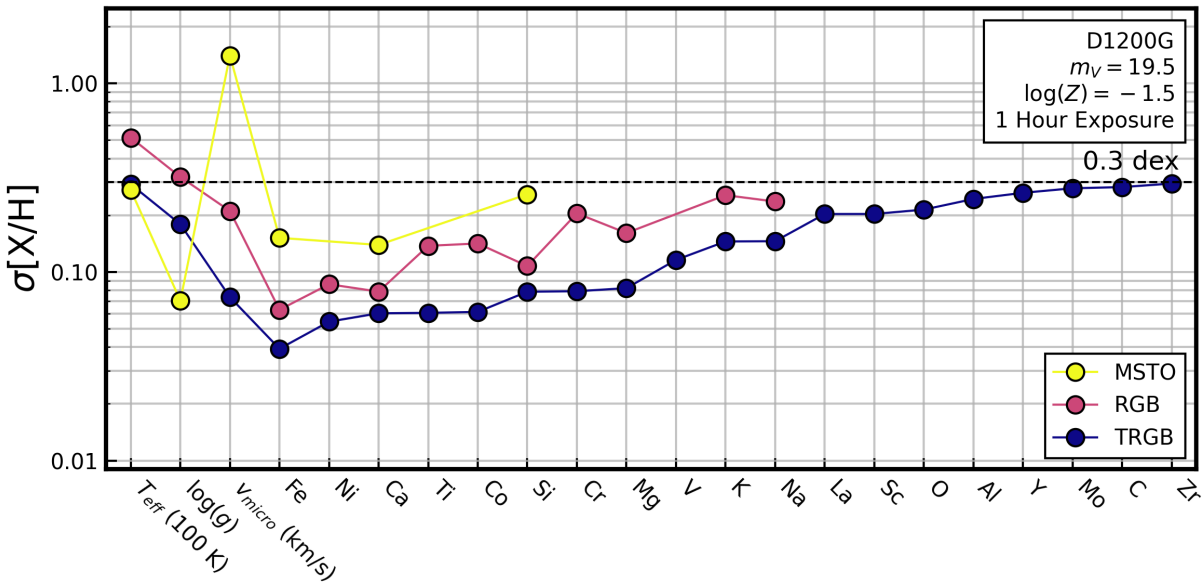


Figure 2.6: D1200G CRLBs for a 1 hour exposure of $\log Z = -1.5$, $m_V = 19.5$ MSTO, RGB, and TRGB stars. The atmosphere parameters for each star can be found in Table 2.1. At low metallicities (such as $\log Z = -1.5$), abundance recovery is more precise for cool giants due to stronger absorption features and less precise for hot subgiants, which have weaker absorption features.

These differences are expected because the absorption features of hot subgiants are significantly weaker than for cool giants. This is especially true for elements like C, N, and O, which are measured primarily from molecular features that are pronounced in TRGB stars but practically nonexistent in MSTO stars. Similarly, Fe, Si, Mg, Al, and other elements whose abundances affect atmospheric structure of a star leave a larger signature in cool, low surface gravity stars than hot, high surface gravity stars. Recovering $T_{\rm eff}$ and $\log g$, on the other hand, can be done more precisely in MSTO stars, due to the strong dependence of the Paschen lines on the atmospheric parameters of the star.

2.5.1.4 CRLBs with Priors

For stars with secure distances (as members of external galaxies typically are), photometry can be used to constrain $T_{\rm eff}$ and $\log g$ to roughly $\pm 100\,{\rm K}$ and $\pm 0.15\,{\rm dex}$, respectively (Kirby, Guhathakurta, Bolte, et al. 2009; Casagrande et al. 2011; Heiter, Jofré, et al. 2015). Knowledge of $\log g$ and Equation (2.13) can also constrain v_{micro} to roughly $\pm 0.25\,{\rm km\,s^{-1}}$ (Holtzman et al. 2015). We can incorporate these photometric estimates as priors on our spectroscopically determined labels as shown in Section 2.3.2.1. To do so we adopt Gaussian priors on these parameter with standard deviations equal to their photometric uncertainties. We once more consider a 1 hr observation of our fiducial $\log Z = -1.5$ RGB star at 50, 100, and 200 kpc.

Figure 2.7 shows the results of the CRLBs assuming Gaussian priors. For reference, we include the CRLBs from Figure 2.3 (top), which assume uniform priors. For the highest S/N case (at 50 kpc; S/N $\sim 75\,{\rm pixel^{-1}}$), the precision on $T_{\rm eff}$ and $\log g$ from D1200G spectroscopy alone is significantly better than the priors. The priors therefore contribute negligible additional information, and the CRLBs only minimally improve. However, in the lowest S/N case (at 200 kpc; S/N $\sim 10\,{\rm pixel^{-1}}$), $T_{\rm eff}$, $\log g$, and $v_{\rm micro}$ are substantially less constrained by the spectroscopy compared to the priors and so nearly all of the information about these stellar labels are coming from the prior. As a result, use of these priors improve the precision of $T_{\rm eff}$, $\log g$, and $v_{\rm micro}$ by factors of 2–6 compared to the uniform prior case.

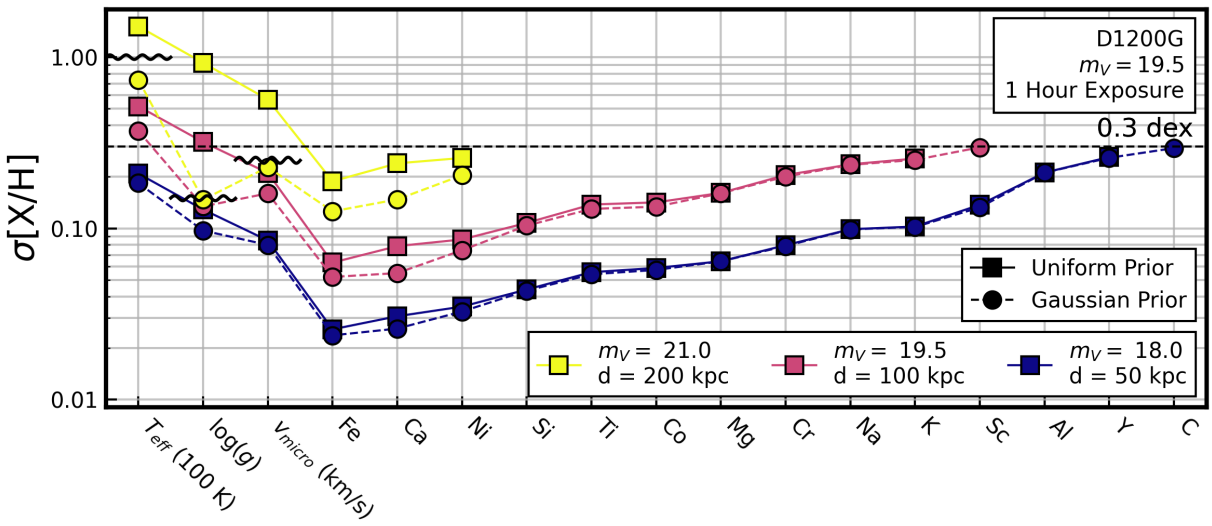


Figure 2.7: Same as the top panel of Figure 2.3 but also including the Bayesian CRLBs assuming $\sigma_{T_{\rm eff},{\rm prior}} = 100\,{\rm K}$, $\sigma_{\log g,{\rm prior}} = 0.15\,{\rm dex}$ and $\sigma_{v_{\rm micro},{\rm prior}} = 0.25\,{\rm km\,s^{-1}}$ (dashed lines). The black wavy lines mark the priors on $T_{\rm eff}$ and $\log g$. In addition to better constrained $T_{\rm eff}$ and $\log g$, the inclusion of priors also improves the precision of abundance determinations, particularly at lower S/N.

In addition, because spectral gradients of $T_{\rm eff}$ and $\log g$ are covariant with the spectral gradients of elements like Fe, Ca, and Ni, priors that better constrain $T_{\rm eff}$ and $\log g$ also lead to improved precision on these chemical abundances. For example, in the case of our faintest star, the Fe, Ca, and Ni abundance precision improves by ∼50% when Gaussian priors on $T_{\rm eff}$ and $\log g$ are included. We expect the inclusion of photometric priors to have more impact when the spectral gradients of different labels are more covariant (i.e., for low-resolution spectra with heavily blended lines and spectra with very limited wavelength coverage and few absorption lines).

2.5.2 Low and Medium Resolution MOS

All other things being equal, high-resolution spectra would be preferable for abundance measurements, as fewer lines are blended, which results in fewer coupled abundance determinations. Unfortunately, as described in Section 2.5.3, high-resolution spectrographs are typically limited to the brightest extragalactic stars due to their high spectral dispersion, relatively low throughput, and limited multiplexing capabilities. As a result, it is not possible at present to efficiently observe large numbers of extragalactic resolved stars with broad wavelength coverage and $R > 10{,}000$ spectroscopy.

Low- and medium-resolution MOSs, on the other hand, provide high multiplexing capabilities, increased throughput, and broad wavelength coverage, enabling them to achieve modest S/N of many faint stars simultaneously in distant systems. Furthermore, as we will show, wavelengths bluer than ∼5000 Å—even at low resolution—are incredibly rich in absorption features, especially for the cool low-mass giants typically observed outside the MW.

Historically, low- and moderate-resolution blue-optical spectra have not been favored for abundance determinations due to the challenge in identifying the continuum and substantial blending of lines (Ting, Conroy, Rix, and Cargile 2017). However, in recent years, advances in spectral fitting techniques have lead to large improvements in abundance recovery from low-resolution blue-optical spectra. Notably, Ting, Rix, et al. (2017) and Xiang, Ting, et al. (2019) have shown that it is possible to measure 16+ elements of ∼6 million MW stars from $R \sim 1800$ LAMOST spectroscopy with a wavelength coverage of 3700–9000 Å. While the small aperture of LAMOST (1.75 m) precludes it from abundance measurements of most stars outside the MW, there are a handful of MOS already in commission that provide similar resolving power and wavelength coverage on 6+ meter telescopes (e.g., Keck/LRIS, LBT/MODS, and MMT/Hectospec). In the following sections, we quantify the potential of these facilities for chemical abundance measurements outside the MW.

2.5.2.1 Blue-Optical MOS on Keck

On the Keck/DEIMOS spectrograph there are several options that provide access to wavelengths bluer than 5000 Å. As listed in Table 2.2, the 900ZD, 600ZD, and 1200B gratings all provide bluer wavelength coverage, but slightly lower resolution, compared to the D1200G

setup. These gratings have already enabled abundance determinations not possible from red-optical spectroscopy, such as the measurement of α-elements in the M31 halo (Escala, Kirby, et al. 2019) and Ba in several dwarf galaxies (Duggan et al. 2018). The 1200B grating is a recent addition to the DEIMOS grating collection and has not been used to measure stellar abundances at the time of writing.

In addition to DEIMOS, the Keck telescopes also host the LRIS MOS, which operates using separate red and blue channels. The 600/4000 grism on the blue arm boasts impressive blue throughput compared to DEIMOS gratings,[19] while the 1200/7500 grating on the red arm provides coverage around the Ca II triplet (Table 2.2). While LRIS has only ever been used for very limited stellar abundance determinations (Shetrone, Siegel, et al. 2009; Lai et al. 2011), it is nonetheless a promising instrument, particularly given the demonstrated success of LAMOST. To quantify the information content accessible in the blue optical by these instrumental setups, we calculate their CRLBs given a 1 hr exposure of our fiducial $\log Z = -1.5$ RGB star at 100 kpc and the relevant ETC configurations for each instrument from Table 2.4.

The forecasted abundance precision for each element is presented in Figure 2.8. Despite their lower resolving powers, instruments with bluer wavelength coverage provide more precise measurements of more elements than D1200G. For example, the 1200B grating on DEIMOS and the 600/4000+1200/7500 LRIS setup enable the recovery of 21 and 22 elements respectively to better than 0.3 dex—about twice that from comparable red-optical spectroscopy at fixed integration time and stellar type. This includes eight r- and s-process elements (Y, Ce, La, Zr, Ba, Sr, Pr, and Eu), which have most, if not all of their absorption features at wavelengths shorter than 5000 Å and are thus largely inaccessible to D1200G and other longer wavelength spectrographs. Information about C and N comes primarily from C_2, CH, and CN absorption bands between 4000 and 5000 Å and to a lesser extent from CN bands between 7000 and 9000 Å.

D1200G does provides comparable or better precision for Fe, Ni, Si, and Co, which have many lines at wavelengths longer than ∼6500 Å, as well as for Ca, Na, and K, which have strong features in the red optical.[20] The improved precision of LRIS is due to a combination of its exceptional throughput down to 3900 Å and the additional wavelength coverage provided by its red arm.[21] However, it is important to remember that LRIS has roughly half the field of view and half the multiplexing as DEIMOS (Table 2.3), meaning that it may ultimately be less efficient for some elements, when the number of stars is included in the calculation

As a reminder, the DEIMOS 600ZD and 900ZD gratings and the LRIS 1200/7500 grating all oversample their spectra with 5 pixels/FWHM. If the pixels in these spectra are not

[19] 25% at 4500 Å compared to 13% for DEIMOS 1200B and 4% for DEIMOS 1200G.

[20] The Ca II triplet at $\lambda\lambda$8498, 8542, 8662, the Na I doublet at $\lambda\lambda$8185, 8197, and the K I doublet at $\lambda\lambda$7667, 7701, respectively.

[21] Though LRIS does lose considerable information for Sc, Na, Cu, Ba, and K in the gap between its red and blue coverage. This can be mitigated to a degree by carefully choosing the dichroic and grating angle employed.

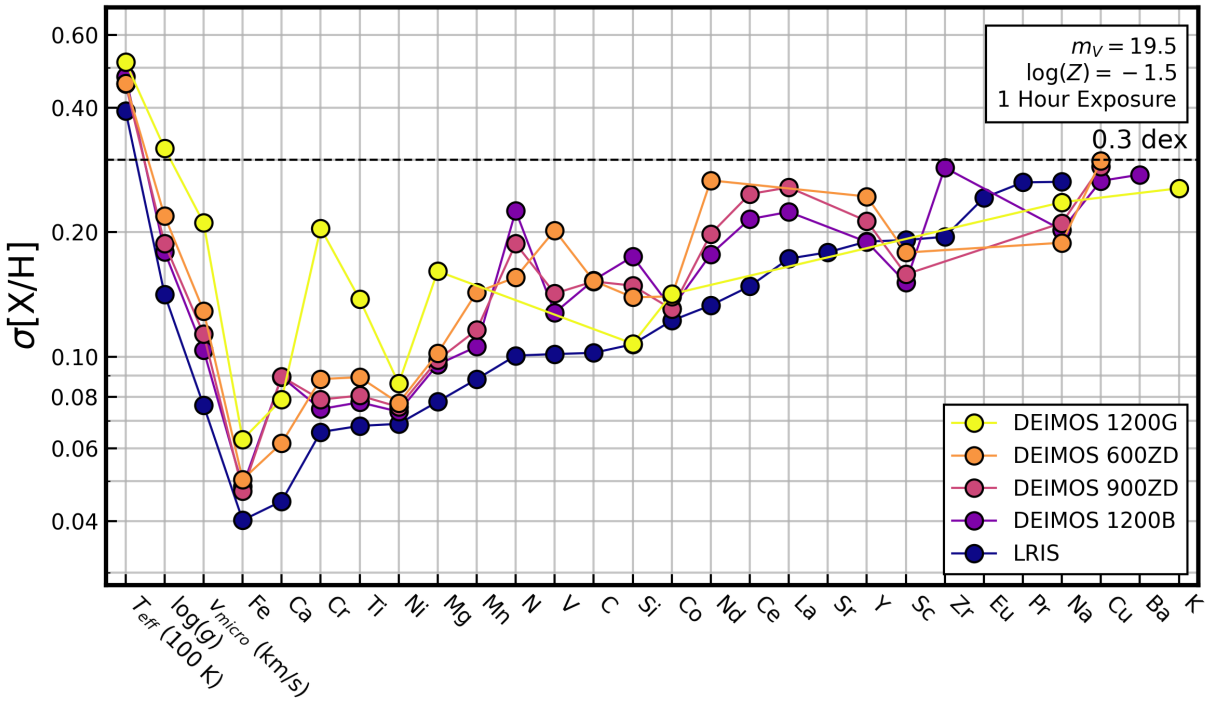


Figure 2.8: Comparison of CRLBs for several MOS setups on Keck/DEIMOS and Keck/LRIS assuming a 1 hr exposure of a $\log Z = -1.5$, $M_V = -0.5$ RGB star at 100 kpc. The LRIS setup includes the spectral coverage of both its blue and red channels. The elements are ordered by decreasing precision as forecasted for LRIS up to 0.3 dex. The CRLB for D1200G is the same as shown previously in Figures 2.3 (top), 2.5, and 2.7.

completely independent as we assume here, the CRLBs we present may be slightly more precise than would be expected in practice (see Section 2.7.4.3).

2.5.2.2 Blue-Optical MOS on other Telescopes

We now turn our attention to blue-sensitive instruments on facilities other than the Keck Telescopes, which include MODS on the LBT, MUSE on the VLT, and Hectospec and Binospec on the MMT.

MODS, like LRIS, operates at low resolution ($R \sim 2000$) across the optical spectrum with a red and a blue arm, and modest multiplexing (Tables 2.2 and 2.3). Other than a recent study on a chemically peculiar ultra metal-poor star in the dwarf galaxy Canes Venatici I (Yoon et al. 2019), MODS has not been utilized for stellar chemical abundance measurements.

While MUSE is not technically an MOS but rather an integral field unit (IFU), it can nonetheless be used effectively for low-resolution multi-object resolved star spectroscopy in

crowded fields. MUSE has already been used to conduct several campaigns for both stellar radial velocity and chemical abundance measurements in globular clusters (e.g., Husser et al. 2016; Kamann, Husser, Wendt, et al. 2016; Kamann, Husser, Dreizler, et al. 2018; Latour et al. 2019), in dwarf galaxies (e.g., Voggel et al. 2016; Alfaro-Cuello et al. 2019; Evans et al. 2019), and in NGC 300 (e.g., Roth et al. 2018; McLeod et al. 2020).

Hectospec, in comparison to MODS, MUSE, and the spectrographs on Keck, has a very large field of view ($1^\circ \times 1^\circ$), which makes it a powerful instrument for spectroscopic observations of very extended stellar populations. For example, Carlin et al. (2009) used Hectospec to measure the kinematics and bulk metallicity of stars in the disrupted MW dwarf galaxy Boötes III. Binospec is a new, complementary MOS to Hectospec with very high throughput, but a significantly smaller field of view and a more limited multiplexing capability (Table 2.3). Both Hectospec and Binospec have a number of gratings that allow for a range in wavelength coverage and resolving power. We examine a few setups we consider to be most applicable to extragalactic stellar spectroscopy (see Table 2.2 for specifics).

Figure 2.9 shows the CRLBs for our fiducial RGB star ($\log Z = -1.5$, $m_V = 19.5$) and a 1 hr exposure. For these observing conditions, MODS is forecasted to recover up to 30 individual elements to better than 0.3 dex. The precision of MODS can be attributed to two key factors: its large, nearly 12 meter effective aperture and its throughput below 4000 Å, which together achieve S/N of $>40\,\mathrm{pixel}^{-1}$ down to 4000 Å and $>10\,\mathrm{pixel}^{-1}$ down to 3500 Å. As discussed in Section 2.5.2.1, these regions become increasingly information rich due to the high densities and strengths of absorption features of many elements.

There are a few specific elements that are worth examining in more detail. Just as with the blue-optimized spectrographs on Keck, the constraints on C and N abundances come predominantly from absorption bands at wavelengths bluer than 5000 Å and (to a lesser extent) between 8000 Å and 1 μm. The sensitivity of MODS across both of these ranges leads exceptional recovery of C and N compared to the other instruments analyzed here. MUSE and the 600 gratings of Hectospec and Binospec do not push nearly as blue (or red) and thus recover C and N abundances less precisely or not at all. While the 270 grating on Hectospec and the 270 and 1000 gratings on Binospec do include most of the blue carbon features, they miss most of the blue nitrogen features and (with the exception of the 270 grating on Binospec) achieve an S/N in this region roughly half that of MODS. As a result they also do not recover C and N as precisely as MODS. In addition to C and N, MODS is also able to recover O to better than 0.2 dex because of strong OH absorption features below 3500 Å and the important role of O in the CNO molecular network (Ting, Conroy, Rix, and Asplund 2018).

Again, it is worth highlighting the precision capable of these blue-optimized spectrographs for heavy r- and s-process elements Nd, Ce, Zr, La, Sr, Y, Eu, Ba, Pr, Dy, Gd, and Sm (in order of decreasing precision for MODS). In addition to those seen in Figure 2.8, the ability to recover Nd, Dy, Gd, and Sm is the direct result of MODS blue sensitivity (discussed further in Section 2.7.1). A few of these are recoverable by MUSE, Hectospec, or Binospec, but measurement is made more difficult due to lower S/N and smaller wavelength coverage.

Given the smaller light-collecting power of MMT, it is reasonable that Hectospec and

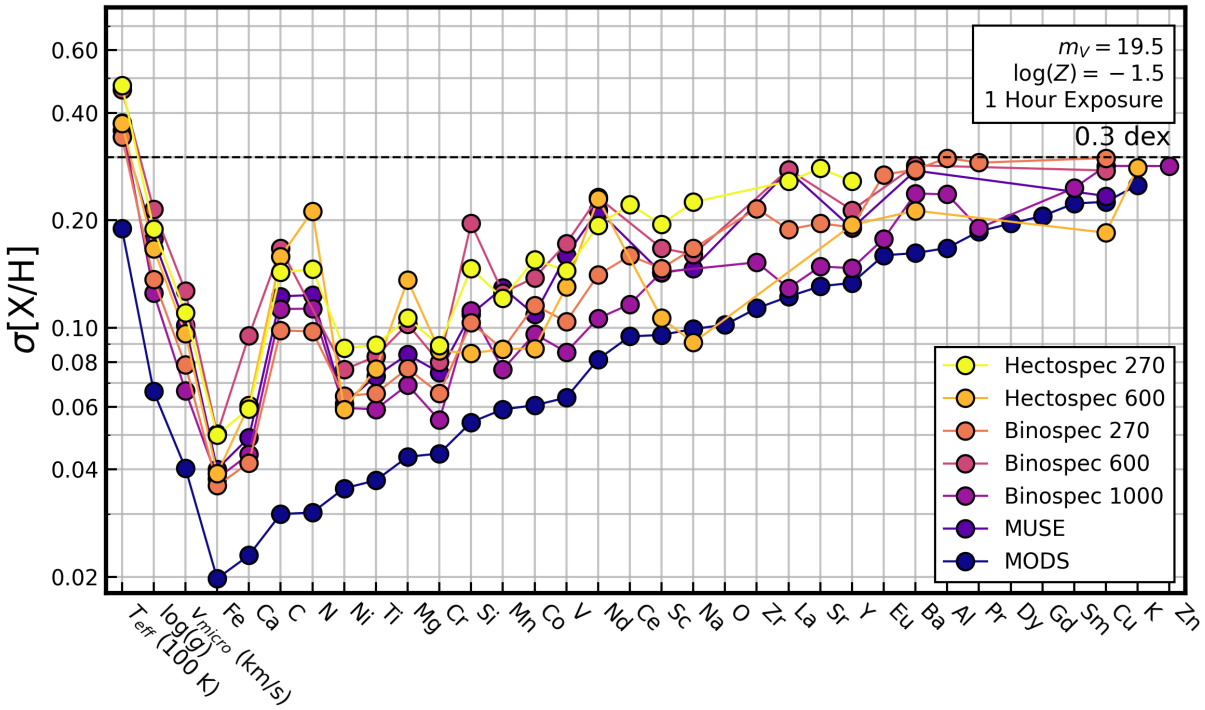


Figure 2.9: Same as Figure 2.8 but for LBT/MODS, MMT/Hectospec, and MMT/Binospec. Elements are ordered by the precision forecasted for LBT/MODS up to 0.3 dex.

Binospec are forecasted to recover fewer elemental abundances and at larger uncertainties. It is nonetheless still interesting to look at them in greater detail and compare the various Hectospec and Binospec settings. Generally the higher throughput of Binospec leads to higher precision measurements, but this of course comes with a diminished field of view and fewer fibers for stars.

Similarly, the increased abundance precision of MODS, MUSE, and other Keck spectrographs is also modulated by much reduced fields of view. The choice between these instruments then ultimately comes down to weighing the importance of detailed abundance patterns versus the importance of a large sample size to the desired science.

We remind the reader that the Hectospec configurations oversample their spectra with 5 pixels/FWHM. If the pixels in these spectra are not completely independent as we assume here, the CRLBs we present may be slightly more precise than would be expected in practice (see Section 2.7.4.3).

2.5.3 Low S/N, High Resolution Spectroscopy

In this section, we consider two classes of high-resolution spectrographs: single-slit echelle spectrographs and multiplexed single-order spectrographs.

2.5.3.1 High Resolution, Single-Slit

High-resolution spectroscopic observations of stars provide precise radial velocities and are the gold standard for chemical abundance determinations. Because high-resolution spectroscopy provides spectra with fewer blended absorption features, spectral abundance determinations preferentially use clean, isolated lines that can be fit with EW methods over blended lines, which require spectral synthesis techniques. By not fitting the entire spectrum of a star simultaneously, some of the chemical information in the spectrum goes un-utilized. By calculating the CRLBs for several high-resolution spectrographs, we illustrate the chemical information that can be accessed through full-spectrum-fitting techniques.

In the context of extragalactic studies, two commonly used single-slit echelle spectrographs are Magellan/MIKE and Keck/HIRES. Both instruments provide high-resolution spectra across the entire optical regime and have been used extensively for abundance measurements in MW globular clusters (e.g., Boesgaard, Stephens, et al. 2000; Venn, Lennon, et al. 2001; Boesgaard, King, et al. 2005; Koch and Côté 2010) and in nearby dwarf galaxies (e.g., Shetrone, Bolte, et al. 1998; Frebel, Simon, and Kirby 2014; Koch and Rich 2014; Frebel, Norris, et al. 2016; Ji, Frebel, Ezzeddine, et al. 2016; Ji, Frebel, Simon, and Chiti 2016; Ji, Frebel, Simon, and Geha 2016; Ji, Simon, et al. 2019).

We also consider two spectrographs on the VLT: UVES and X-SHOOTER. UVES is a high-resolution spectrograph with a more limited wavelength coverage (only 4800–6800 Å) but is capable of observing up to eight stars at a time when connected with the Fibre Large Array Multi Element Spectrograph (FLAMES) fiber feed.[22] It has been used to observe RGB stars in MW globular clusters (e.g., Alves-Brito et al. 2006) and in nearby dwarf galaxies (e.g., Shetrone, Venn, et al. 2003; Letarte et al. 2006; Hill et al. 2019; Lucchesi et al. 2020). X-SHOOTER has also been used to measure abundances of bright stars in dwarf galaxies (Starkenburg, Hill, Tolstoy, François, et al. 2013; Spite et al. 2018) and provides slightly lower resolution than MIKE, HIRES, and UVES but significantly higher throughput and broader wavelength coverage.[23]

As discussed in Section 2.4.1.3, a 1 hr exposure of a $m_V = 19.5$ RGB star is typically insufficient for high-resolution spectrographs to overcome the read-noise-limited regime of faint object spectroscopy. Instead we consider a more realistic 6 hr (∼1 night) of integration, which yields S/N > 15 (10) pixel^{-1} at 4500 Åand S/N > 20 (20) pixel^{-1} at 7500 Åfor HIRES (MIKE) when adopting the ETC configurations in Table 2.4.

[22] In this way, it straddles the boundary of the single-slit, multi-order spectrographs discussed in this section and the highly multiplexed, single-order spectrographs discussed in Section 2.5.3.2.

[23] The NIR arm of X-SHOOTER extends wavelength coverage to 2.48 μm, but due to the limitations of our line list we only consider wavelengths shorter than 1.8 μm.

Figure 10 shows the CRLBs for HIRES, MIKE, FLAMES-UVES, and X-SHOOTER. As expected, high-resolution spectra provide very precise detailed chemical abundance patterns. HIRES, MIKE, and X-SHOOTER are forecasted to measure a dozen elements to nearly 0.01 dex and over 30 elements to better than 0.3 dex. UVES, with its smaller wavelength coverage and lower S/N (5–10 pixel^{-1}), is still forecasted to recover over 20 elements. This high precision is predicted despite the low S/N (<20 pixel^{-1} redward of 4500 Å) of these observations, demonstrating the potential power of full spectrum fitting applied to high-resolution spectra. While at low S/N any given absorption feature might be only weakly informative, the ensemble of all spectral features still provides strong constraints on the chemical abundances of a star.

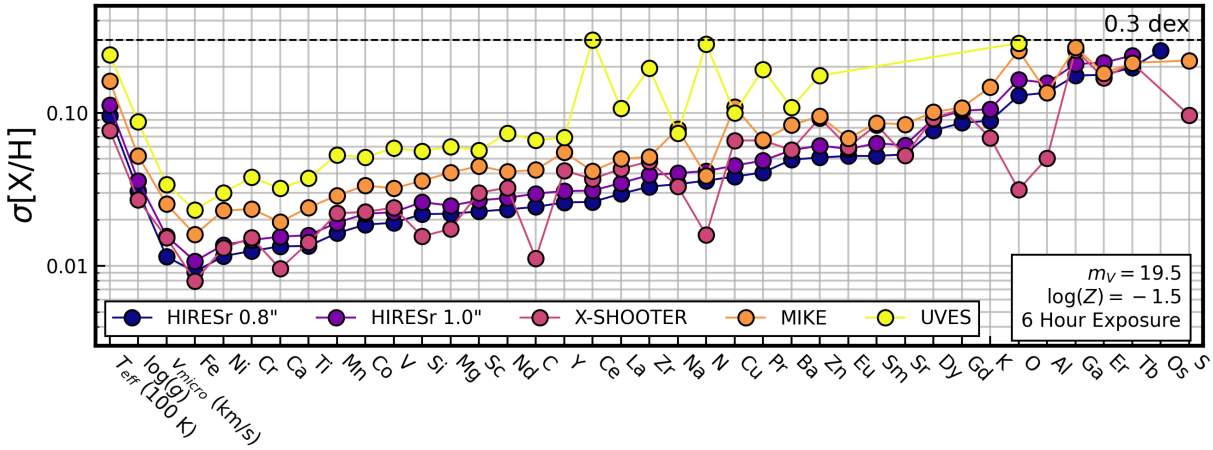


Figure 2.10: Comparison of CRLBs for high-resolution single-slit echelle spectrographs Keck/HIRES, Magellan/MIKE, and VLT/X-SHOOTER assuming a 6 hr exposure of a $\log Z = -1.5$, $M_V = -0.5$ RGB star at 100 kpc. The elements are ordered by decreasing precision as forecasted for HIRES up to 0.3 dex. The CRLBs suggest that even at low S/N (∼15–20 pixel^{-1}), the chemical information content of high-resolution spectra is considerable.

The chemical information for many of the elements in Figure 2.10 can be traced to the same large numbers of features below ∼5000 Å as previously discussed in Section 2.5.2. While these absorption features are still subject to blending, the higher resolution of these instruments increases the rms depth of the absorption feature and alleviates degeneracy between elements. This results in increased abundance precision over low-resolution instruments at fixed wavelength coverage. We can see this effect when comparing the CRLBs of the two HIRES settings, which have the same wavelength coverage but different resolving powers—the CRLBs scale with resolving power $\sigma_{CRLB} \propto R^{-1/2}$ as expected for instruments with the same wavelength range.[24]

[24] A factor of R^{-1} from the scaling of the absorption feature rms depth and a factor of $R^{1/2}$ from the scaling of S/N with dispersion. For these two HIRES settings, $R^{-1/2} \sim 0.85$.

In addition to elements previously discussed in Sections 2.5.1 and 2.5.2, HIRES can recover the abundances of neutron-capture elements Sm, Er, Tb, and Os to better than 0.3 dex. At $R \sim 50{,}000$ there are nearly 100 Sm lines with gradients $>5\%\,\mathrm{dex}^{-1}$ and over 30 lines with gradients of 10%–30% dex^{-1} in the HIRES wavelength range—all of which are below 4500 Å. The same spectrum has ∼15 (5) absorption lines with gradients of $>5\%\,\mathrm{dex}^{-1}$ ($10\%\,\mathrm{dex}^{-1}$) absorption lines for Er (Tb) blueward of 5000 Å. Os can be recovered to ∼0.3 dex from no more than five absorption lines with $>5\%\,\mathrm{dex}^{-1}$ gradients.

The bluer wavelength coverage of MIKE is largely offset by its lower resolving power ($R \sim 28{,}000$) and very low S/N ($<5\,\mathrm{pixel}^{-1}$) below 5000 Å. Nevertheless, MIKE achieves slightly better precision for Tb and Er, which have two to three times more lines between 3500 and 3900 Åthan they do at wavelengths longer than 3900 Å. For MIKE spectra, the recovery of N is aided by strong molecular absorption bands at λ3550 and λ3800 and another in the red at λ9150. Its higher precision for Al and S compared to HIRES is the result of additional atomic absorption lines beyond 8500 Åand its higher S/N in the red.

X-SHOOTER, despite its lower resolution ($R \sim 10{,}000$), recovers most elements as precisely as, if not better than, MIKE and HIRES. For C, N, and O, X-SHOOTER can achieve precisions two to three times better than MIKE and HIRES as a result of its larger wavelength coverage. It is sensitive to both the CNO molecular bands in the blue optical and the NIR molecular features beyond 1 μm. Si, Mg, Na, Al, K, and S also have a handful of absorption features in the NIR, enabling one to two times higher precision with X-Shooter than MIKE and HIRES. Furthermore, because the NIR is generally less dense with absorption features, the gradients for these elements are less degenerate with other stellar labels and can thus be more precisely recovered.

The comparatively lower precision of FLAMES-UVES can be attributed to its shorter (and redder) wavelength coverage, which does not include nearly as much of the high-information density spectral regions as the other spectrographs considered here. Furthermore, the S/N is roughly two to three times lower than that of MIKE or HIRES. Depending on the desired science, however, the multiplexing capabilities of UVES may more than make up for its lower throughput and wavelength coverage.

At low S/N (e.g., $5\,\mathrm{pixel}^{-1}$), there may be a concern that the assumptions of Gaussianity, which underlies the CRLB, may not be valid. However, we show in Appendix 2.10.5 that the CRLBs are robust to the level of ∼0.01 dex down to S/N $\sim 5\,\mathrm{pixel}^{-1}$. Thus, we believe non-Gaussianity to have a minimal impact on the CRLBs, especially compared to other practical limitations (e.g., model fidelity) that make it difficult to fully realize the precision forecasted by the CRLBs.

UVES and the UVB arm of X-SHOOTER oversample their spectra with 5 pixels/FWHM. If the pixels in these spectra are not completely independent as we assume, the CRLBs may not be as precise as we present here (see Section 2.7.4.3).

2.5.3.2 High-Resolution, Single-Order

Another approach to high-resolution spectroscopy involves using order-blocking filters that block all but one order of the echelle spectrum. Doing so allows for improved multiplexing, but limits the observed wavelength to a small window of 50–300 Å. Historically, the primary application of these instruments for extragalactic archaeology has been the efficient measurement of precise radial velocities in dwarf galaxies (e.g., Walker, Mateo, Olszewski, Gnedin, et al. 2007; Walker, Mateo, and Olszewski 2009), but these spectra clearly contain chemical information as well.

We consider three such high-resolution, single-order, fiber-fed MOS: VLT/FLAMES-GIRAFFE, MMT/Hectochelle, and Magellan/M2FS. Due to the nature of order blocking in these instruments, there is great flexibility in deciding what small portion of spectrum to observe. In this work, we will only look at spectral regions targeted by existing observations and save a detailed analysis of the optimal wavelength windows for a future paper. For M2FS, this includes a "HiRes" and a "MedRes" setting around the Mg I b triplet ($\lambda\lambda$5183, 5172, 5167), which have been used for membership determination and [Fe/H] measurement in several MW satellites (e.g., Walker, Mateo, Olszewski, Gnedin, et al. 2007; Walker, Mateo, and Olszewski 2009; Walker, Mateo, Olszewski, Bailey, et al. 2015; Walker, Mateo, Olszewski, Koposov, et al. 2016). The RV31 order-blocking filter was used on Hectochelle for similar purposes (e.g., Walker, Mateo, and Olszewski 2009; Walker, Olszewski, et al. 2015; Spencer et al. 2017) and is also utilized by the H3 MW halo survey (Conroy, Bonaca, et al. 2019; Conroy, Naidu, et al. 2019). On FLAMES-GIRAFFE, five setting have been used by the DART (Dwarf Abundances and Radial Velocities Team) program to measure various abundances and radial velocities in Local Group dwarf galaxies: LR8, HR10, HR13, HR14A, and HR15 (e.g., Hill et al. 2019; Theler et al. 2019). Details for all of these instruments and settings can be found in Table 2.2. Just as with the previous high-resolution CRLBs, we consider 6 hr of integration on our $\log Z = -1.5$ RGB star at 100 kpc and the ETC configurations in Table .

Figure 2.11 shows the forecasted precision for these single-order echelle spectrographs. As expected, the limited wavelength coverage of these setups severely reduces their chemical abundance recovery compared to the full-optical, high-resolution spectrographs presented in Figure 2.10. Even most low-resolution spectrographs can achieve comparable or better abundance recovery in a fraction of the time as presented in Figures 2.8 and 2.9. This is because the information content scales proportionally with the square root of the number of absorption features. A smaller wavelength range means fewer lines for a given element and worse precision.

Nevertheless, given the narrow wavelength range covered by these orders and the low S/N ($\sim$15–30 pixel^{-1}), it is promising that more than a handful of elements beyond Fe can be recovered to better than 0.3 dex. We first consider the abundance precision for M2FS and Hectochelle (Figure 2.11; top), which cover 5100–5300 Å. This narrow region of the spectrum contains numerous absorption lines of Fe, and to a lesser extent also of Ni, Ti, Co, Cr, and Nd, which enable their recovery. All three filters were designed to include the Mg I b triplet

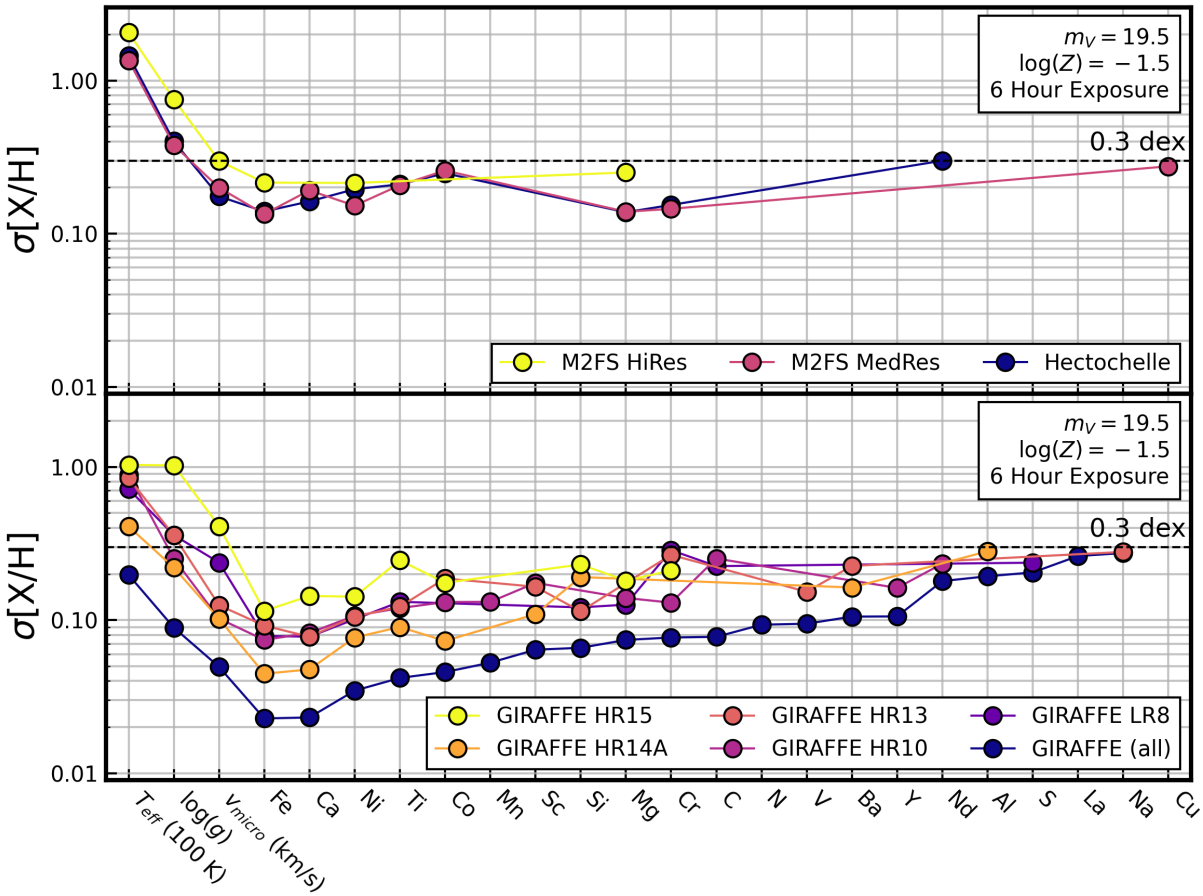


Figure 2.11: Same as Figure 2.10 but for multiplexed, single-order echelle spectrographs. CRLBs for Magellan/M2FS and MMT/Hectochelle are included in the top panel, and CRLBs for various VLT/FLAMES-GIRAFFE orders are included in the bottom panel. Elements are ordered by the precision forecasted for a combined analysis of all five GIRAFFE orders shown. The CRLBs suggest that even very small regions of spectrum, when well chosen, may contain non-negligible chemical information.

and as a result Mg can also be measured. There are also a few (<5) strong (∼10%–30% dex^{-1} at $R \sim 32{,}000$) lines each for Ca, Sc, Y, and Cu in this wavelength range that enable the M2FS MedRes configuration and Hectochelle to recover these elements. Hectochelle achieves slightly higher precision due to its higher resolving power. Because the HiRes filter on M2FS has a more limited wavelength range, it misses a considerable fraction of these lines and thus cannot measure these abundances as precisely.

Next we consider GIRAFFE, which has several orders that span the entire optical spectrum. Fe, Ca, Ni, Ti, and Co all have numerous strong lines (>10% dex^{-1} at $R \gtrsim 20{,}000$)

below 7000 Å, enabling their recovery by all high-resolution order-blocking filters. Mn, however, has the majority of its strongest lines between 5300 and 5600 Å and is thus only recovered by HR10. The same is approximately true for Y and Nd. Ba has two moderate absorption features (>10% dex^{-1} at $R \gtrsim 20{,}000$) at λ6143 and λ6499 in the HR13 and HR14A filters, respectively, but is better recovered in HR14A because of its higher S/N and resolving power. The combination of throughput and resolution enables HR14A to achieve higher precision for its recoverable elements than the other individual filters, though its redder wavelength coverage precludes it from measuring elements whose lines reside primarily at wavelengths bluer than 6000 Å.

For reference, we also include in the bottom of Figure 2.11 the CRLB for the combined analysis of all five GIRAFFE orders as was done in Hill et al. (2019). It is clear that by combining the many information-carrying absorption features across all orders provides a significant improvement in the possible stellar label precision and enables the measurement of elements that no individual filter alone could recover (e.g., N and La). However, to achieve the S/N and abundance precision found here, would require 6 hr of integration on each of the five GIRAFFE orders for a total of 30 hr of integration. Still, it is useful to compare this precision to that of low-resolution MOS and high-resolution single-slit echelle spectrographs. While low-resolution, blue-optical spectroscopy can achieve similar precision abundance determinations for a similar number of stars in a small fraction of the time, the kinematic information in these observations is limited—at $R \sim 2000$, the precision of radial velocity measurements is only $\sigma_{RV} \sim 150\,\mathrm{km\,s^{-1}}$, which is good enough for membership determination, but not for detailed kinematic studies. In contrast, $R \sim 20{,}000$ spectra yield $\sigma_{RV} \sim 5\,\mathrm{km\,s^{-1}}$, which are precise enough for stellar multiplicity determinations, orbit reconstruction, and dark matter mass measurements. Furthermore, these high-resolution observations will be less prone to systematics incurred by model imperfections in blended lines.

A drawback to high-resolution, single-slit echelle spectrographs is the amount of time required to build up large samples of stars. In 30 hr of integration time, assuming 6 hr per pointing and ignoring overheads, HIRES, MIKE, and X-SHOOTER could observe five stars, while five echelle orders (6 hr each) could be acquired by GIRAFFE for ∼100 stars. Ultimately, the choice of instrument and observing strategy is highly dependent on the science case and whether higher abundance precision or a larger sample size is most valuable and whether precise radial velocities are needed. However, in the specific case of chemodynamical studies of dwarf galaxies, where both chemical and kinematic information are desired for a large number of stars, it may be worth trading in full optical coverage for specific wavelength regions and higher multiplexing.

2.6 Forecasted Precision of Future Instruments

In this section, we forecast the precision achievable by instruments currently in their construction or design stages. Our lengthy, but incomplete, list includes JWST/NIRSpec, 30

meter class ELTs, and several planned survey facilities (e.g., MSE, FOBOS). Because many of these instruments are still undergoing conceptual and practical revisions, the specifications we adopt in this section are estimates based on the best currently available information.

2.6.1 JWST/NIRSpec

The unprecedented angular resolution of the Near-Infrared Spectrograph (NIRSpec) on JWST opens up a new domain of crowded-field extragalactic stellar spectroscopy that is currently at or beyond the limits of the most powerful ground-based telescopes (e.g., faint stars in the disk of M31 or beyond the Local Group). In this analysis, we consider four of the nine NIRSpec MOS disperser-filter combinations whose details can be found in Table 2.2. We consider 6 hr of integration and a $\log Z = -1.5$ TRGB star at a magnitude of $m_V = 21$, which is similar to observing such a star in M31 or at the edge of the Local Group.

Figure 2.12 shows the CRLBs for JWST/NIRSpec. We predict that NIRSpec can recover between 13 and 17 individual elemental abundances to better than 0.3 dex despite low resolution of these spectra ($R < 3000$) and the faintness of the target star. This is quite promising for the future of extragalactic stellar spectroscopy as the field moves toward more distant and crowded extragalactic systems. For comparison, ground-based observations are presently limited to measuring only [Fe/H], bulk α-element enhancements, and a few other elements in the M31's halo and satellites (e.g., Collins et al. 2013; Vargas, Gilbert, et al. 2014; Escala, Kirby, et al. 2019; Gilbert et al. 2019; Kirby, Gilbert, et al. 2020).

Figure 2.12 also shows that for the same filter (i.e., wavelength coverage) the slightly higher resolution of the G140H grating provides an advantage in precision over the G140M grating despite the reduced S/N ($100\,\mathrm{pixel}^{-1}$ versus $160\,\mathrm{pixel}^{-1}$ at $1.2\,\mu$m). Just as in Section 2.5.3, this is consistent with the CRLBs scaling with $R^{-1/2}$ at fixed wavelength coverage.

Further, we see that the redder F100LP filter provides better abundance precision than the blue F070LP filter. This is due to a combination of factors including larger wavelength coverage and marginally higher S/N achieved by the F100LP filter. Though it is true that blue-optical wavelengths are rich in information, the situation changes in the red, where molecular bands in the NIR are more information rich than the red optical.

In fact, the abundance precision benefits greatly from information contained at wavelengths longer than $1.4\,\mu$m provided by the F100LP filter. These redder wavelengths include numerous molecular features like the strong H_2O absorption lines that extend to $1.8\,\mu$m. Also included are bands of CN ($\lambda 1.1\,\mu$m), OH ($\lambda 1.4\,\mu$m), and CO ($\lambda 1.5\,\mu$m), features, which enable precise determinations of C, N, and O. In addition to Fe, Si, and Mg, which have absorption features somewhat uniformly distributed from 7000 Å to $1.8\,\mu$m, the F100LP filter also enables precise recovery of Mn, which has $\sim$10 lines between 1.2 and $1.4\,\mu$m with strengths greater than $1\%\,\mathrm{dex}^{-1}$ (at $R = 2700$).

The redder wavelength coverage of the F100LP filter also allow for more precise recovery of $T_{\rm eff}$ and $\log g$. This is the result of both Paschen lines at $\lambda\lambda 1.05$, 1.09, and $1.28\,\mu$m and Brackett lines redward of 1.46. These lines are all sensitive to atmospheric parameters and thus provide strong constraints on $T_{\rm eff}$ and $\log g$ (and to a lesser extent Fe, Si, Mg, and Al).

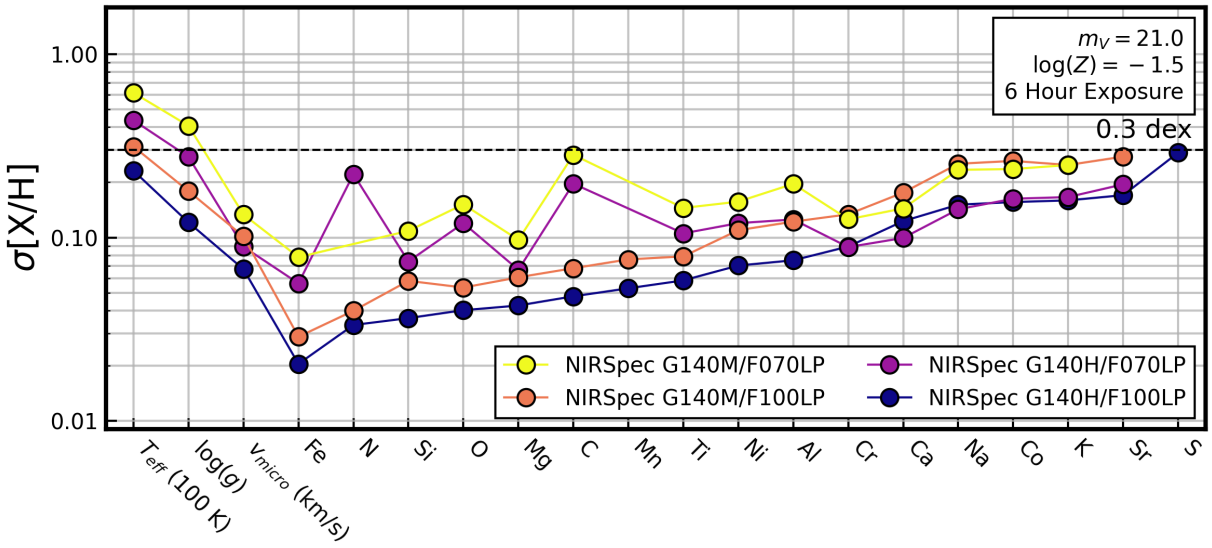


Figure 2.12: CRLBs for four gratings on JWST/NIRSpec assuming a 6 hour exposure of a $\log Z = -1.5$, $m_V = 21$ TRGB star. The elements are ordered by decreasing precision as forecasted up to 0.3 dex. These CRLBs represent the abundance precision that can be measured for RGB stars in M31 or in dwarf galaxies at the edge of the Local Group.

The bluer wavelength coverage of the F070LP filter does provide better recovery for Ti, Ca, Na, and Cr. Constraints on Ti abundance come from several TiO bands blueward of 1 μm and constraints on Cr come from roughly a dozen weak ($<2\%\,\mathrm{dex}^{-1}$ at $R = 2700$) lines blueward of 1.2 μm. The precision of Ca and Na is a result of the Ca I triplet at $\lambda\lambda$8498, 8542, 8662 and Na I doublet at $\lambda\lambda$8185, 8197 as discussed previously in Section 2.5.2.1.

We conclude by noting potential challenges in achieving the NIRSpec CRLBs. The reported elemental precision is strongly contingent on the information content of complicated molecular features. As a result, the abundances measured by NIRSpec may be quite sensitive to assumptions of the model atmosphere, molecular network, and line lists employed. Achieving the reported CRLBs and avoiding large systematics at $R < 3000$ will require careful treatment of this portion of the spectrum.

In addition, due to the rigid nature of NIRSpec's mechanical slit mask, it will frequently be the case that stars will lie slightly off the center of their slit. In addition to a small cut in S/N to lost light, this introduces deviations to the expected LSF of the spectrum. Accounting for this effect will be important for abundance recovery to approach the forecasted precision and avoid systematics caused by variations in the LSF. Efforts to calibrate NIRSpec early in the lifetime of JWST should help to mitigate this issue.

2.6.2 Extremely Large Telescopes

The advent of extremely large telescopes (ELTs) with apertures in excess of 30 meters have the potential to revolutionize extragalactic archaeology. Their higher angular resolution and increased light-collecting power will enable the spectroscopic observation of resolved stars in some of the most distant and compact systems in and around the Local Group. The Thirty Meter Telescope (TMT; 30 meter aperture), the European-Extremely Large Telescope (E-ELT; 39 meter aperture), and the Giant Magellan Telescope (GMT; 24.5 meter aperture) all have plans for a highly multiplexed spectrographs—TMT/WFOS, E-ELT/MOSAIC, GMT/GMACS, and GMT/G-CLEF.

2.6.2.1 Low-Resolution ELT MOS

We first consider the three low-resolution spectrographs, WFOS, MOSAIC, and GMACS, which all enable observations of 100+ stars across the full optical spectrum at resolving powers between $R \sim 1000$ and $R \sim 5000$. The configurations we consider are listed in Table 2.2. As in Section 2.5.2, we assume a 1 hr observation of our fiducial $\log Z = -1.5$ RGB star with $m_V = 19.5$ and the ETC configurations in Table 2.4.

Figure 2.13 presents the CRLBs for these ELT spectrographs. We predict that all three optical ELT spectrographs are capable of measuring 30–40 elemental abundances to better than 0.3 dex. In addition to all Fe-peak elements and most α-elements, this includes 22 neutron-capture elements spanning all three r- and s-process peaks. Of these, 12, 9, and 8 can be recovered to better than 0.1 dex by GMACS (G3), MOSAIC (HMM-VIS), and WFOS (B2479/R1392), respectively.

Many of these elements have only weak features below 4000 Å, which necessitate high S/N in the blue optical and near-UV for their recovery. Tb and Tm, for example, have ∼20 absorption lines with 1%–3% dex^{-1} gradients at $R \sim 3500$, but nearly all are found at wavelengths shorter than 4000 Å. Similarly, Pd, Os, and Hf have fewer than 10 absorption lines of similar strengths, which are also predominantly located blueward of 4000 Å. The strongest line of Th is at λ4019 with a gradient of $\sim$1.5% dex^{-1}, while ∼20 weaker (0.5%–1.0% dex^{-1}) features exist between 3100 and 4000 Å. Despite the limited chemical information, spectrographs on ELTs are capable of measuring these elements because their large-aperture telescopes and blue wavelength coverage can achieve S/N ∼ 100 at 4000 Å.

The informative power of blue-optical spectroscopy can be further seen in the comparatively poorer abundance recovery of the MOSAIC HMM-Vis and HMM-NIR settings. Because the optical arm only extends to 4500 Å, it cannot capitalize on the information-rich, near-UV stellar spectrum. The NIR is expected to recover even fewer abundances than the optical arm due to the lower information density beyond 8000 Å. Nevertheless there are some elements (e.g., Ca, Si, Sr, O, Al, and S) whose absorption features are better observed in the NIR. CN absorption in the red and NIR also allow for recovery of C and N to a similar degree to what can be done with spectra down to 4500 Å. We note, however, that because the JWST NIRSpec ETC was repurposed to provide S/N in the NIR for MOSAIC, the S/N

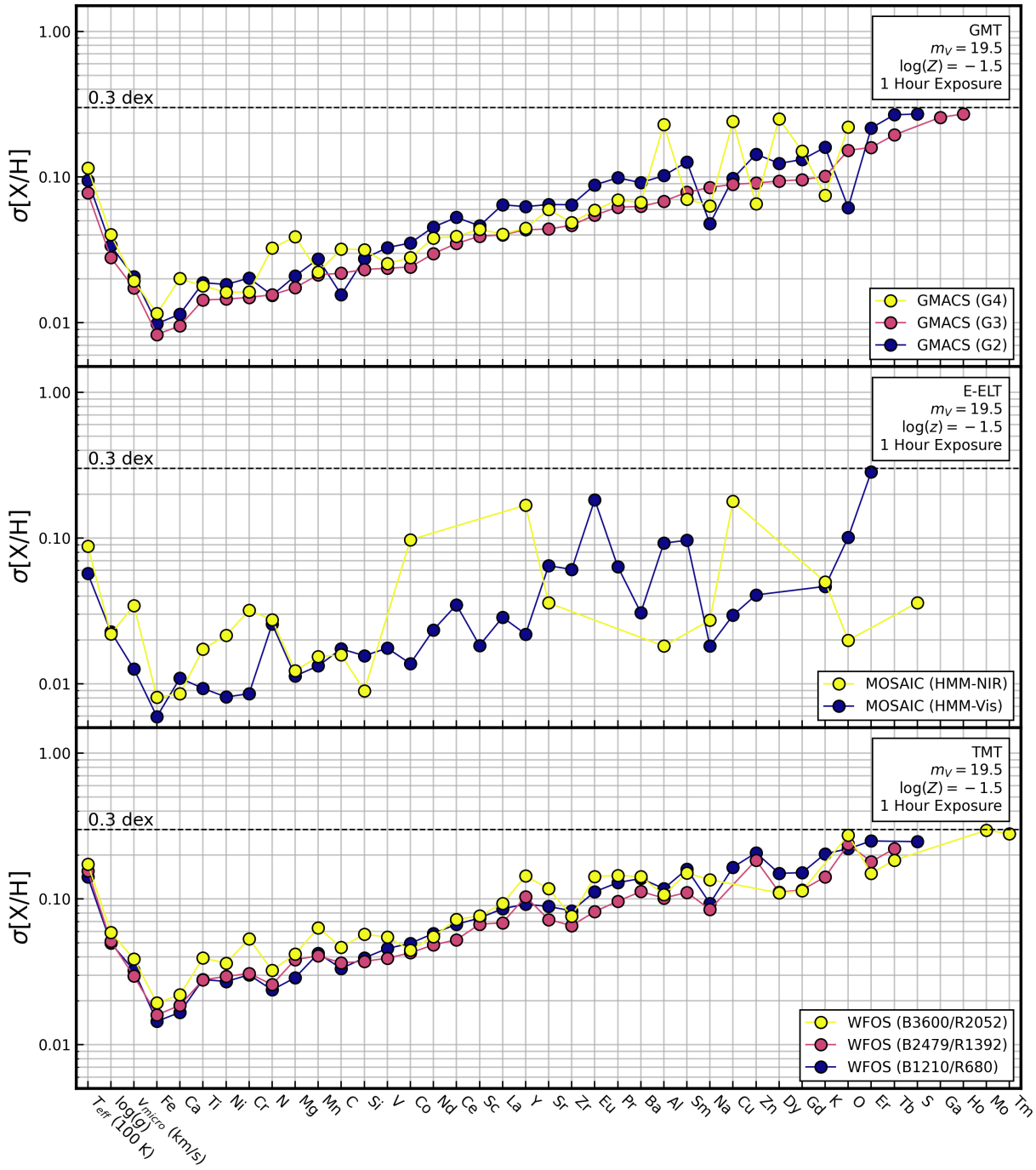


Figure 2.13: Same as Figure 2.8 but for the low-resolution ELT spectrographs GMT/G-MACS, E-ELT/MOSAIC, and TMT/WFOS.

used here does not include the effects of troublesome NIR telluric features. As a result, we expect the abundance precision of MOSAIC HMM-NIR spectra to be noticeably worse in practice.

Figure 2.13 (top) illustrates the trade-offs in S/N, wavelength coverage, and resolution at a fixed number of detector pixels for three different GMACS gratings. As predicted by Ting, Conroy, Rix, and Cargile (2017), the abundance precision of a detector with fixed pixel real estate under the assumption of uniform distribution of chemical information is relatively invariant of the resolving power. Of course, there are slight differences in the expected precision of the gratings. For many elements, G2 ($R = 1000$) performs more poorly than the higher resolution gratings, which is likely due to strongly blended lines at $R = 1000$ and the resulting increased covariance between elements. It is also apparent that the chemical information is not uniformly distributed; there are several abundances (e.g., Cr, C, Ba, Al, Dy, Gd, and K) which the G4 grating recovers noticeably worse if not at all because the absorption features of these elements lie outside of its reduced wavelength coverage. These elements are predominantly those with few strong features that lie below 4200 Å. Similar conclusions can be drawn from a comparison of the three WFOS grating combinations.

2.6.2.2 High-Resolution ELT MOS

Here, we consider G-CLEF, a GMT first-light, fiber-fed echelle spectrograph. While it is primarily optimized for very high-resolution ($R \sim$ 100,000) single-slit spectroscopy across the optical, it will also feature a MOS mode that will combine modest multiplexing, Keck/HIRES-like spectra, and a 24.5 meter aperture telescope that will dramatically increase the feasibility of high-resolution spectroscopy of stars beyond the immediate vicinity of the Local Group (see Tables 2.2 and 2.3 for details). We calculate the S/N using the G-CLEF ETC given the same observational conditions used for the forecasting of existing high-resolution instruments (see Table 2.4).

Figure 2.14 shows the CRLBs of G-CLEF with the HIRES 1″.0 CRLBs for comparison. We forecast that G-CLEF observations will recover 30 elements to better than 0.1 dex (and nearly 40 to 0.3 dex) similar to HIRES and the other single-slit high-resolution spectrographs analyzed previously in Section 2.5.3.1. In addition to achieving HIRES-like abundance recovery, the multiplexing of G-CLEF enables the simultaneous observation of up to 40 stars at a time. This dramatically increasing the feasibility of high-resolution studies of substantial numbers of stars in extragalactic systems (for both chemistry and kinematics).

The reason G-CLEF does not achieve substantially better abundance precision than its 10 meter class analogs appears to be largely a consequence of G-CLEFs lower predicted throughput. Despite having a much larger light-collecting power, G-CLEF acquires roughly the same S/N as Keck/HIRES at wavelengths shorter than 6000 Åwhere most of the chemical information resides. G-CLEF achieves higher S/N ($\sim$35 pixel^{-1} compared to $\sim$20 pixel^{-1}) at longer wavelengths, but this only yields small improvements in abundance precision. Furthermore, the bluer wavelength coverage of G-CLEF has S/N < 5 pixel^{-1} and thus provides little additional information.

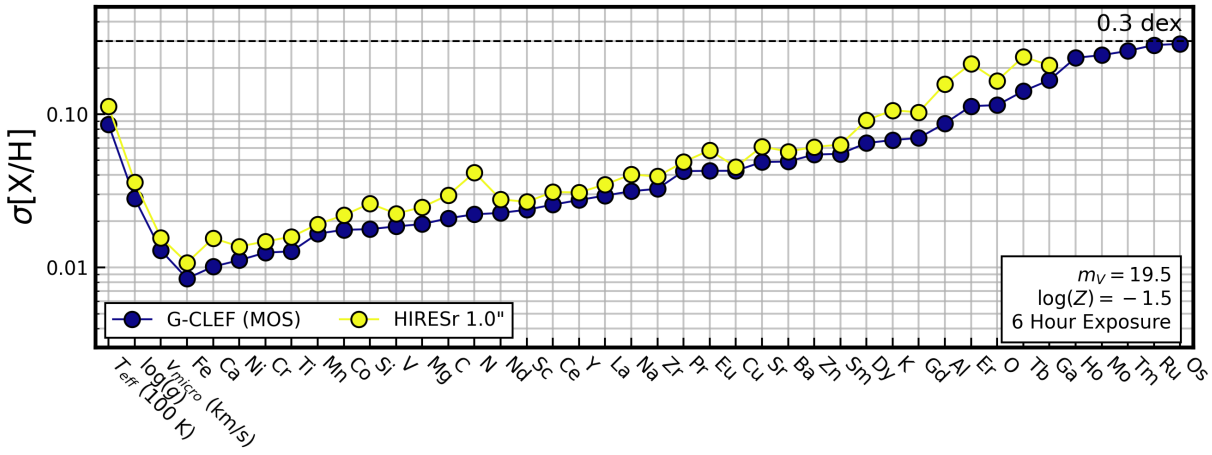


Figure 2.14: Same as Figure 2.10 but for high-resolution ELT spectrograph GMT/G-CLEF.

2.6.3 Spectroscopic Surveys

Galactic archaeology in the MW has been revolutionized by several large-scale spectroscopic surveys, including RAVE, SEGUE, LAMOST, GALAH, APOGEE, and DESI,[25]. These surveys have collected millions of stellar spectra from which detailed abundance patterns have been measured. The success of these surveys in the realm of stellar abundance measurements is in part due to the high quality and homogeneity of the spectra collected. This has allowed for rigorous, self-consistent analyses, the implementation of data-driven approaches, and the refining of stellar models. However, similarly ambitious observing campaigns outside the MW are in their early stages, primarily because it requires a dedicated survey instrument on a 10 meter class telescope.

The next decades is poised to bring the field of extragalactic stellar spectroscopy its first large sample of homogeneously collected stellar spectra. For example, the Prime Focus Spectrograph (PFS) on Subaru will begin science observations in early 2020. PFS will dedicate ∼100 nights to surveying M31's disk and halo, making it the largest extragalactic stellar spectroscopic survey to date (Tamura et al. 2018).

The MSE will replace the Canada–France–Hawaii Telescope with an 11.25 meter dedicated survey telescope, while FOBOS is a next-generation instrument proposed for the Keck telescopes with time dedicated for a stellar (extra)galactic archaeology survey. Both MSE and FOBOS are much earlier in their conceptual design and plan to be on sky by ∼2030 (Bundy et al. 2019; MSE Science Team et al. 2019).

The details for these spectrographs can be found in Tables 2.2 and 2.3. For all three survey

[25]In Appendices 2.10.4 and 2.10.6, we forecast the precision of the ongoing LAMOST MW survey and the recently begun DESI survey of MW halo stars. For forecasted precision of other MW surveys we refer the reader to Ting, Conroy, Rix, and Cargile (2017).

instruments we consider our standard 1 hr of integration time of our fiducial $\log Z = -1.5$, $m_V = 19.5$ RGB star and the ETC configurations in Table 2.4.

We present the abundance precisions of PFS, MSE, and FOBOS for this observing scenario in Figure 2.15. All three spectrographs are capable of similar chemical abundance precision to blue-optimized spectrographs considered in Sections 2.5.2.1 and 2.5.2.2 (e.g., DEIMOS 1200B, LRIS, and MODS), recovering >20 elements to better than 0.3 dex. As seen in previous analyses, there are only minor differences between the low- and medium-resolution setting on PFS and MSE. The increase in resolution is roughly canceled out by decreases in S/N and wavelength coverage. In this comparison, the additional wavelength coverage beyond 1 μm by the NIR and red arms of PFS and MSE (low-res) provide improved precision of Si and Al, but not C, N, and O, which would require even redder spectra that extend past 1.4 μm.

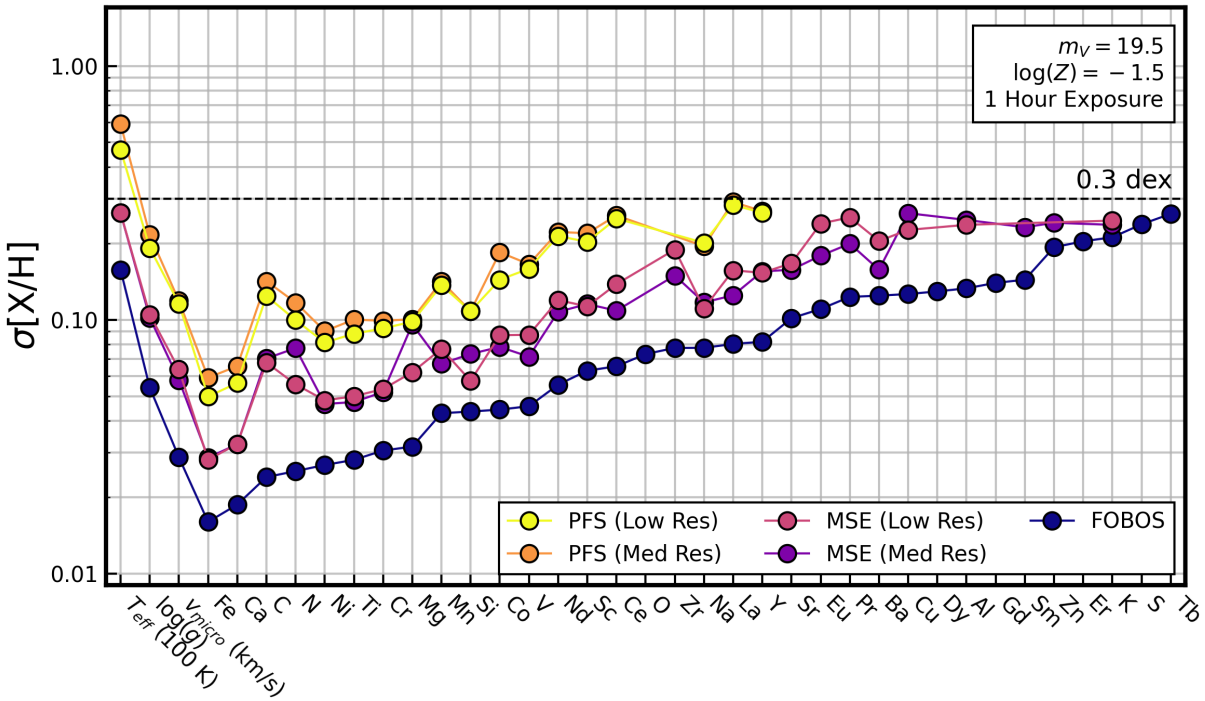


Figure 2.15: Same as Figure 2.8 but for the survey instruments PFS, MSE, and FOBOS.

Despite the relatively similar specifications of these three survey spectrographs, there is a considerable spread in their forecasted abundance precision. This can be attributed to two predominant factors. The first and most important factor is the S/N of the observations. Throughout most of the optical, PFS achieves an S/N only one-half to three-quarters that of FOBOS and MSE. In addition, the blue sensitivity of FOBOS enables an S/N $> 10\,\mathrm{pixel}^{-1}$ down to 3500 Åfor these observations, while the S/N of MSE and PFS drops below an S/N

of $10\,\mathrm{pixel}^{-1}$ at $\sim$4000 Å.

The second factor contribution to the higher precision predicted for FOBOS is its higher wavelength sampling (6 pixels/FWHM), which is nearly twice that of MSE and PFS. Even holding all other instrument specifications constant (e.g., wavelength coverage, resolving power, S/N), the higher sampling alone leads to a 2 improvement in the forecasted precision. Of course, oversampling the spectrum by this degree in practice would likely lead to increased correlations between adjacent pixels, resulting in a smaller improvement than our naïve scaling with $n^{-1/2}$ predicts (see Appendix 2.10.3).

2.7 Discussion

2.7.1 Information Rich Blue Spectra

In the context of extragalactic spectroscopy (i.e., at medium and low resolution), a key result of this paper is the importance of the blue-optical spectrum for measuring abundances. Spectral regions bluer than $\sim$4500 Å are rich in absorption features of α elements and r- and s-process elements and overall enable the recovery of more than double the number of elements than red-optical-only wavelengths. This finding echoes the power of low-resolution, blue-optical spectra highlighted in Ting, Conroy, Rix, and Cargile (2017) and demonstrated by Xiang, Ting, et al. (2019) with LAMOST spectra.

Figures 2.16 and 2.17 summarize the power of blue-optical spectroscopy for abundance recovery. To generate these figures, we have simulated a spectra with $R \sim 2000$ and 5000, respectively, and a spectral sampling of 3 pixels per resolution element for a $\log Z = -1.5$ RGB star. We then computed the CRLB for each element for the 2000 Å wavelength regions shown on the x-axis. We assume a K2V SED, constant throughput with wavelength, and an S/N of $100\,\mathrm{pixel}^{-1}$ at 6000 Å ($\sim 40\,\mathrm{pixel}^{-1}$ at 3000 Å; $\sim 55\,\mathrm{pixel}^{-1}$ at 1.5 μm). Each cell is color-coded by the CRLB precision.[26]

Figures 2.16 and 2.17 show that the largest number of elements can be recovered in the spectrum spanning 3000–5000 Å. In this range, 38 (49) elements are recovered to a precision of <0.3 dex for $R = 2000$ (5000). The number of elements available drops to 28 (34) in the 2000 Å range between 4000 and 6000 Å, indicating the rich information available below 4000 Å.

In the 5000–7000 Å range, 18 (22) elements can be recovered. As the wavelength coverage shifts redder, fewer elements are precisely measurable. $AtR = 2000$, no elements, including Fe, can be measured from 2000 Åregions between 1.2 and 1.5 μm. This is because there are few absorption features for any elements—Fe with only $\sim$20 lines with gradients larger than $1\%\,\mathrm{dex}^{-1}$ has the strongest of any element in this portion of the spectrum. The paucity of lines means there is little information to break the degeneracy between the poorly constrained $T_{\rm eff}$ and $\log g$ ($\sigma_{T_{\rm eff}} > 300$ K and $\sigma_{\log g} > 1.5$ dex) and the elemental abundances. Applying

[26] To first order, the precision of a given element from a combination of two or more wavelength windows can be found by taking the inverse square sum of the precision in the relevant wavelength ranges.

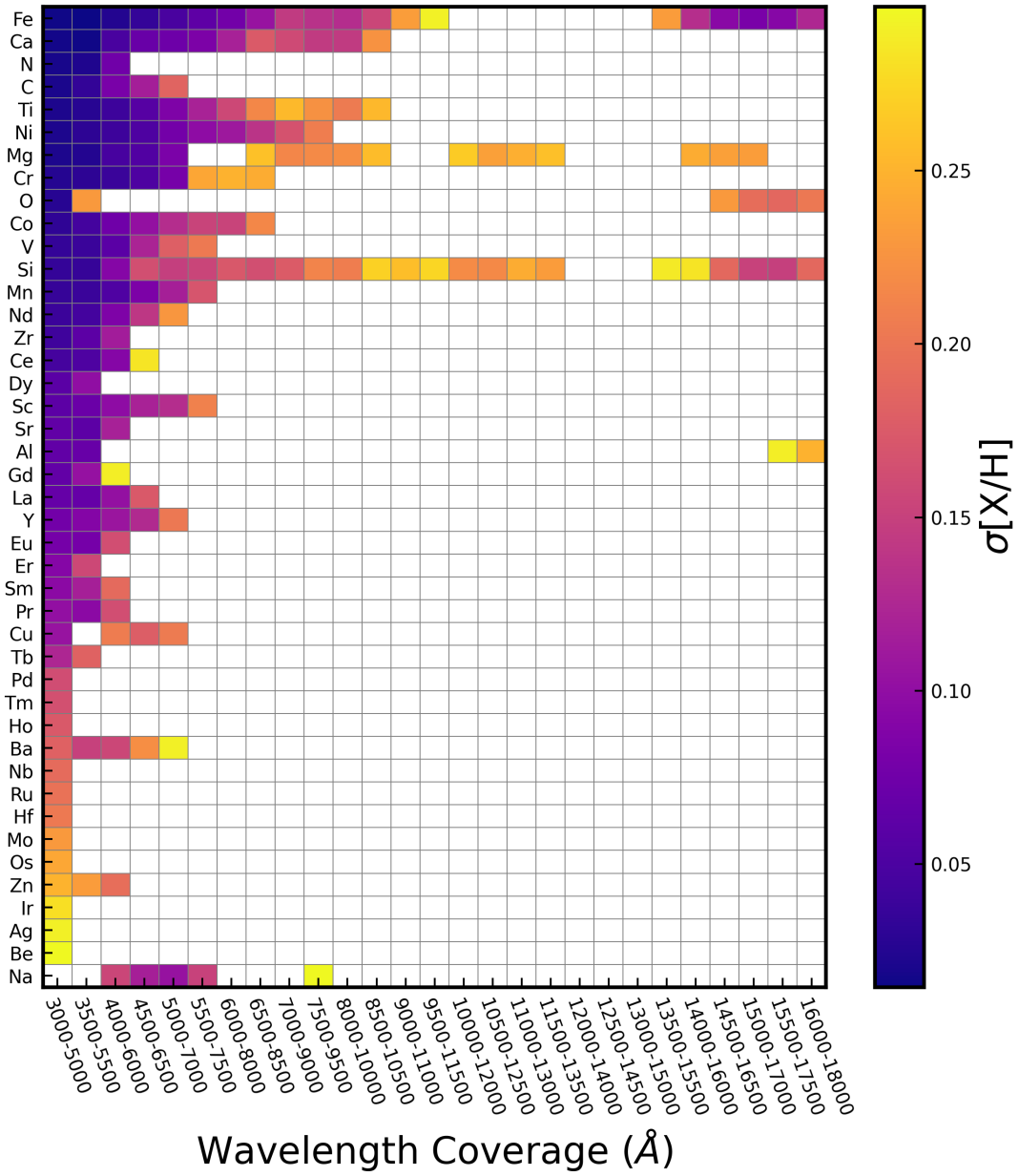


Figure 2.16: CRLBs for a $\log Z = -1.5$ RGB star observed in 2000 Å wavelength regions from 3000 Å to 1.8 μm, assuming $R = 2000$, $n = 3$ pixels/FWHM, constant throughput, a K2V stellar SED, and S/N= 100 pixel^{-1} at 6000 Å. This figure demonstrates the high density of chemical information found at wavelengths shorter than 4500 Å, especially for many neutron-capture elements.

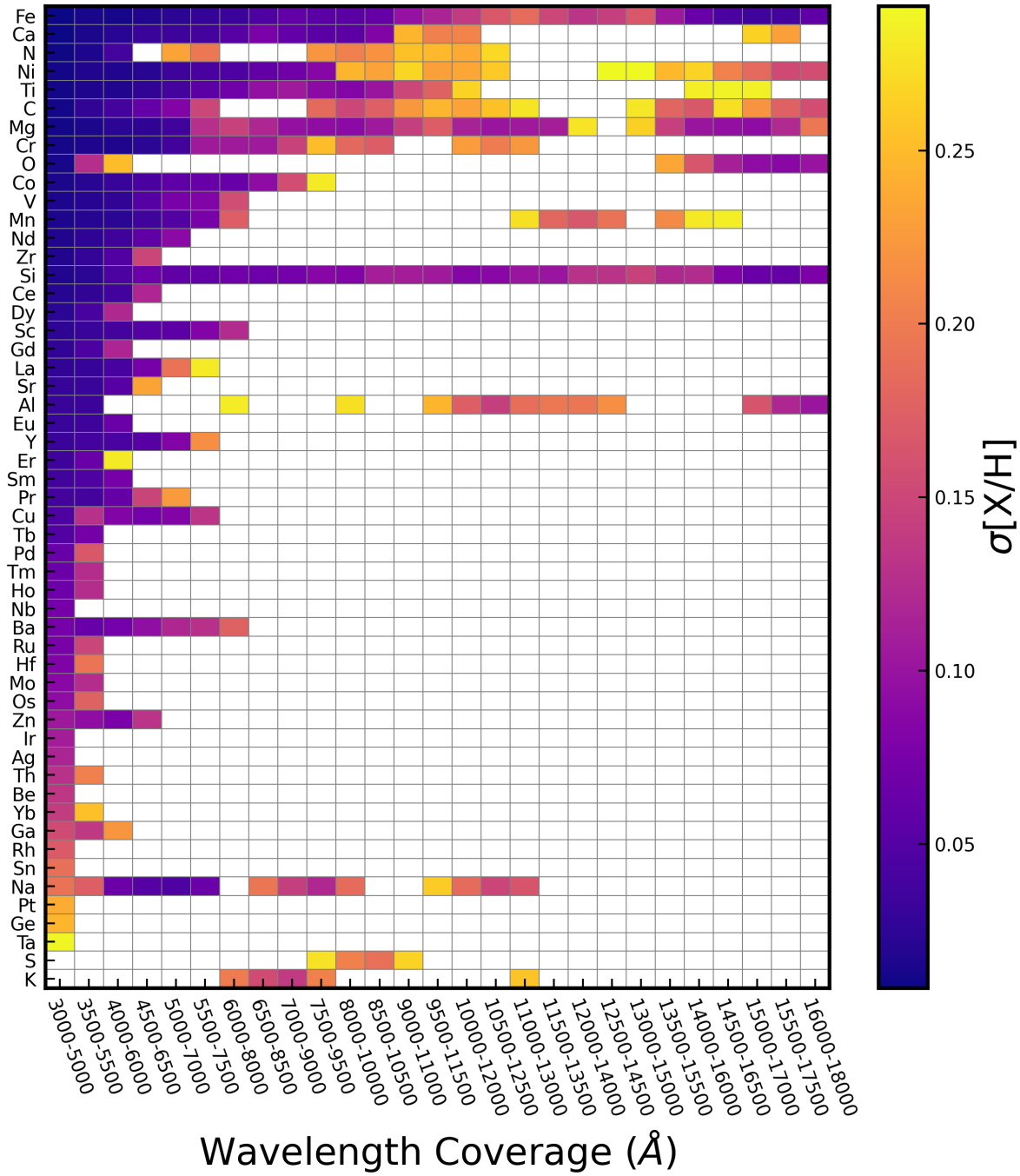


Figure 2.17: Same as Figure 2.16, except for $R = 5000$.

the same priors as in Section 2.5.1.4 enables the recovery of Fe, Si, and Mn to better than 0.3 dex. As the wavelength coverage moves farther into the near-IR (1.5–1.8 μm), the number of elements that can be recovered increases as a result of molecular features (e.g., H2O and CO) and larger numbers of Fe, Si, Mg, and Al lines (see APOGEE results: Ness, Hogg, et al. 2015; García Pérez et al. 2016; Ting, Conroy, Rix, and Cargile 2019).

Beyond increasing the number of elements that can be recovered, the blue optical is rich in the absorption lines of neutron-capture elements. For this reason, the blue-optical portion of the spectrum has long been targeted by high-resolution spectroscopy (e.g., Sneden and Parthasarathy 1983; Cowan et al. 2002; Sneden, Cowan, et al. 2003; Hansen et al. 2015). However, as shown in Figure 2.18, these elements have strong gradients even at low resolution ($R \sim 2000$). Sr and Eu, for example, have a handful of absorption lines between 3500 and 4500 Å with gradients of 4%–8% dex^{-1}. Other elements, like Zr, Ce, and Nd, have a forest of weaker ($\sim$2% dex^{-1}) absorption lines that extend blueward of 4500 Å. The results of Figures 2.16–2.18 together indicate that full spectral fitting methods have the potential to recover neutron-capture elements outside the immediate vicinity of the MW.

The high information density of the blue optical also introduces challenges to abundance recovery. For example, the large number of lines makes it challenging to define a continuum. Most spectral fitting routines operate on normalized spectra and the lack of a clearly defined continuum introduces additional sources of uncertainties into the fitting process. A second challenge is the blending of absorption lines. The blending of spectral features is not inherently a problem for full spectral fitting, provided that all stellar labels are fit simultaneously to account for degeneracies. However, doing so requires a high degree of trust in the stellar atmosphere models, radiative transfer treatment, and line lists. When lines are resolved, individual lines that are imperfectly modeled (e.g., from non-LTE or 3D effects) can be isolated and ignored. But when lines are severely blended as they are in the blue optical, identifying and masking (or calibrating) problematic lines become a nontrivial, but crucial, endeavor.

Finally, blue-optical spectra will typically have lower S/N than redder observations of the cool RGB stars we are considering—their flux peaks at $\sim$6100 Å. To achieve the same S/N at 3000 Å as at 6100 Å requires at least 50% longer integration times in the blue.[27] We have attempted to take this into account by using ETCs with SEDs of cool stars to determine realistic S/Ns of our the observing scenarios.

Taken together, the challenges of dealing with line blending and lower S/N has meant that medium- and low-resolution blue-optical spectroscopy has seldom been used for extragalactic stellar chemical abundance measurements. These difficulties, however, do not invalidate the enormity of the information content contained in the near-UV and blue portions of the stellar spectrum. Given the current designs of upcoming instruments and surveys, we will soon be awash in low-resolution blue stellar spectroscopy and the potential for major advances in abundance determinations. Fully taking advantage of this data set will not be trivial and will take significant investments in stellar models, instrumental calibrations, and spectral

[27] Assuming a constant throughput and a K2V stellar SED.

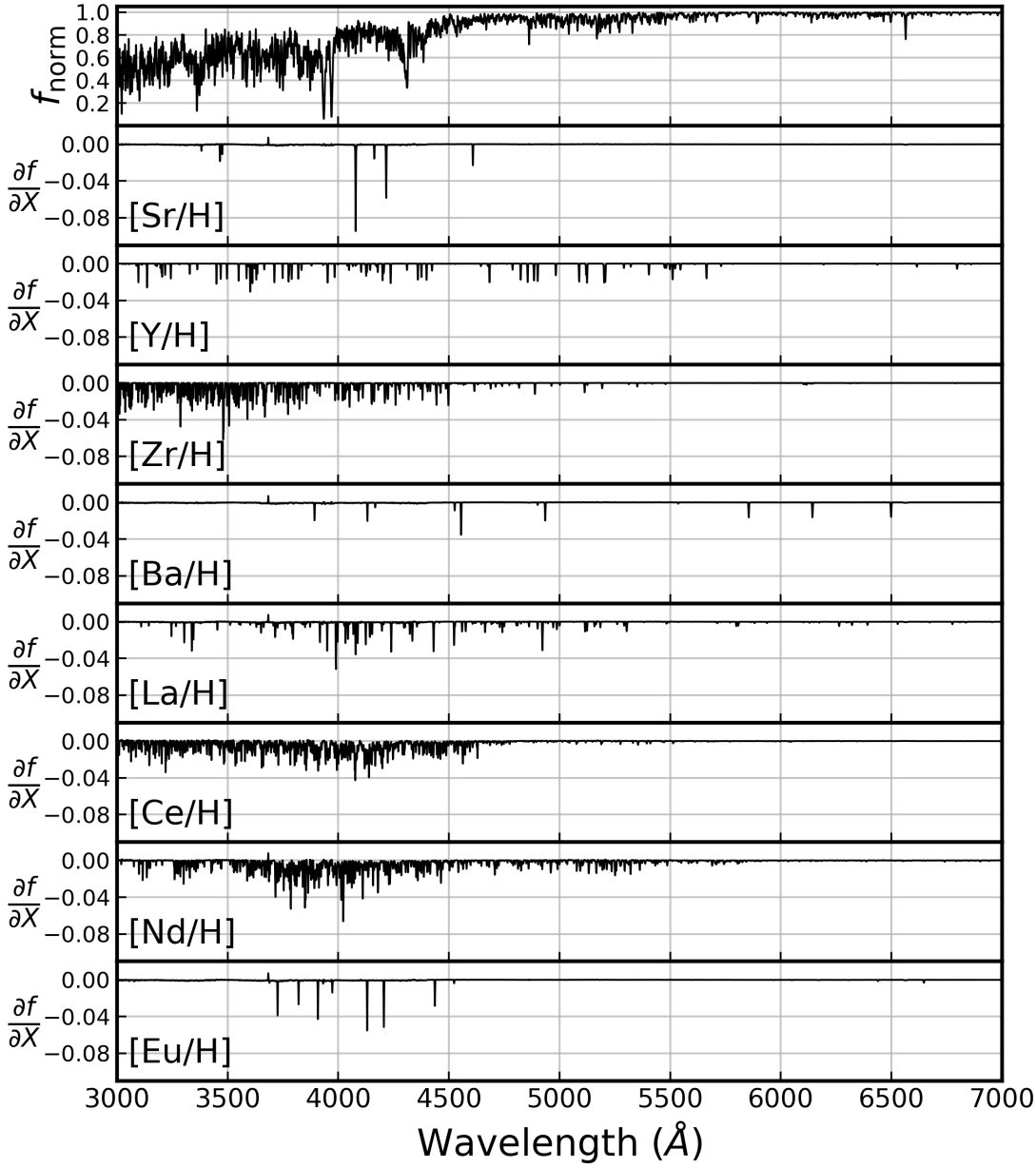


Figure 2.18: (Top) Spectrum of a $\log Z = -1.5$ RGB star convolved down to $R = 2000$. (Below) Gradients of the spectrum with respect to r-/s-process elements recoverable by LBT/MODS given the setup in Section 2.5.2.2. Most of the information for these elements is at wavelengths shorter than 4500 Å. Not shown in this figure are three modest Sr lines with gradients of $1\%\,\mathrm{dex}^{-1}$ between 1.0 and 1.1 μm and a handful of weak Y lines (all with gradients of $<0.5\%\,\mathrm{dex}^{-1}$) that lie red-ward of 7000 Å.

fitting techniques, but we believe that it will be well worth the investment.

2.7.2 Stellar Chemistry Beyond 1 Mpc

At present, a full night (∼6 hr) of observing time on a 10 meter telescope is necessary to measure [Fe/H], [α/Fe], and a few individual elemental abundances in stars as faint as $m_V \sim 23$ (e.g., Vargas, Gilbert, et al. 2014; Escala, Kirby, et al. 2019; Gilbert et al. 2019; Kirby, Gilbert, et al. 2020). While this enables the measurement of stellar metallicities in the halo of M31 with current facilities, measuring elemental abundances in systems at greater distances and stellar densities is currently out of reach, due to long integration times, read-noise limitations, and crowding. Outside the Local Group, stellar spectroscopy is not possible for resolved stars. However, both JWST/NIRSpec and the ELT spectrographs will excel in the observation of faint stars in crowded systems. They provide Hubble-like angular resolution ($\lesssim$0″.2) for spectroscopy, can achieve reasonable S/N for faint stars in modest integration times, and are sensitive to the spectral features of many elements (see Sections 2.6.1 and 2.6.2.1).

Figure 2.19 illustrates the potential of JWST and the ELTs for resolved star spectroscopy in and beyond the Local Group. Here we plot the CRLB for several elements as a function of distance for two telescope configurations: JWST/NIRSpec (G140H/100LP) and GMT/GMACS (G3; see Table 2.2). For these calculations, we assume 6 hr observations of a $\log Z = -1.5$ TRGB star (see Table 2.1) and replace the CRLBs of individual α elements (O, Ne, Mg, Si, S, Ar, Ca, and Ti) with a CRLB for [α/H].[28] The CRLBs indicate that JWST and GMACS will be able to measure the Fe abundance to 0.3 dex in individual stars out to 4.4 and 5.0 Mpc, respectively.[29] GMACS is capable of recovering α abundances, primarily through Ca features and to a lesser extent from Ti, Mg, and Si features, out to 4.5 Mpc. For NIRSpec, α is recovered through a combination of Si, O, and Mg features (in order of decreasing importance) out to 3.5 Mpc. The small wiggles in the G3 S/N at 5000 Å (and CRLBs) seen beyond 25 Mpc are the result of interpolation errors in the extraction of data from the GMACS ETC at low S/N.

We also calculate the Bayesian CRLB using the same Gaussian priors as in Section 2.5.1.4 ($\sigma_{T_{\rm eff},\rm prior} = 100\,\rm K$, $\sigma_{\log g,\rm prior} = 0.15\,\rm dex$, and $\sigma_{v_{\rm micro},\rm prior} = 0.25\,\rm km\,s^{-1}$). The middle panel of Figure 2.19 illustrates that these priors can improve the precision of C and α (N, Fe, and α) by up to a factor of 2 (1.5) for JWST (GMACS) observations of faint stars.

In addition to Fe and α, NIRSpec and GMACS are capable of recovering a handful of other individual abundances at a distance of ∼3 Mpc—N, C, and Mn for NIRSpec and C, Ni, Cr, Co, N, and V for GMACS. These elements can all be measured to better than 0.2 dex at 2 Mpc and 0.1 dex at 1 Mpc. Other elements not shown that can also be recovered to 0.3 dex out to 1 Mpc include Mn, Nd, Sc, Ce, La, Zr, Y, Pr, Sm, Ba, Na, K, Al, Sr, Eu, Cu, Gd,

[28] The gradients for α were calculated as in Section 2.4.2 except that offsets were applied to all α-element abundances in lockstep instead of individually.

[29] We note that the S/N for both instruments is quite low beyond 4 Mpc: $<10\,\rm pixel^{-1}$ for NIRSpec and <5 (<10) $\rm pixel^{-1}$ for GMACS at 5000 (8000) Å.

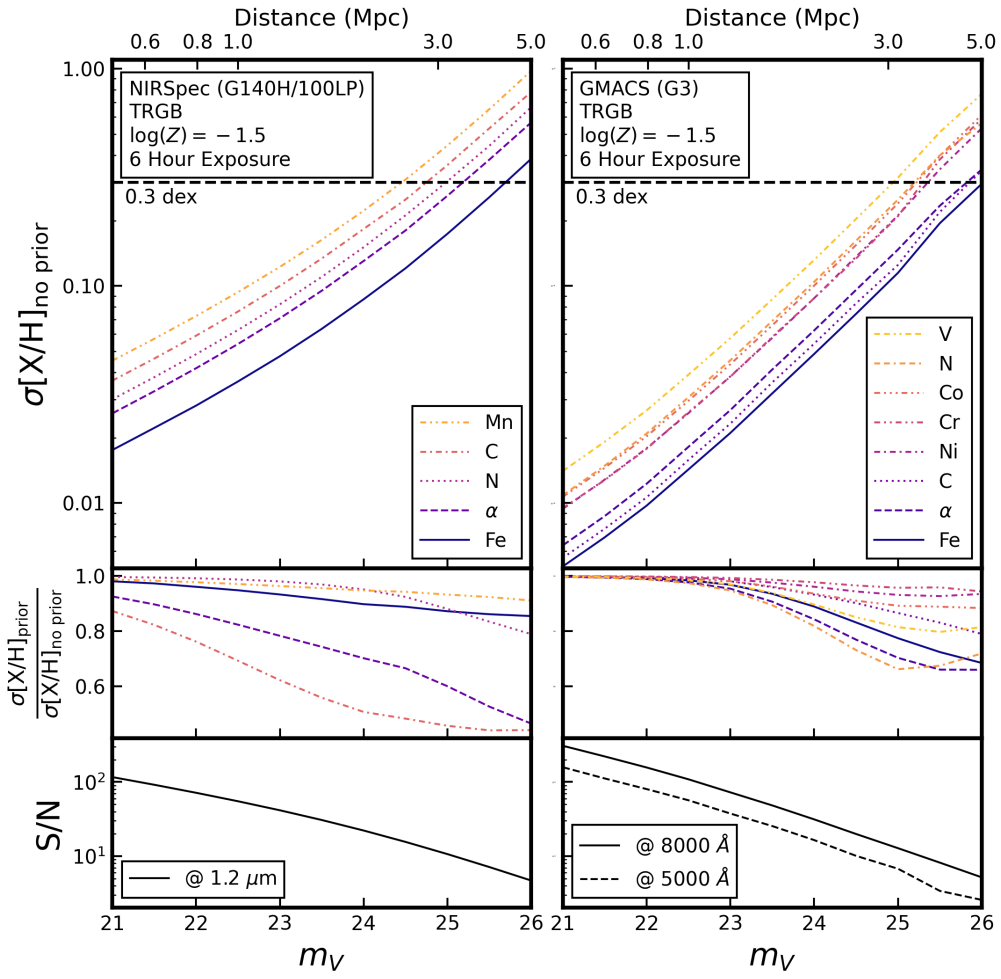


Figure 2.19: CRLBs for the JWST/NIRSpec G140H/100LP (left) and the GMT/GMACS G3 (right) setups given a 6 hr observation of a log $Z = -1.5$ TRGB star as a function of apparent magnitude and distance. The middle panels show how the CRLBs improve when assuming Gaussian priors of $\sigma_{T_{\rm eff},\rm prior} = 100\,\rm K$, $\sigma_{\log g,\rm prior} = 0.15\,\rm dex$, and $\sigma_{v_{\rm micro},\rm prior} = 0.25\,\rm km\,s^{-1}$. The S/N at a characteristic wavelength is plotted in the bottom panels for each instrument. Small wiggles in the G3 S/N at 5000 Å (and CRLBs) are due to interpolation errors in the extraction of data from the GMACS ETC at low S/N. JWST and ELTs will enable the recovery of Fe and α to better than 0.3 dex beyond 4 Mpc and out to $\sim$3 Mpc for a handful of other elements.

Zn, and Dy for GMACS and Ni, Al, and Cr for JWST. This would not only enable precise chemical abundance measurements of stars in M31 and its satellites, but also enable detailed chemical enrichment studies of galaxies at the periphery of the Local Group and beyond, including potential new faint galaxy discoveries by LSST. Though we did not explicitly compute the CRLBs as a function of distance for TMT/WFOS and E-ELT/MOSAIC, we expect that each of these powerful facilities have similar abundance recovery potentials for stars outside the Local Group.

2.7.3 Planning Observations

For stellar abundance work, selecting the appropriate spectrograph, setup, and exposure time for a specific science case can be daunting given the large number of facilities and instrumental configurations. This can often lead to inefficiencies in observational strategies.

As illustrated in Sections 2.5 and 2.6, the CRLB provides a useful and quantitative way to evaluate abundance recovery for a given spectroscopic setup. As an example, consider the comparison of Keck spectrographs and gratings in Figure 2.8, which displays the numerous trade-offs of each setup on an element-by-element basis. LRIS generally provides the most chemically informative spectra, but if high multiplexing is a priority, the 1200B grating on DEIMOS is likely the better choice. However, if a specific element is of interest (e.g., Ca), one of the lower-resolution DEIMOS gratings might be more valuable than the 1200B grating.

Given the simplicity in its computation, we suggest that CRLBs should be standardized as part of observational planning for resolved star spectroscopic abundance measurements as a logical extension of the standard ETC usage. An ETC determines the S/N of a spectrum based on the integration time and observing conditions, and the CRLB in turn relates that S/N into an expected abundance precision. Figure 2.3 provides a clear example of how calculating CRLBs for an instrument can inform an observing strategy. If the intended science goals necessitate simply measuring Fe and an α element out beyond 100 kpc, an hour-long exposure with the D1200G grating will likely suffice, allowing for a handful of fields to be observed in a night. However, if the science requires measuring specifically the 2.3 element, Mg, an integration time of three or more hours is necessary per field and a different observing strategy is required.

2.7.4 Caveats and Assumptions

In this section we discuss in more detail the assumptions adopted in our calculation of CRLBs, namely that (1) the model spectra perfectly reproduce real stellar spectra, (2) the likelihood and noise properties are Gaussian, and (3) that adjacent pixels are uncorrelated. We save a more technical discussion of the CRLB for a biased estimator for Appendix 2.10.1.

2.7.4.1 Model Fidelity

Model fidelity is a fundamental assumption inherent in all problems of parameter estimation. The CRLB of stellar spectra is no exception to this as the gradient spectra used in the above calculations are strongly dependent on the physical assumptions and spectral line lists that underpin any spectral synthesis model. It is important to keep in mind that the CRLB makes no claims about the accuracy of stellar label measurements, merely the possible precision. Nevertheless, incomplete or incorrect line lists will leave out or misplace spectral information, while models that assume 1D atmospheres in local thermodynamic equilibrium (LTE) may incorrectly predict the spectral response to varying stellar labels for non-LTE lines. It is thus important to strive for consistency and consider the CRLBs calculated using the models relevant to the spectral fitting that will be conducted. While comparing CRLBs of different models is a valuable exercise to evaluate systematics in the predicted CRLBs, this should not be done to pass judgment on model quality.

A common practice in full-spectrum fitting is the masking of spectral regions that are known to be poorly fit by the spectral model to avoid introducing potential systematics into the analysis. Often the poor fit is due to non-LTE effects, but may also be the result of 3D effects, poorly calibrated oscillator strengths, or an incomplete (or incorrect) line list (see Nissen et al. 2018, and references therein). When these regions are masked, so too is the information that it holds. In such a case the appropriate CRLB should be calculated with gradient spectra masked in the same regions (as we do in Section 2.5.1.1), resulting in a higher uncertainty for the stellar labels. We note, however, that because information adds in quadrature, masking 90% of the lines only worsens the CRLB by a factor of $\sim$3. For a more thorough analysis of the dependence on masked regions, see Ting, Conroy, Rix, and Cargile (2017).

Another underlying challenge for our CRLBs is the assumption that the continuum can be perfectly determined. In the red-optical and near-infrared region of the spectrum, lines are sufficiently sparse that even at $R \sim 2000$ identifying the continuum and dividing it out is routine. Unfortunately, the many absorption features in the blue optical and UV, make it challenging to define a stellar continuum. Instead, a pseudo-continuum is defined using a polynomial function (or some smoothing kernel) and divided out, potentially introducing systematics or additional uncertainty in the normalized flux that will worsen the precision. By similarly normalizing the model spectra (instead of using the true continuum), any systematics introduced through imperfect normalization can be minimized.

Knowledge of the instrumental LSF is necessary to fit observed spectra with model spectra at the same resolving power. In this work, we have assumed a constant LSF. However, in practice, the LSF is not always known to great precision and can vary from object to object depending on where in the field of view the star lies. Use of the wrong LSF is thus another means by which systematics may be introduced into the fitting of stellar labels. Ting, Conroy, Rix, and Cargile (2017) showed that at least at moderate resolution ($R \sim 6000$) and high S/N ($>200\,\mathrm{pixel}^{-1}$), mismatched LSFs only bias stellar label recovery for differences in broadening greater than $10\,\mathrm{km\,s}^{-1}$ and is unlikely to affect the measurement precision.

Spectral fitting at lower resolving powers should be even less sensitive to mismatches in LSF.

In addition, when using rest-frame synthetic spectra, it is necessary to properly determine and correct for the radial velocity of stars. As with the continuum normalization and LSF, we have not quantified the uncertainty in stellar labels that is introduced when the radial velocity is fit simultaneously with other stellar labels. We expect any changes in the CRLBs to be small given that radial velocity is unlikely to correlate with other stellar labels. We will pursue this analysis in a future study.

Even with perfect spectral models, continuum normalization, and instrument characterization, fully extracting the chemical information content of a spectrum requires fitting the full wavelength range (as opposed to measuring EWs) for all stellar labels simultaneously. This is particularly important at low and moderate resolution to account for the degeneracies between labels introduced by blended spectral features. In practice, this can be computationally challenging owing to the high dimensionality of the stellar label space and the large runtimes needed to generate even 1D LTE stellar atmospheres.

Despite these challenges, the future of extragalactic stellar spectroscopy looks bright as steady progress is being made in all of the aforementioned areas. Attempts to incorporated non-LTE and 3D effects into stellar atmosphere and radiative transfer models have been undertaken by a number of groups (e.g., Caffau, Ludwig, et al. 2011; Bergemann, Lind, et al. 2012; Amarsi, Lind, et al. 2016). Several groups have committed to further refining line lists through the identification of unknown (or misplaced) lines in stellar spectra (e.g., Shetrone, Bizyaev, et al. 2015; Andreasen et al. 2016) and the improved calibration of transition oscillator strengths (e.g., Pickering et al. 2001; Aldenius et al. 2007; Pehlivan Rhodin et al. 2017; Laverick et al. 2018). Lastly, full spectrum-fitting techniques have made major strides with spectral "emulators" trained through data-driven (e.g., `the Cannon`; Ness, Hogg, et al. 2015), (e.g., `the Payne`; Ting, Conroy, Rix, and Cargile 2019) or combined (e.g., `the DD-Payne`; Xiang, Ting, et al. 2019) methods, which bypass the computationally expensive stellar atmosphere and radiative transfer calculations.

The above challenges to achieving the precision predicted by the CRLBs should not dissuade from the use of CRLBs. Instead, the precision forecasted by the CRLBs provide strong motivation for the continued efforts toward understanding stars, their atmospheres, and their spectra.

2.7.4.2 Assumptions of Gaussian Posteriors

Implicit in the derivations of Equations (2.1), (2.2), and (2.5) was that of Gaussian likelihoods and uncertainties. When these conditions are not met, the CRLB will inaccurately predict measurement errors and degeneracies between stellar labels. In such situations, a more accurate estimate of the achievable precision can be found using Bayesian sampling techniques. A comparison of the CRLB and the precision predicted by HMC sampling in the low-S/N limit is performed in Appendix 2.10.5, and we find it robust down to an S/N of 5 in the case of D1200G (assuming a constant S/N with wavelength).

2.7.4.3 Pixel-to-pixel Correlation

Throughout this study we simplify our analysis by setting the correlation between adjacent pixels to zero when calculating the CRLBs.[30] In practice, however, most spectrographs are designed to oversample their spectra such that the number of pixels per resolution element is larger than the Nyquist sampling ($\sim$2 pixels/FWHM).[31] As a result, adjacent pixels will show some correlation and not be truly independent as we have assumed.

While this is unlikely to make a large difference for most spectrographs, which only slightly oversample their spectra (3–4 pixels/FWHM), the pixel-to-pixel correlation of spectrographs that more highly oversample (e.g., Hectospec, FLAMES-UVES, FOBOS, and some DEIMOS and LRIS gratings) may be non-negligible in practice. If instead we believe that only 2 pixels per resolution element are informative then the CRLBs should be a factor of $\sqrt{2}$ $(\sqrt{3})$ larger than presented for spectrographs with a sampling of 4 (6) pixels/FWHM as the CRLBs scale as $n^{-1/2}$. More realistically, additional sampling beyond the Nyquist limit will yield pixels that are still informative, just less so than wholly independent pixels. Thus, we expect the increase in the CRLB to be considerably less than a factor of $\sqrt{2}$ $(\sqrt{3})$ when the correlation of adjacent pixels are taken into account. In Appendix 2.10.3, we present an illustrative example of the impact of wavelength sampling and pixel-to-pixel correlation on the CRLBs.

2.8 Chem-I-Calc

Forecasting stellar label recovery for spectroscopic observations is crucial to planning realistic observational campaigns and for validating the reported precision of spectral fitting analyses. However, there are far more combinations of instruments, observational conditions, and stellar targets than can be presented in a single paper. To make the calculation of stellar CRLBs convenient to the astronomical community, we have developed the open-source Python package, `Chem-I-Calc`—the Chemical Information Calculator (Sandford 2020).[32]

The `Chem-I-Calc` Python package provides all the tools necessary to perform all of the computational work presented in this paper, excluding the generation of high-resolution spectra. All of the calculations performed in this paper are included in a Jupyter Notebook on the `Chem-I-Calc` Github repository along with several other helpful tutorials and instructions for downloading the synthetic spectra described in Section 2.4. A basic usage example is included in Appendix 2.10.7. The code base is designed to be easy to modify for users that need more flexibility in their CRLB calculations (e.g., for incorporating wavelength-dependent resolution, alternative stellar models, or masking of specific wavelength regions).

[30] A similar simplification is employed nearly ubiquitously in the measurement of chemical abundances from stellar spectroscopy.

[31] For most instrumental LSFs, the Nyquist sampling is somewhat larger than 2 pixels/FWHM (see Robertson 2017).

[32] `https://github.com/NathanSandford/Chem-I-Calc`

While `Chem-I-Calc` is ready to be used in its current state, it is still under active development. Over time we expect to add additional commonly used spectrographs as presets and include a larger range of stellar types and metallicities as reference stars. We gratefully welcome community feedback and contributions to the Python package.

2.9 Summary

Current and future generations of powerful, highly multiplexed spectrographs on large-aperture telescopes make accessible an enormous wealth of chemical information in the spectra of stars outside the MW. Already these instruments have observed the spectra of tens of thousands of individual stars in extragalactic systems, enabling the measurement of their abundance patterns (e.g., Suda et al. 2017, and references therein). With the advent of large-scale extragalactic spectroscopic surveys and ELTs, the number of stars outside the MW with observed spectra will increase by at least an order of magnitude (Takada et al. 2014; Bundy et al. 2019; MSE Science Team et al. 2019).

The majority of these spectra will be acquired at low and moderate resolution ($R < 10000$) and feature heavy blending of spectral lines, necessitating the entire spectrum be fit for all stellar labels simultaneously. Recently, novel full-spectral fitting techniques (e.g., The Cannon; Ness, Hogg, et al. (2015), The Payne; Ting, Conroy, Rix, and Cargile (2019), and The DD-Payne; Xiang, Ting, et al. (2019)) applied to stellar spectra from MW surveys have proven capable of measuring dozens of elemental abundances from low-resolution spectra.

With the field of extragalactic stellar spectroscopy poised for substantial growth, it is imperative that we understand the chemical information content of the spectra we collect and the precision to which it enables the recovery of elemental abundances. To that end, we have employed CRLBs to quantify the information content of extragalactic stellar spectra and forecast chemical abundance precision for 41 existing, future, and proposed spectrograph configurations on 14 telescopes. Here we summarize our findings.

1. The CRLB is an efficient method for computing the expected precision of stellar labels determined via full spectral fitting. We find that the precision of literature abundances for the commonly used DEIMOS 1200G grating and the LAMOST MW survey are within a factor of 2 of our CRLBs.

2. Low- and moderate-resolution spectroscopy at blue-optical wavelengths ($\lambda \lesssim 4500\,\text{Å}$) are incredibly information rich, enabling the recovery of two to four times as many elemental abundances as red-optical spectroscopy ($5000 \lesssim \lambda \lesssim 10000\,\text{Å}$) at similar resolutions. Further, low-resolution, blue-optical spectroscopy is capable of constraining the abundances of several neutron-capture elements (e.g., Sr, Ba, La, Eu).

3. High-resolution ($R \gtrsim 20000$) spectra contain substantial chemical information even at low S/N ($\sim 10\,\text{pixel}^{-1}$). Maximizing the precision of abundance recovery from high-resolution spectra benefits from full spectral fitting over EW techniques.

4. Even small (∼100-500 Å) windows of low-S/N, high-resolution spectra can constrain [Fe/H] and a handful of other elements to better than 0.3 dex.

5. JWST/NIRSpec and ELTs can recover 10–30 elements for red giant stars throughout the Local Group and [Fe/H] and [α/Fe] for resolved stars in galaxies out to several Mpc with 6 hr (∼1 night) of integration time.

6. Our analysis strictly concerns the precision, not accuracy, of chemical abundance measurements. In practice, imperfect stellar models, line lists, and data reduction can introduce systematics that can bias abundance measurements and hinder attainment of near-CRLB precision. Further investment in the development of stellar models and spectral analysis is necessary to maximally use the chemical information content of the spectra collected.

7. CRLBs, like ETCs should be used when planning stellar spectroscopic observations or developing spectroscopic instrumentation. To facilitate the calculation of CRLBs, we present `Chem-I-Calc`, an open-source Python package for calculating CRLBs of arbitrary spectrograph configurations.

2.10 Appendices

2.10.1 Appendix: Biased CRLB

A fundamental assumption adopted in this work is that of perfect models that accurately reproduce observed stellar spectra. However, as in most of astrophysics and as we discussed in Section 2.7.4, this is not the case in practice. Many spectral features are poorly modeled due to 3D and non-LTE effects, miscalibrated oscillator strengths and transition wavelengths, and imperfect reductions. While these systematic errors primarily affect the accuracy of abundance measurements, they also invalidate our assumption that the MLE, $\hat{\theta}$, is an unbiased estimator of the true stellar labels and may also change the expected precision of the abundance measurements.

If the bias of a particular spectral model is known, this can be included in the prediction of stellar label precision using the "biased" or "misspecified" CRLB:

$$\sigma_{\text{biased},\alpha} = \sqrt{([I + D]F^{-1}[I + D]^T)_{\alpha\alpha}}, \tag{2.17}$$

where F is the FIM as defined in Equation (2.6), I is the identity matrix, and D is the bias gradient matrix:

$$D = \left[\frac{\partial b}{\partial \theta_\alpha}\right]_{\hat{\theta}}, \tag{2.18}$$

where b is the bias of your labels given by

$$b(\hat{\theta}) = E(\theta) - \hat{\theta}. \tag{2.19}$$

Because evaluating the bias is both model and instrument dependent, it is beyond the scope of this paper. However, we note that in the simple case of a uniform bias (i.e., measuring the surface temperature of all stars to be 100 K too hot), the normal and biased CRLB are the same. In the more complicated (and realistic) case where the bias is dependent on the stellar labels (i.e., the surface temperature is measured to be 100 K too hot in giant stars but 100 K too cold in dwarf stars) the biased CRLB will differ from normal CRLB. Depending on the direction and amplitude of the bias, this may result in either better or worse precision than in the unbiased case.

The main challenge in practice is not that the CRLBs cannot be used in the presence of bias, but that the bias needs to be known a priori for the CRLB—or any forecast of precision—to be computed accurately.

2.10.2 Appendix: CRLB Calculation

For instruments whose observations span noncontiguous wavelength ranges, the gradient spectra (and 1D S/N arrays) for each of the wavelength ranges are concatenated together. This technique can also be used to combine observations from potentially complementary instruments or observing campaigns, though we do not consider any here. All combinations of wavelength ranges examined in this work are forced to be non-overlapping to avoid a more complicated treatment of the spectral covariance matrix. This is done even though it means ignoring the additional information that an overlapping region of spectrum might provide.

From this point, the calculation of the CRLBs from the gradient spectra and spectral covariance is simply a matter of matrix multiplication and inversion. However, because the gradient spectrum for some labels is much larger than for others (i.e., Fe compared to Nb), the FIM may be nearly singular and thus unstable to inversion. We take several steps to avoid matrix inversion problems and calculate robust CRLBs:

(i) We divide the spectral gradient with respect to $T_{\rm eff}$ by 100.

(ii) If $F_{\alpha\alpha} < 1$ for any label, α, we set $F_{\alpha j} = F_{i\alpha} = 0$ and $F_{\alpha\alpha} = 10^{-6}$

(iii) We compute the Moore-Penrose pseudo-inverse of the FIM (Moore 1920; Penrose 1955).

The purpose of (i) is to place $df/dT_{\rm eff}$ on roughly the same scale as df/d[X/H]. This keeps the eigenvalue of the FIM with respect to $T_{\rm eff}$ from dwarfing those of the other labels. As a result, the CRLB for $T_{\rm eff}$ is in units of 100 K. Step (ii) avoids zero eigenvalues for labels with very little information in the spectrum. It also removes the covariance of these labels with all other labels, which would otherwise make the matrix nearly singular. This results in a CRLB of $\sim 10^3$ for these labels, which can safely be ignored. Finally, by calculating the pseudo-inverse instead of the true inverse of the FIM in (iii), we avoid numerical instabilities when attempting to invert near-singular matrices.

When including prior information into our CRLB calculations, we add the inverse variance of these priors to the relevant diagonal entries of the FIM as outlined in Equation (2.12)

before inverting the FIM as before. To be rigorously Bayesian, we ought to state that we do this for all labels, including those with uninformative priors with zero inverse variance.

2.10.3 Appendix: Wavelength Sampling and Pixel Correlations

To illustrate the impact of assuming the independence of all pixels in the calculation of the CRLB, we consider the simple case where each resolution element is sampled by 3 pixels and all adjacent pixels are correlated by some fraction, c. In such a scenario, the flux covariance is no longer the diagonal matrix presented in Equation (2.15), but now has diagonal-adjacent terms equal to $c\sigma^2$, where $\sigma = (\mathrm{S/N})^{-1}$ at each pixel:

$$\Sigma = \begin{bmatrix} \sigma^2(\lambda_1) & c\sigma^2(\lambda_1) & & & & \\ c\sigma^2(\lambda_2) & \sigma^2(\lambda_2) & c\sigma^2(\lambda_2) & & & \\ & & \ddots & & & \\ & & c\sigma^2(\lambda_{N-1}) & \sigma^2(\lambda_{N-1}) & c\sigma^2(\lambda_{N-1}) \\ & & & c\sigma^2(\lambda_1) & \sigma^2(\lambda_N) \end{bmatrix}. \tag{2.20}$$

Figure 2.20 shows the impact of assuming adjacent pixels are 10%, 30%, 50%, and 99% correlated on the CRLB as applied to our fiducial D1200G observation. For comparison, we also include the CRLBs assuming one, two, three, and four completely uncorrelated pixels per resolution element. As expected under the assumption of independent pixels, the CRLBs scale as $n^{-1/2}$, where n is the number of pixels per resolution element.

When adjacent pixels have correlations of 10%, 30%, and 50%, the CRLBs are roughly 8%, 23%, and 35% larger respectively than in the uncorrelated case. These CRLBs are equivalent to calculating the CRLB assuming $n = 2.6$, 2.0, and 1.6 independent pixels per resolution element respectively. In the extreme case that all three pixels are nearly 100% correlated with each other, there is effectively only one independent pixel per resolution element and the CRLB approaches the $n = 1$ pixel/FWHM CRLB or 3 times what is found with uncorrelated pixels.

A more realistic treatment of pixel correlation would require adopting a kernel describing the correlation of pixels beyond just the adjacent ones. This, however, requires a deep knowledge of each instrument, which is beyond the scope of this paper.

2.10.4 Appendix: Comparison with LAMOST DD-Payne Abundances

In Section 2.5.1.1, we found our CRLBs for D1200G to be in good agreement with the precision reported by Kirby, Xie, et al. (2018). D1200G observations of metal-poor RGB stars, however, provide only a single point of comparison between our forecasts and what might be expected in practice. Because so few full spectral fitting techniques are currently used in extragalactic contexts, similar comparisons are quite challenging.

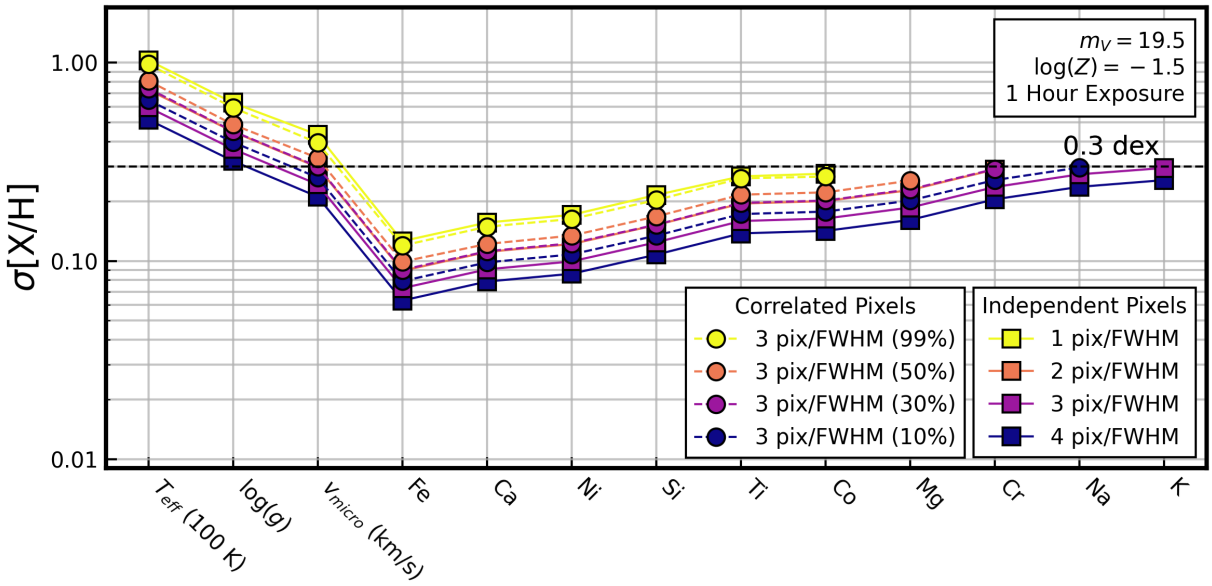


Figure 2.20: D1200G CRLBs for a 1 hr exposure of a $\log Z = -1.5$, $m_V = 19.5$ RGB star assuming various wavelength samplings and pixel-to-pixel correlations. CRLBs assuming uncorrelated pixels but varying wavelength sampling are represented by squares and solid lines. CRLBs assuming 3 pixels/FWHM but varying degrees of correlation between adjacent pixels are represented by circles and dashed lines. For completely independent pixels, the CRLBs scale proportionally $n^{-1/2}$, where n is the number of pixels per resolution element.)

Instead, we turn to an example within the Galaxy to provide an additional comparison. Specifically, we compare our CRLBs to the internal precision reported by Xiang, Ting, et al. (2019) for observations of MW stars by the LAMOST spectrograph (Cui et al. 2012). Xiang, Ting, et al. (2019) employed the DD-Payne[33] for full-spectral fitting and used repeat observations to quantify the internal precision of their measurements.

Because LAMOST observed primarily MW stars, we calculate the CRLBs for a typical solar-metallicity K-Giant star ($T_{\rm eff} = 4800\,{\rm K}$, $\log g = 2.5$, $v_{\rm micro} = 1.7\,{\rm km\,s^{-1}}$, $\log Z = 0$ and solar abundance patterns). To estimate the S/N of the LAMOST spectra, we use the mean flux variance from several LAMOST spectra of giant stars with a g-band S/N of $50\,{\rm pixel^{-1}}$. As in our comparison to Kirby, Xie, et al. (2018), we make several cuts on the sample in order to fairly compare the reported precision with our CRLBs, which we list in Table 2.6. These cuts leave the reported precision for approximately 6000 stars.

Because Xiang, Ting, et al. (2019) report their abundance precision in terms of [X/Fe], we add σ[[Fe/H] in quadrature to σ[[X/Fe] so that the CRLBs are on the same scale. Xiang,

[33] `The DD-Payne` is a hybrid spectral model that is trained on high-resolution measurements from GALAH and APOGEE and regularized on *ab initio* spectral gradients.

$4600 < T_{\rm eff}\ (\rm K) < 5000$
$2.3 < \log g < 2.7$
$-0.1 < [\rm Fe/H] < 0.1$
$-0.1 < [\alpha/\rm Fe] < 0.1$
$40 < g\text{-band S/N } \left(\rm pixel^{-1}\right) < 60$
χ^2 Flag = good
[X/Fe] Flag = 1

Table 2.6: Cuts on LAMOST DR5

Ting, et al. (2019) do provide estimated systematic uncertainties for their measurements, but because CRLBs are a measure of precision and not accuracy, we do not include them in this comparison.

Figure 2.21 shows the reported measurement precision of these stars compared to our LAMOST CRLBs. Similar to our comparison with Kirby, Xie, et al. (2018), we find that most abundances reported by Xiang, Ting, et al. (2019) are within a factor of ∼2 of our CRLBs. The largest difference is in the precision of $T_{\rm eff}$, which is reported to be 27 K, nearly three times larger than our predicted precision (10 K). This is not wholly unreasonable given the subtle and highly model-dependent effects that $T_{\rm eff}$ has on spectral features. The reported precision for Fe (0.029 dex) is also more than a factor of 2 larger than our forecast (0.013 dex)—though the absolute difference is quite small. We suspect this is driven by the larger uncertainties found for $T_{\rm eff}$ and $\log g$ by Xiang, Ting, et al. (2019) and the substantial correlation these labels have with Fe in giant stars.

Interestingly, we find that the precision reported for Ni, O, and C outperforms the CRLB by a factor of 1.2, 1.7, and 2.1. We suspect that this might be the result of "gradient aliasing" in the `DD-Payne`, whereby the model picks up spectral gradient features from elements other than the one it attributes them to. This is a common challenge in data-driven methods, and while Xiang, Ting, et al. (2019) attempted to mitigate it by regularizing the model with *ab initio* spectral gradients, some gradient aliasing may remain. For the remaining abundances, there are several reasons why slightly poorer precision might be expected in practice, including model fidelity and imperfect calibrations (see Section 2.7.4 for further discussion). Together, the comparisons conducted here and in Section 2.5.1.1 illustrate that the CRLBs are quite reasonable representations of contemporary abundance measurements.

2.10.5 Appendix: Validation of CRLBs

To validate the robustness of the CRLBs, we infer the stellar labels of a mock spectrum at various S/Ns using an *ab initio* trained spectral model and an HMC sampling method and compare the precision of this inference with the precision forecasted by the CRLBs. We outline the process of training the spectral model in Appendix 2.10.5.1 and fitting the mock

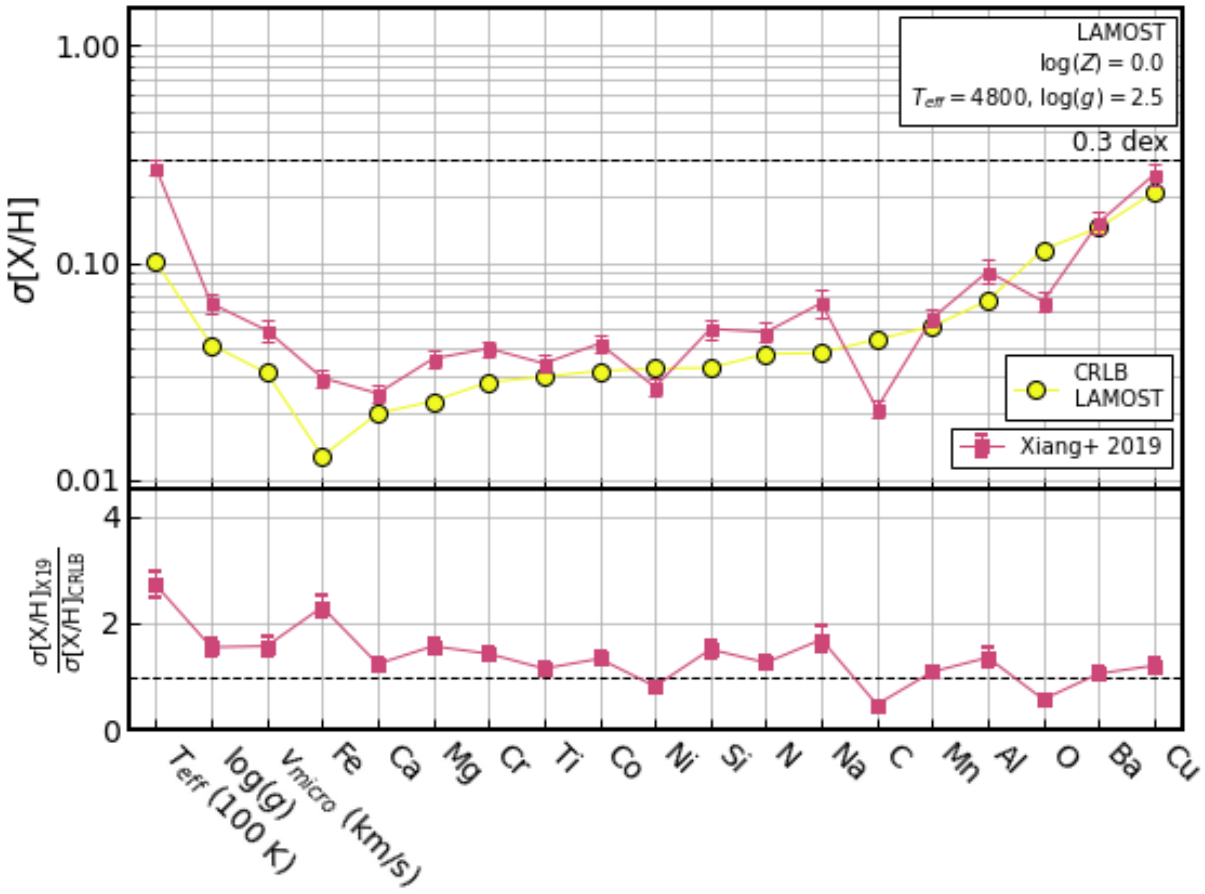


Figure 2.21: (Top) LAMOST CRLBs for a typical solar-metallicity K-Giant with a g-band S/N of 50 pixel^{-1} overplotted with the internal precision of ∼6000 comparable stars report by Xiang, Ting, et al. (2019). Error bars denote the upper and lower quartiles of the sample. (Bottom) The ratio of the forecasted LAMOST CRLBs to the reported precision for each stellar label. As found with the comparison to Kirby, Xie, et al. (2018) in Figure 2.4, the measurement uncertainties for most elements are generally a factor of $\lesssim$2 larger than the CRLBs. The reported precision for Ni, C, and O slightly outperform the CRLBs, which may be the result of additional spectral information included by the data-driven model of *ab initio* that is not incorporated in our purely *ab initio* model.

spectrum in Appendix 2.10.5.2. The results of this comparison are presented in Appendix 2.10.5.3.

2.10.5.1 Appendix: Training a Spectral Model

Training a spectral model requires a large set of stellar spectra with known labels that span the relevant parameter space. To generate this training set, we randomly drew 10^4 stellar labels from the following uniform distribution:[34]

$$\begin{aligned} T_{\rm eff} &\sim \mathcal{U}(4500\ {\rm K}, 5000\ {\rm K}), \\ \log g &\sim \mathcal{U}(1.5, 2.1), \\ v_{micro} &\sim \mathcal{U}(1.4\ {\rm km\ s^{-1}}, 2.4\ {\rm km\ s^{-1}}),\ \text{and} \\ [{\rm X/H}] &\sim \mathcal{U}(-0.5, 0.5), \end{aligned}$$

where in this case X refers to a smaller subset of elements: Fe, Ca, Ni, Si, Ti, Mg, and Co. We only considered 7 elements, limiting the model to 10 stellar labels, to simplify the training process. These specific elements were chosen as they are the most precisely recovered elements by the D1200G setup (see Section 2.5.1 and Table 2.2). The bounds of the uniform distributions are chosen to center on the parameters of our fiducial RGB star (Table 2.1) and span roughly two times the D1200G (S/N = 50 pixel^{-1}) CRLB for each stellar label, assuming the Gaussian priors of $\sigma_{T_{\rm eff}} = 100\,{\rm K}$, $\sigma_{\log g} = 0.15$, and $\sigma_{\rm micro} = 0.25\,{\rm km\,s^{-1}}$ used previously in Section 2.5.1.4. Spectra were generated and convolved to instrumental resolution as previously described in Section 2.4.2.

Withholding 2500 spectra for validation, we train an updated version of `The Payne`[35] (details in Table 2.7). To aid the training process, the labels are normalized according to

$$\theta_{i,{\rm scaled}} = \frac{\theta_i - \theta_{i,{\rm min}}}{\theta_{i,{\rm max}} - \theta_{i,{\rm min}}} - 0.5, \tag{2.21}$$

where $\theta_{i,{\rm min}}$ and $\theta_{i,{\rm max}}$ are the minimum and maximum values represented in the training and validation data sets. After 10^5 training steps, which takes roughly 4 hr on an Nvidia K80 GPU, the model that minimized the L1 mean loss on the validation spectra is chosen as the final model.

We compare *ab initio* spectra from our validation set to spectra generated with the same labels using `The Payne` and find mean interpolation errors of individual pixels to be less than 0.1%. These errors are much smaller than typical observational uncertainties in the normalized spectra.

2.10.5.2 Appendix: Fitting Mock Spectra with HMC Sampling

The mock spectrum is generated using The Payne at the labels of the fiducial $\log Z = -1.5$ RGB star to avoid introducing any bias that may have been introduced in the training of the

[34] $\mathcal{O}(10^3)$ stellar spectra would likely have been sufficient, but opted to generate 10^4 to further reduce emulation errors.

[35] `https://github.com/tingyuansen/The_Payne`

# Training Spectra	7500
# Validation Spectra	2500
# Spectra / Batch	512
# Hidden Dense Layers	2
# Neurons / Layer	300
Activation Function	Leaky ReLU
# Training Steps	10^5
Loss Function	L1 Mean
Optimizer	Rectified Adam
Learning Rate	10^{-3}
Interpolation Errors	$< 0.1\%$

Table 2.7: Details of `The Payne`.

spectral model—recall that we are interested in precision, not accuracy, here. We assume a constant S/N across the entire spectrum, which manifests as an uncertainty in each pixel of $\sigma = f(\lambda)/(\mathrm{S/N})$, where $f(\lambda)$ is the normalized flux of the model. With the same mock spectrum, we perform the fitting assuming a range in S/N from 5 to 200 pixel^{-1} that is constant across the entire wavelength coverage.

With only 10 stellar labels and likelihoods that we believe to be close to Gaussian, using an MCMC sampling technique would likely be adequate for this scenario. However, because our neural network spectral emulator is differentiable, we opt to use an HMC sampler, making it readily adapted for inference with many more labels where an MCMC sampler might face convergence problems.

We adopt the Gaussian likelihood function in Equation (2.1) and the following priors:

$$\begin{aligned} T_{\mathrm{eff}} &\sim \mathcal{N}^*(4750\ \mathrm{K}, 100\ \mathrm{K}), \\ \log g &\sim \mathcal{N}^*(1.8, 0.15), \\ v_{micro} &\sim \mathcal{N}^*(1.9\ \mathrm{km\ s^{-1}}, 0.25\ \mathrm{km\ s^{-1}}), \text{ and} \\ [\mathrm{X/H}] &\sim \mathcal{U}(-0.5, 0.5), \text{ and} \\ [\mathrm{X^*/H}] &\sim \delta(0.0), \end{aligned}$$

where $\mathcal{N}^*(\mu, \sigma)$, represents a normal distribution truncated at the limits of the training set so that the model does not extrapolate. Here, X* refers to elements that the CRLB predict cannot be recovered to better than 0.3 dex at the given S/N. These elements are held fixed at solar value, which is equivalent to applying a delta function prior at [X/H] = 0.0. The fixed labels at each S/N are displayed in Table 2.8. For each S/N we perform the HMC sampling using 24 parallel chains. Each chain begins with 3000 burn-in samples, which are discarded, followed by another 3000 samples, which constitute our posterior sample.

S/N (pix^{-1})	Fixed Labels
5, 10	[Ni/H], [Si/H], [Ti/H], [Co/H], [Mg/H]
15	[Si/H], [Ti/H], [Co/H], [Mg/H]
20	[Co/H], [Mg/H]
30, 50, 100, 200	None

Table 2.8: Fixed stellar labels at each S/N.

2.10.5.3 Appendix: Comparison to the CRLB

In Figure 2.22, we plot the difference between the precision predicted by the CRLBs and the standard deviation of the mock fit posteriors for each S/N. In the calculation of the CRLBs, we include the same priors on $T_{\rm eff}$, $\log g$, and $v_{\rm micro}$ used in the HMC sampling. In addition, for each S/N, we only consider the gradients for the stellar labels that are left free in the sampling (see Table 2.8), thus holding all other labels fixed at solar values. Instead of calculating spectral gradients from *ab initio* spectra, we calculate the gradients from our trained spectral model to exclude any systematics introduced by interpolation errors of the model.

In general, we find the CRLBs and the standard deviations of the mock fits to be in agreement at the 0.01 dex level down to an S/N of 10 pixel^{-1} and at the 0.02 dex level down to an S/N of 5 pixel^{-1}. At very high S/N (200 pixel^{-1}), the CRLBs accurately predict the precision of the $v_{\rm micro}$ and all chemical abundances, only very slightly under-predicting the precision of $T_{\rm eff}$ by 1 K and $\log g$ by 0.01 dex. As the S/N decreases to 20 pixel^{-1}, the difference grows to 5 K and 0.02 dex in $T_{\rm eff}$ and $\log g$, respectively, and the CRLBs slightly over-predict the precision for Si, Ti, and Mg by no more than 0.01 dex. All of these differences remain relatively small compared to the typical precision found for these labels and are the result of the posteriors of these labels being slightly non-Gaussian (negatively skewed).

As the S/N decreases further, the precision of both the mock fit and the CRLB become prior dominated for $T_{\rm eff}$, $\log g$, and $v_{\rm micro}$, resulting in a smaller difference in the precision of $T_{\rm eff}$. This is not the case for the precision of $\log g$ and $v_{\rm micro}$ due to the difference between the Gaussian prior included in the CRLB calculation and the truncated Gaussian included in the HMC sampling. Still, the differences are only $\sim$0.02 dex, which is quite minor in relation to the expected precision at S/N < 15 pixel^{-1}. Thus, we find that the CRLB is a robust predictor of stellar label precision down to at least an S/N of 15 pixel^{-1}.

2.10.6 Appendix: DESI CRLBs

DESI is a fiber-fed MOS that covers a wavelength range from 3600 to 9800 Å with a resolving power of 2000–5000. The primary science goal of the DESI survey is not galactic archaeology, nor is the 4 meter Mayall telescope it is mounted on large enough to efficiently observe

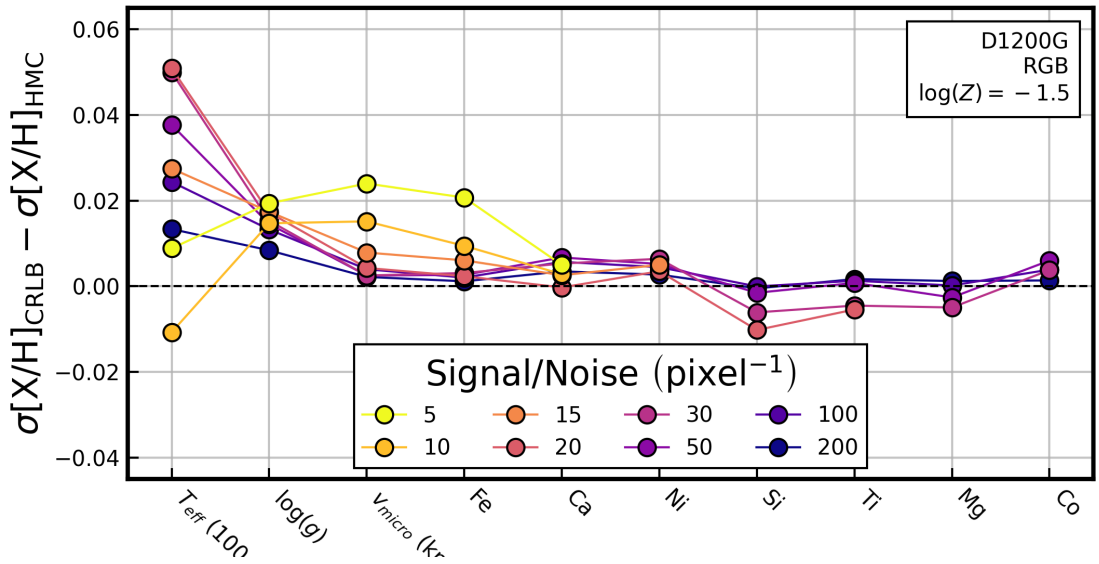


Figure 2.22: The difference between the CRLB and the stellar label precision found through HMC sampling for a $\log Z = -1.5$ RGB star observed with the D1200G setup. A constant S/N across the wavelength coverage was assumed. Differences are small ($\lesssim 5\,\mathrm{K}$ for σ_{T_eff}; $\lesssim 0.02\,\mathrm{dex}$ for $\sigma_{\log g}$; $\lesssim 0.02\,\mathrm{km\,s^{-1}}$ for $\sigma_{v_\mathrm{micro}}$; and $\lesssim 0.02\,\mathrm{dex}$ for $\sigma_{[\mathrm{X/H}]}$), indicating that the CRLB is a robust predictor of stellar label precision down to at least S/N $\sim 15\,\mathrm{pixel^{-1}}$.

resolved stars in dwarf galaxies. Nevertheless, it is a particularly interesting spectrograph for stellar chemical abundance measurements. When observing conditions are too poor for faint galaxy work, DESI will target bright galaxies, filling unused fibers with MW stars. This will yield spectra for roughly 10 million MW stars. In addition to many thin- and thick-disk stars, these deep observations are expected to reach MSTO stars in the MW's halo out to 30 kpc, allowing for a dramatically improved understanding of the stellar halo's chemical composition. In addition, the DESI instrumental design has been a major inspiration for current and next-generation survey instruments that will be targeting stars in dwarf galaxies.

Thus, while DESI will not be observing dwarf galaxy stars, we still think it valuable to present the theoretical abundance precision achievable by DESI in the MW halo. For these calculations we assume a uniform S/N of $30\,\mathrm{pixel^{-1}}$, which should be achievable for stars of $m_r = 16.5$–18 in a short 5–10 minute exposure (DESI Collaboration, Aghamousa, et al. 2016). The spectroscopic configuration used is given in Table 2.2. Because DESI will be able to observe down to the MSTO in the halo, we calculate the CRLBs for MSTO, RGB, and TRGB stars as done for D1200G in Section 2.5.1.3.

In Figure 2.23, we plot the CRLBs for DESI, illustrating its capability to extend the precise chemical abundance measurements of MW-disk surveys out to the MW's halo. As

seen for D1200G in Figure 2.6, abundance recovery is more precise for cool giants due to stronger absorption features and less precise for hot subgiants, which have weaker absorption features.

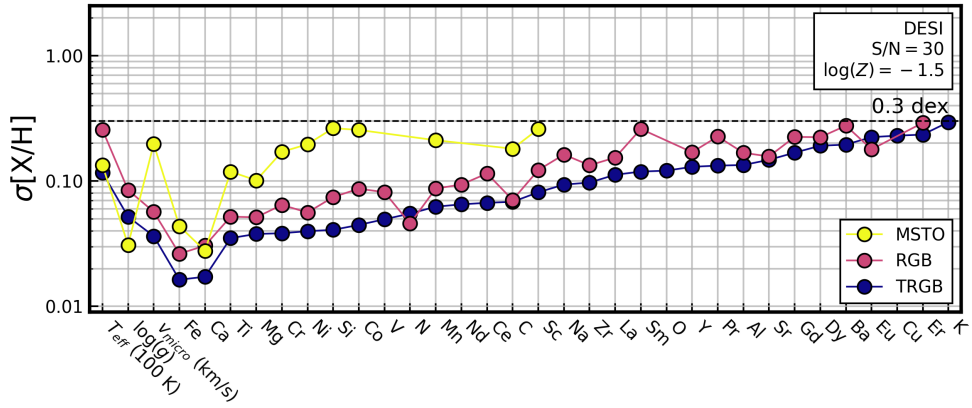


Figure 2.23: DESI CRLBs of $\log Z = -1.5$ MSTO, RGB, and TRGB stars with a constant S/N of 30 pixel^{-1}. The atmosphere parameters for each star can be found in Table 2.1. Just as for D1200G, abundance recovery is more precise for cool giants and less precise for hot sub-giants.

2.10.7 Appendix: `Chem-I-Calc` Example

Here we provide a basic usage example for `Chem-I-Calc`. More in-depth documentation, including installation instructions, can be found at `https://github.com/NathanSandford/Chem-I-Calc`.

1a. Set Reference Star

Most users will likely want to instantiate a `ReferenceSpectra` object from a pre-computed reference star that reasonably represents the characteristics of the observed star(s). For this example, we consider a RGB star with $\log(Z) = -1.5$ as our reference star.

```
from chemicalc import reference_spectra as ref

RGB = ref.ReferenceSpectra(reference="RGB_m1.5")
```

A complete list of pre-computed reference stars and a detailed description of the `ReferenceSpectra` object can be found at `https://chem-i-calc.readthedocs.io/en/latest/user/refstars.html`.

1b. Set Spectrograph Configuration

Next, we will need to instantiate an `InstConfig` object, which defines the specific properties of the spectrograph. This can be done using one of the many pre-configured instrument setups included in `Chem-I-Calc` by calling `AllInst.get_spectrograph` (see `d1200g` below). Alternatively, a custom instrument configuration can be created by calling `InstConfig` directly with the desired parameters (see `my_spec` below).

```
from chemicalc import instruments as inst

d1200g = inst.AllInst.get_spectrograph("DEIMOS 1200G")
my_spec = inst.InstConfig(
    name="My Spectrograph",
    res=5000,     # Resolving Power
    samp=3,       # Pixels / Resolution Element
    start=6000,   # Blue Wavelength Bound (in Angstrom)
    end=10000,    # Red Wavelength Bound (in Angstrom)
)
```

A detailed description of the `InstConfig` object and a list of the pre-configured instruments can be found at `https://chem-i-calc.readthedocs.io/en/latest/user/instruments.html`.

1c. Set Spectrograph Configuration

Before we calculate the CRLBs, we must also set the Signal/Noise (S/N) of our observation using the `set_snr` method of the `InstConfig` object. This method can take the following types of arguments:

- An integer or float: This applies a S/N that is constant with wavelength (see `d1200g` below).
- An array of integers or floats: This applies a wavelength-depend S/N that is interpolated onto the wavelength grid of the instrumental configuration (see `my_sppec` below).
- An ETC query from `chemicalc.s2n` (e.g., `Sig2NoiseDEIMOS`)

```
import numpy as np

d1200g.set_snr(100)  # Set constant S/N of 100

my_snr = np.vstack(
    [
        np.linspace(
            my_spec.start_wavelength,
            my_spec.end_wavelength,
            100,
        ),  # wavelength array
```

```
        np.linspace(
            50,
            100,
            100,
        )  # S/N array
    ]
)
my_spec.set_snr(my_snr)  # Set wavelength-dependent S/N
```

For more information on acceptable S/N inputs, including how to use the ETC querying functionality of `Chem-I-Calc`, see `https://chem-i-calc.readthedocs.io/en/latest/user/snr.html`.

2. Convolve Reference Spectra to Instrument Resolution

Next, we convolve the high-resolution ($R \sim 300,000$) reference spectra down to the resolving power of the instrumental setups by passing our `InstConfig` objects to the `convolve` method of our `ReferenceSpectra` object. If the wavelength grid of the instrument is large, this may be somewhat computationally taxing.

```
RGB.convolve(d1200g)
RGB.convolve(my_spec)
```

3. Calculate Gradient Spectra

Next, we calculate the partial derivatives of the reference spectrum with respect to the stellar labels using the `calc_gradient` method of the `ReferenceSpectra` object. This method takes as an argument either the name of an instrument setup (e.g., "DEIMOS 1200G") or an `InstConfig` object (e.g., `my_spec`).

```
RGB.calc_gradient("DEIMOS 1200G")
RGB.calc_gradient(my_spec)
```

4. Calculate CRLBs

Before calculating the CRLBs, we use the `init_crlb_df` function to initialize an empty `pd.DataFrame` with indices corresponding to the stellar labels of `ReferenceSpectra`. Then we calculate the CRLBs using the `calc_crlb` function for each `InstConfig` and store the results in a column of the CRLB DataFrame.

```
from chemiclac.crlb import init_crlb_df

CRLB_example = init_crlb_df(RGB)

CRLB_example["DEIMOS 1200G"] = calc_crlb(RGB, d1200g)
CRLB_example["My Spectrograph"] = calc_crlb(RGB, my_spec)
```

5. Apply Cutoff and Sort CRLBs

Using the `sort_crlb` function, we can sort the DataFrame of CRLBs in order of decreasing precision and remove all CRLBs larger than a specified cutoff value (here 0.3 dex). Setting the argument `fancy_labels=True` replaces the labels for effective temperature, surface gravity, and microturbulent velocity with LaTeX formatted labels for plotting.

```
from chemiclac.crlb import sort_crlb

CRLB_example = sort_crlb(CRLB_example, cutoff=0.3, fancy_labels=True)
```

6. Plot CRLBs

Finally we can plot the CRLBs for our observing scenario.

```
from chemiclac.plot import plot_crlb

fig = plot_crlb(
    CRLB_example,
    labels="Example CRLBs\n$\log(Z)=-1.5$ RGB",
    cutoff=0.3,
    cutoff_label_yoffset=0.02,
    ylim=(0.009, 1.7)
)
```

For more plotting options, see `https://chem-i-calc.readthedocs.io/en/latest/user/plotting.html`.

Chapter 3

Validating Stellar Abundance Measurements from Multi-Resolution Spectroscopy

In the previous chapter, I forecasted the chemical abundance precision from extragalactic stellar spectroscopy with current and future spectrographs. A primary conclusion of that work was that, theoretically, low-resolution spectroscopy is capable of measuring detailed stellar chemical abundance patterns previously considered only accessible to high-resolution observations. Achieving these predicted precisions, however, is non-trivial as the full-spectrum fitting techniques required to measure chemistry from the blended features of low-resolution spectra are heavily reliant on the adopted stellar spectral models, which we know to be imperfect. In this chapter, I provide a test chemical abundance recovery as a function of resolution, providing a first-order validation of the forecasts in Chapter 2.

This chapter has been adapted from the paper Sandford, Weisz, et al. (2023), which was published in the Astrophysical Journal Supplement Series. It was coauthored by the following individuals, and is included in this thesis with their express permission: Daniel Weisz, Yuan-Sen Ting.

3.1 Abstract

Large-scale surveys will provide spectroscopy for ∼50 million resolved stars in the Milky Way and Local Group. However, these data will have a high degree of heterogeneity and most will be low-resolution ($R < 10{,}000$), posing challenges to measuring consistent and reliable stellar labels. Here, we introduce a framework for identifying and remedying these issues. By simultaneously fitting the full spectrum and *Gaia* photometry with `the Payne`, we measure ∼30 abundances for 8 metal-poor red giants in M15. From degraded quality Keck/HIRES spectra, we evaluate trends with resolution and signal-to-noise (S/N) and find that (i) ∼20

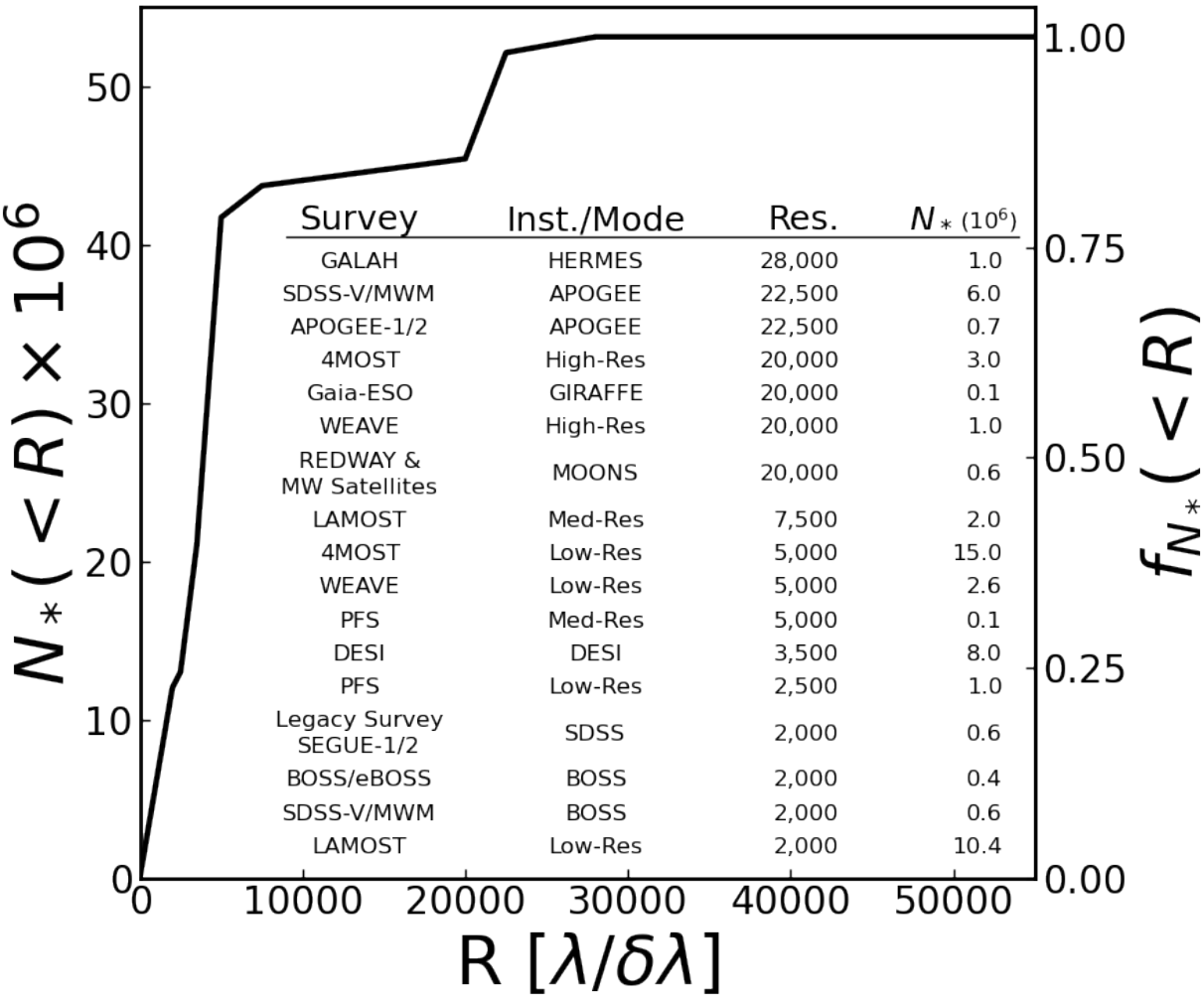
abundances are recovered consistently within $\lesssim$0.1 dex agreement and with $\lesssim$0.05–0.15 dex systematic uncertainties from 10,000 $\lesssim R \lesssim$ 80,000; (ii) for 9 elements (C, Mg, Ca, Sc, Ti, Fe, Ni, Y, Nd), this systematic precision and accuracy extends down to $R \sim 2500$; and (iii) while most elements do not exhibit strong S/N-dependent systematics, there are non-negligible biases for 4 elements (C, Mg, Ca, and Dy) below S/N $\sim$ 10 pixel^{-1}. We compare statistical uncertainties from MCMC sampling to the easier-to-compute Cramér-Rao bounds and find that they agree for $\sim$85% of elements, indicating the latter to be a reliable and faster way to estimate uncertainties. Our analysis illustrates the great promise of low-resolution spectroscopy for stellar chemical abundance work in the low-metallicity regime, and ongoing improvements to stellar models (e.g., 3D-NLTE physics) will only further extend its viability to more stars, more elements, and higher precision and accuracy.

3.2 Introduction

Astronomy is in the midst of a multi-decade golden era of stellar spectroscopy. Large spectroscopic surveys (e.g., APOGEE; Majewski et al. (2017), GALAH; De Silva et al. (2015), LAMOST; Cui et al. (2012), Gaia-ESO; Gilmore et al. (2012), Gaia-RVS; Recio-Blanco et al. (2022), DESI; Cooper et al. (2023)), are mapping the detailed chemical abundance patterns of millions of stars across the Milky Way (MW), and in doing so have ushered in a renaissance of chemodynamical studies seeking to piece together the complex formation history of the MW and its satellite system. Meanwhile, deep observations with 6+ meter telescopes have pushed the limits of resolved star spectroscopy beyond the MW and have begun unveiling the chemical evolution of other Local Group (LG) galaxies (e.g., Kirby, Xie, et al. 2018; Escala, Kirby, et al. 2019; Gilbert et al. 2019), including those that are relics from the early universe (e.g., Tolstoy, Hill, et al. 2009; Simon 2019, and references therein).

Over the course of the coming decade, the next iteration of ambitious stellar spectroscopic surveys (e.g., WEAVE; Dalton et al. (2016), SDSS-V; Kollmeier et al. (2017), PFS; Tamura et al. (2018), MOONS; Taylor et al. (2018), 4MOST; de Jong et al. (2019), FOBOS; Bundy et al. (2019)) will deliver an order-of-magnitude gain in the number of stars for which detailed chemical abundance patterns can be measured. By $\sim$2030, stellar spectra will be acquired for roughly 50 million resolved stars throughout the MW and LG (Figure 3.1). Spectrographs on next-generation large-aperture space- and ground-based telescopes (e.g., JWST; Gardner et al. (2006), GMT; Fanson et al. (2020), TMT; Skidmore et al. (2015), E-ELT Gilmozzi et al. (2007)) will further supplement these surveys; their unparalleled sensitivity and light-collecting power enabling spectroscopic observations out to several Mpc, far beyond the capabilities of existing ground-based facilities (Sandford, Weisz, et al. 2020).

However, the vast increase in data volume and availability made possible by these past, present, and future observations also pose newfound technical challenges. The combination of these large and numerous spectroscopic datasets will feature a high degree of heterogeneity across wavelength regime, signal-to-noise (S/N), and spectral resolving power ($R \equiv \lambda/\delta\lambda$), all of which can introduce complications in deriving consistent and reliable stellar chemical



Survey	Inst./Mode	Res.	N_* (10^6)
GALAH	HERMES	28,000	1.0
SDSS-V/MWM	APOGEE	22,500	6.0
APOGEE-1/2	APOGEE	22,500	0.7
4MOST	High-Res	20,000	3.0
Gaia-ESO	GIRAFFE	20,000	0.1
WEAVE	High-Res	20,000	1.0
REDWAY & MW Satellites	MOONS	20,000	0.6
LAMOST	Med-Res	7,500	2.0
4MOST	Low-Res	5,000	15.0
WEAVE	Low-Res	5,000	2.6
PFS	Med-Res	5,000	0.1
DESI	DESI	3,500	8.0
PFS	Low-Res	2,500	1.0
Legacy Survey SEGUE-1/2	SDSS	2,000	0.6
BOSS/eBOSS	BOSS	2,000	0.4
SDSS-V/MWM	BOSS	2,000	0.6
LAMOST	Low-Res	2,000	10.4

Figure 3.1: Forecasted number of stars observed by large spectroscopic surveys by ∼2030 as a function of spectral resolving power. Surveys with very limited wavelength coverage suitable (e.g., RAVE, Gaia-RVS, H3) are excluded. Surveys with fewer than 10^5 stars are also excluded as they contribute to the figure imperceptibly. Survey overlap is not considered. In 2030, ∼75% of the >50 million observed stellar spectra in the MW and LG will be taken at $R < 10{,}000$.

abundance measurements (Jofré, Heiter, and Soubiran 2019, and references therein).

As can be seen in Figure 3.1, the majority (75%) of the resolved star spectra acquired in the next decade will be obtained at "low resolution" ($R < 10{,}000$), where lower dispersion, higher throughput, and improved multiplexing provide both better observational efficiency and access to fainter and more distant stars. For these same reasons, the relative prolificity

of low-resolution stellar spectroscopy becomes more pronounced with increasing distance—very few stars beyond a few hundred kpc will have high-resolution spectroscopy of modest or higher S/N ($\gtrsim$40 pixel^{-1}) available. The trade-off is that low-resolution stellar spectroscopy suffers from severe blending of absorption features, which necessitates full spectral modeling and robust synthetic stellar spectra to precisely and accurately measure detailed chemical abundance patterns.

While the combination of low-resolution spectroscopy and full spectral fitting has lead to enormous scientific gains (e.g., Kirby, Guhathakurta, Bolte, et al. 2009; Kirby, Guhathakurta, Simon, et al. 2010; Ting, Rix, et al. 2017; Kirby, Xie, et al. 2018; Ting, Hawkins, et al. 2018; Xiang, Ting, et al. 2019; Wang, Hayden, et al. 2022), a variety of questions remain about the fidelity of abundance recovery in the low-resolution regime given their heavy reliance on synthetic stellar models. Namely, a major concern is that most spectral models used for full-spectrum fitting do not or do not fully capture the 3D and non-local thermodynamic equilibrium (NLTE) effects of the stellar atmosphere on line formation. Similarly, despite ongoing and sustained efforts (e.g., Lawler et al. 2013; Ryabchikova et al. 2015; Den Hartog et al. 2019; Smith et al. 2021, to just name a few contributions), there are many atomic and molecular transitions that are missing or imperfectly calibrated in the linelists employed by these spectral models.

For high-resolution observations, imperfections in the spectral model can be sidestepped by simply ignoring problematic features. But for low-resolution observations, poorly modeled spectral features become blended and inseparable from neighboring features and may introduce systematic biases and uncertainties into the measured chemical abundances if they are not handled carefully (Nissen et al. 2018). Given the ongoing proliferation of low-resolution stellar spectroscopy and the crucial role that low-resolution observations will play in extragalactic chemical abundance measurements, quantifying and addressing the systematics incurred as a function of resolution will be of the utmost importance. Without a firm grasp of these systematics, it will be difficult to draw robust conclusions across the disparate datasets, especially between the high-resolution studies that define our understanding of the MW and the low-resolution studies that provide our only window into galaxies beyond 1 Mpc.

It is relatively common practice in low-resolution stellar chemical abundance studies to correct for systematic biases, quantify systematic uncertainties, or otherwise validate the fidelity of low-resolution measurements by comparing these measurements with high-resolution literature measurements for a subset of stars (e.g., Kirby, Guhathakurta, Simon, et al. 2010). In may cases, however, these cross-validations are themselves quite heterogeneous, featuring measurements made with both full-spectrum fitting techniques and classical equivalent width (EW) fitting techniques, which frequently employ a great diversity of model atmospheres, spectral synthesis codes, and line lists (e.g., see Table 9 of Kirby, Guhathakurta, Simon, et al. 2010). While many studies (e.g., Bedell et al. 2014; Hinkel et al. 2016; Jofré, Heiter, Worley, et al. 2017; Blanco-Cuaresma 2019; Arentsen et al. 2022) have attempted to quantify methodological, instrumental, or model-oriented systematics, we are aware of no studies to date, which perform a comparison of abundance measurements as a function of resolving power using solely full-spectrum fitting techniques.

It is worth taking a moment to mention that for some scientific purposes, namely kinematic studies, high-resolution low-S/N ($\gtrsim$5 pixel^{-1}) spectra is sufficient. In these instances, multi-element abundance measurements are not attempted as historically only high-resolution spectra with moderate to high S/N ($\gtrsim$ 40 pixel^{-1}) have been deemed sufficient (Jofré, Heiter, and Soubiran 2019). In large part, this is because EWs are challenging to measure precisely in noisy spectra and can lead to biased results (e.g., Heiter, Soubiran, et al. 2014; Smiljanic et al. 2014). Consequently, high-resolution spectroscopy, even with large 10-m telescopes like Keck, has been limited to relatively bright stars ($r < 19.5$), excluding all but the brightest RGB stars in nearby dwarf galaxies (Simon 2019). Full spectrum fitting techniques, however, are predicted to better leverage the information content of low S/N spectra—even if a single noisy absorption line is only weakly informative, the ensemble of all spectral features should still provide strong constraints on the chemical abundances of a star (Ting, Conroy, Rix, and Cargile 2017; Sandford, Weisz, et al. 2020). While applications of full spectrum fitting to high-resolution stellar spectroscopy are becoming more common place, most are concerned with bright MW stars for which acquiring high-S/N spectra is relatively easy. The utility of low-S/N high-resolution spectra for chemical abundance measurements, especially for extragalactic metal-poor stars, has yet to be thoroughly demonstrated.

In this paper, we quantify the systematic biases and uncertainties in stellar chemical abundance measurements as a function of resolution and S/N by applying self-consistent full-spectrum fitting techniques to initially exquisite Keck/HIRES spectra ($R >$ 50,000, S/N $>$ 100 pixel^{-1}) that we have artificially degraded to lower resolution and S/N ($R \sim 2500$; S/N $\sim$ 5 pixel^{-1}). By fitting real observations from a single instrument, as opposed to mock spectra or observations from multiple spectrographs, we capture the impact of model inaccuracies on stellar label recovery when propagated to lower resolutions, while reducing complicating factors associated with different instruments, reduction pipelines, observing conditions, and stellar models.

Our sample consists of 8 metal-poor red giant branch (RGB) stars in MW globular cluster M15 with a rich history of previous study on which we validate our measurements. We choose to specifically investigate metal-poor RGB stars in this paper as these will make up the majority of stars observed extragalactically (e.g., in stellar streams and dwarf galaxies), where low-resolution spectroscopy will be the nearly universal observing mode. As such, a detailed understanding of the resolution- and S/N-dependence of abundance measurements in these metal-poor stars is crucial to place the majority of extragalactic resolved star spectroscopy on firm footing. Measurements for solar-metallicity stars and dwarf stars, on the other hand, benefit from their ubiquity in MW surveys where the fidelity of low-resolution measurements can be empirically verified against high-resolution measurements. That said, an extension of the methodology presented in this paper to a larger sample of stars with a wider range in metallicities and stellar parameters would prove enormously useful in quantifying the isolated impacts of resolution and S/N on the stellar chemical abundance measurements more broadly. Such work will be pursued in a future study.

This paper is structured as follows. In Section 3.3, we describe the archival data and their

degradation to lower resolution and S/N. We present our full-spectrum fitting techniques in Section 3.4. In Section 3.5, we present our results as a function of resolution and S/N. We discuss our primary findings in Section 3.6 and present our conclusions in Section 3.7.

3.3 Observations

3.3.1 Archival Data

We use publicly available archival spectra from the Keck Observatory Archive[1] taken with the HIRES instrument on the Keck I Telescope (Vogt et al. 1994). In total, we analyze 40 individual spectra of 8 RGB stars in the M15 globular cluster. Observations span the wavelength range 3160–8370 Å and provide nominal resolving powers ($R = \lambda/\delta\lambda$) from 37,500 to 86,600. In addition to archival Keck/HIRES spectroscopy, we also employ Gaia DR3 photometry (Gaia Collaboration et al. 2022) to better constrain stellar fundamental parameters (e.g., $T_{\rm eff}$, $\log g$). We apply extinction corrections to this photometry using the Schlafly et al. (2011) dust map, the Gaia extinction coefficients from Collaboration et al. (2018), and adopting $R_V = 3.1$.

Table 3.1 provides a list of the stars analyzed in this work, and Table 3.2 provides a summary of the spectroscopic observations. Figure 3.2 shows the location of these stars on the Gaia DR3 color-magnitude diagram of probable M15 members as determined by Vasiliev et al. (2021). All of the stars considered in this study reside on the upper portion of the RGB.

3.3.2 Data Reduction

All archival data were reduced using version 1.3.1 of the `PypeIt` data reduction pipeline (Prochaska et al. 2020)[2]. At the time of reduction, `PypeIt` did not support Keck/HIRES data, so a few minor alterations to the reduction code were necessary, which we summarize below.

Echelle orders were manually identified for each observational setup by matching preliminary wavelength solutions to the HIRES Echelle Format Simulator[3]. Spectral orders were discarded if $\gtrsim$50% of their extent fell off or between detectors—no attempt was made to stitch together orders that spanned multiple detectors. As a result, order 67 (5280–5370 Å) was discarded from the C147Hr and C316Hr programs.

Wavelength calibrations were performed using the `PypeIt` "reidentify" method, in which the observed arc spectra are cross-correlated against archival arc spectra. Appropriate archival spectra for each setup were adopted from the `MAKEE` data reduction package[4].

[1] `https://koa.ipac.caltech.edu/`

[2] `https://pypeit.readthedocs.io`

[3] `https://www2.keck.hawaii.edu/realpublic/inst/hires/HIRES-efs-master/efs.html`

[4] `https://sites.astro.caltech.edu/~tb/makee/`

Kustner ID (Kustner 1921)	2MASS ID (Skrutskie et al. 2006)	Other IDs	$m_{G,0}$	$G_{BP,0} - G_{RP,0}$
K341	J21295492+1213225	CBG 4099	12.39	1.59
K386	J21295562+1210455	CBG 40825	12.32	1.62
K431	J21295618+1212337	S1	12.62	1.53
K462	J21295666+1209463	...	12.45	1.58
K583	J21295856+1209214	...	12.32	1.61
K731	J21300053+1211369	ARP I-63, CBG 45062	13.99	1.29
K934	J21300480+1211469	ARP I-62	14.17	1.26
K969	J21300637+1206592	S8	13.11	1.40

Table 3.1: M15 stars analyzed in this work. For brevity, we will refer to stars throughout this work using their Kustner IDs. Alternative identifiers are as follows: ARP = Arp (1955), CBG = Carretta, Bragaglia, Gratton, and Lucatello (2009), and S = Sandage (1970). G-band magnitudes and BP-RP colors are from Gaia DR3 and corrected for extinction (Gaia Collaboration et al. 2022).

Default `PypeIt` methods and algorithms were employed for bias subtraction, flat-fielding, flexure correction, cosmic ray rejection, sky subtraction, and object extraction. After extraction, the stellar spectra were velocity corrected into the Heliocentric reference frame using the default `astropy`[5] Solar System ephemeris. To minimize information loss, repeat observations of the same star are not stacked, but fit individually. A "stacked" measurement is obtained by combining the posteriors of fits to individual exposure using a hierarchical model (see Section 3.4.3.4).

We do not formally flux calibrate the 1D extracted spectra but rather fit for the pseudo-continuum simultaneously with its atmospheric parameters and elemental abundances (see Section 3.4.2.3). As a part of the pseudo-continuum fitting, we define a scaled blaze function for each order, which we extract from the combined flat-field calibration frame and scale to the flux of each observed spectral order.

In Figure 3.3, we present a sample order from one of the reduced archival observations. The scaled blaze function for the order is over-plotted in red, and the adopted observational masks (described in Section 3.3.3) are included as vertical shaded bands. A complete library of the reduced spectra analyzed in this work can be made available upon request.

3.3.3 Observational Masks

In Figure 3.3, we illustrate the three types of observational masks adopted to flag pixels with large observational artifacts or uncertainties and exclude them in our spectral fitting analysis. The telluric absorption mask (blue shaded regions), includes all pixels that contain

[5]https://www.astropy.org/

Kustner ID (Kustner 1921)	Wavelength Range (Å)	Resolution ($\lambda/\delta\lambda$)	Date (DD-MM-YYYY)	Program ID	Program PI	Exposures
K341	3650–5200	45,000	03-09-1997	U09H	R. Kraft	5×1800s
K386	3650–5200	45,000	04-09-1997	U09H	R. Kraft	7×1800s
K431	3840–8370	86,600	09-09-2011	C316Hr	E. Kirby	4×1770s
K431	3840–8370	86,600	17-09-2011	C316Hr	E. Kirby	1500s
K462	3650–5200	45,000	03-09-1997	U09H	R. Kraft	8×1800s
K583	3650–5200	45,000	04-09-1997	U09H	R. Kraft	6×1800s
K731	3840–8370	37,500	10-06-2008	C147Hr	J. Cohen	1000s
K731	3840–8370	37,500	11-06-2008	C147Hr	J. Cohen	1000s
K934	3840–8370	37,500	11-06-2008	C147Hr	J. Cohen	800s
K969	3840–8370	86,600	09-09-2011	C316Hr	E. Kirby	3×1725s, 3×1475s

Table 3.2: Summary of archival observations analyzed in this work. All raw data are available on the Keck Observatory Archive. Several archival HIRES observations of M15 stars are omitted from this study because they lack suitable flat-field exposures for `PypeIt` reductions or lack Gaia photometry.

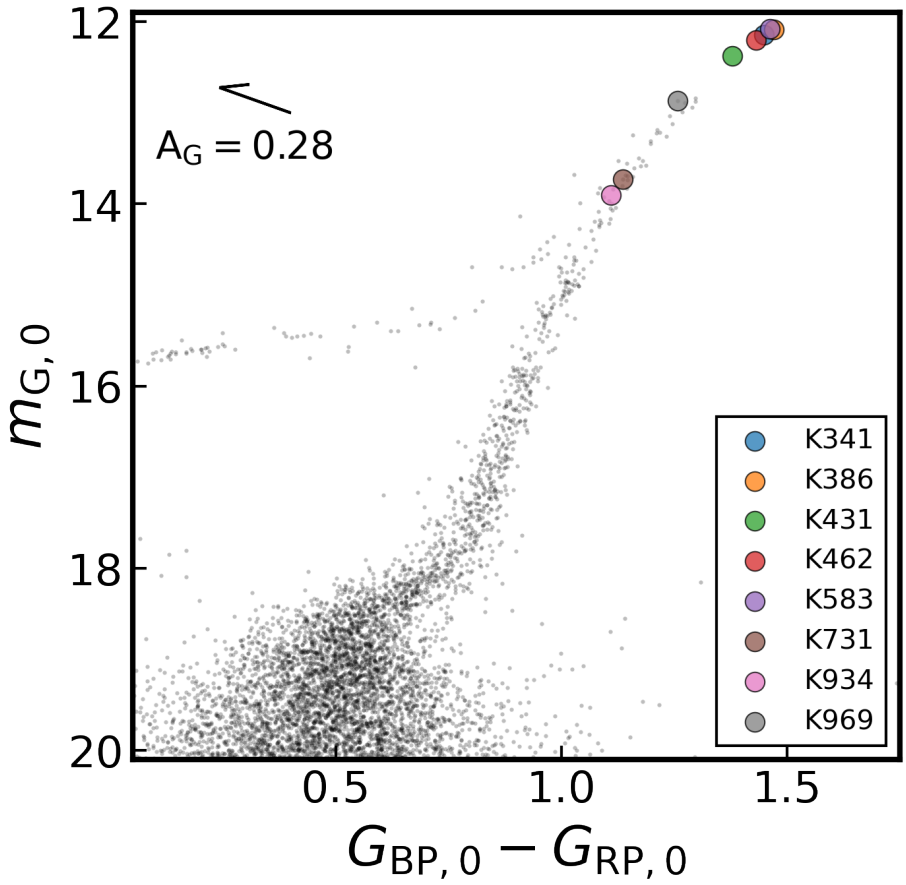


Figure 3.2: Gaia DR3 color-magnitude diagram of likely M15 members as identified by Vasiliev et al. (2021). Stars analyzed in this work are represented by filled circles, which are all located on the upper part of the RGB. The median extinction correction applied to the sample is denoted by the arrow in the upper left-hand corner of the figure.

strong telluric contamination, as identified in the "List of Telluric Lines" provided by `MAKEE`[6]. The detector boundary mask (gray shaded regions) includes the first 64 and last 128 pixels of every order in the C147Hr and C316Hr programs, which exhibit strongly non-linear response functions that bias polynomial fits to the spectral continuum[7]. Lastly, the bad pixel mask (purple shaded regions) includes all hot pixels, improperly subtracted sky lines, and cosmic rays as identified automatically with `PypeIt` or by visual inspection.

[6] `https://www2.keck.hawaii.edu/inst/common/makeewww/Atmosphere/atmabs.txt`

[7] The Older U09H program observations do not exhibit strong non-linear effects near the detector boundaries, so no detector boundary mask is necessary.

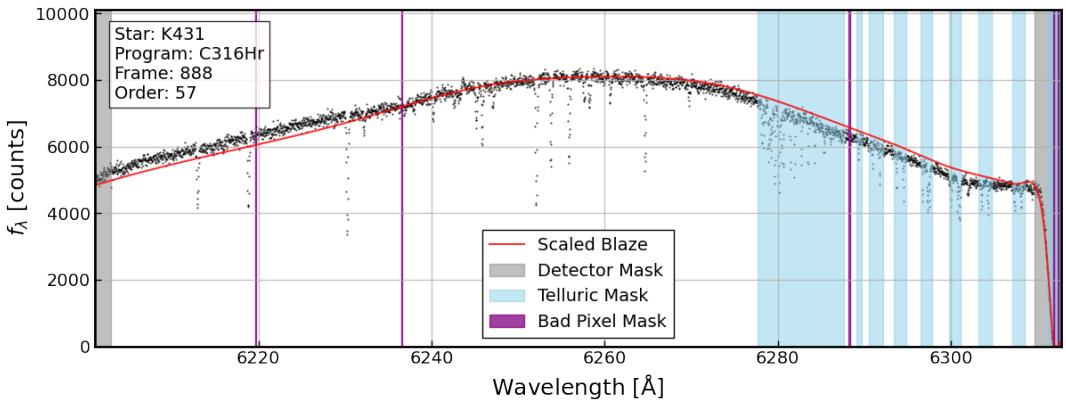


Figure 3.3: A sample order from one reduced archival observations (black points) illustrating the types of masks we apply to the data. The solid red line represents the scaled blaze function, which we use for the zeroth-order continuum determination. Deviations from the observed continuum are accounted for using a polynomial as described in Section 3.4.2.3. The gray, blue, and purple shaded regions represent the detector boundary mask, the telluric mask, and the bad pixel mask respectively. Pixels that lie within these observational masks are ignored in the spectral fitting analysis.

3.3.4 Post-Processing Observations

A primary goal of this paper is to self-consistently test the robustness of stellar spectroscopic label recovery as a function of spectral resolving power and S/N using real (as opposed to mock) data. Specifically, we consider stellar label recovery along two axes: i) as a function of resolution at fixed integration time and ii) as a function of S/N at fixed resolution. In order to satisfy these conditions using archival data from only one spectrograph, we apply several post-processing operations to the data (e.g., to degrade resolution or S/N), which we now describe.

3.3.4.1 Varying Resolution at Fixed Integration Time

Because the archival spectra are all taken at high resolution, testing stellar label recovery at lower resolution requires that we artificially degrade the resolving power of the archival spectrum and repeat our analysis at each resolution. We perform this degradation by convolving each archival spectrum to successively halved resolving powers down to $R \sim 2500$—a factor of 16–32 lower than the native instrumental resolution. The convolution of a sample order from one reduced archival spectrum is presented in Figure 3.4.

Here, and throughout this paper, we perform spectral convolutions assuming that the instrumental broadening kernel, $\mathcal{F}_v^{\rm inst}$, is well-described by a zero-mean Gaussian with con-

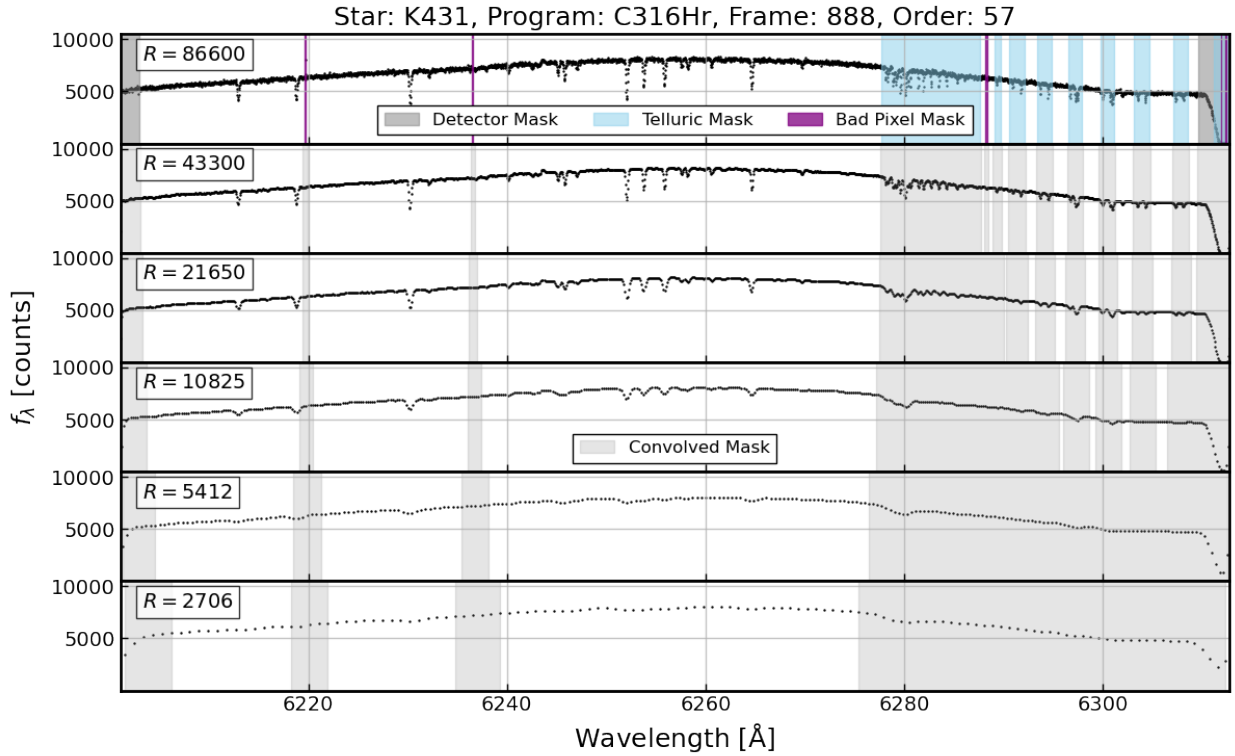


Figure 3.4: An illustration of the effects of varying spectral resolution on the observational masks using the same sample order and observational masks from Figure 3.3 (top). Lower panels depict the observed order convolved to lower resolutions by successive factors of 2. As the spectral resolving power decreases, the observational masks (light grey bands) grow to include pixels impacted by the broadening of masked features. The spectrum is also re-binned as it is convolved to lower resolution to maintain a constant $N_{\rm pix}$/FWHM. The S/N of the spectrum scales with $R^{-1/2}$ as a result of this re-binning.

stant width, $\sigma_{\rm inst} = 1/2.355R$, where R is the spectral resolving power of the instrumental configuration used in the observation. We also assume that R is constant as a function of wavelength though this is not strictly true in practice. Given the initial resolving power of an observation, R_0, we achieve the desired resolving power, R, by convolving each order of the initial spectrum with a Gaussian kernel of width

$$\sigma_{\rm inst} = \left[(2.355R)^{-2} - (2.355R_0)^{-2}\right]^{1/2}. \tag{3.1}$$

We perform these convolutions via multiplication of the spectrum and the broadening kernel in Fourier-space which increases computational efficiency and better preserves spectral information. An identical convolution is applied to the flux uncertainty of each order.

However, convolving observational data has several unavoidable consequences that must be handled properly for a self-consistent analysis. First, by convolving the spectra on their native wavelength grid results in spectra that are oversampled (i.e., $N_{\rm pix}/{\rm FWHM} \gtrsim 3$). For example, a spectrum with $N_{\rm pix}/{\rm FWHM} \sim 3$ at $R = 40{,}000$ would have $N_{\rm pix}/{\rm FWHM} \sim 6$ at $R = 20{,}000$ and $N_{\rm pix}/{\rm FWHM} \sim 48$ at $R = 2500$, which is unrealistically over-sampled. Instead, to more appropriately emulate low resolution observations, we downsample the spectra by a factor of R_0/R to maintain constant $N_{\rm samp} \sim 3\,{\rm pixels/FWHM}$. This downsampling is performed using the using the `SpectRes`[8] Python package (Carnall 2017). Importantly, `SpectRes` re-bins the spectra and its uncertainties in a manner that conserves flux, resulting in the S/N of the convolved and downsampled spectra increasing as the resolution is decreased according to ${\rm S/N} \propto R^{-1/2}$.

Second, convolution also complicates the use of the observational masks described in Section 3.3.3. The convolution kernel not only broadens spectral features, but also sky lines, detector artifacts, and bad pixels, causing them to "spill out" from the existing masks. Our solution for this is to treat our masks as binary arrays with 0's corresponding to masked pixels and 1's corresponding to unmasked pixels. We then broaden and interpolate these masks in the same manner as the observed spectrum and expand them to include any pixels where the convolved mask is <0.99—that is, any region where a masked pixel contributes >1% of its flux. For bad pixels with extremely outlying values, this can still lead to substantial contributions to unmasked pixels. To mitigate this, we replace all bad pixels with the mean value of the nearest non-masked pixel prior to convolution. Broadened observational masks are represented in Figure 3.4 by light grey vertical bands.

A third complication is potential edge effects. To illustrate the issue, consider the pathological example of a strong absorption line with a central wavelength that lies just outside the range of an observed spectral order. At high resolution, the absorption from this line might be completely excluded from the observed order. But at low resolution, the line might be broadened to the point where its wings bleed into the observed order. Convolving the observed spectrum artificially as we do in this study, would completely omit the contribution of this broadened line, introducing additional systematic error into the analysis. For spectra from the C147Hr and C316Hr programs, the detector boundary masks are sufficient to exclude any edge effects. For spectra from the U09H program, we implement a one pixel mask at each end of each order and proceed with the mask convolution procedure described above. This will similarly exclude any potential edge effects.

As a result of expanding the observational masks, a greater fraction of the spectrum is masked at lower resolution. For example, in the C316Hr observations ∼10% of the pixels are masked at $R \sim 80{,}000$ versus ∼25% at $R \sim 2500$, and in the U09H observations ∼1% of the pixels are masked at $R \sim 45{,}000$ versus ∼7% at $R \sim 2500$. While larger contamination from telluric lines is to be expected at lower resolution, it is not typically the case for cosmic rays, hot/dead pixels, and detector edge effects. This is a minor, but necessary, trade-off in our choice to use the same exposures at multiple resolutions. We believe the value in using

[8] `https://spectres.readthedocs.io/en/latest/`

real data (as opposed to synthetic spectra) greatly outweighs these minor complications.

3.3.4.2 Varying S/N at Fixed Resolution

The archival spectra was taken with specific science goals in mind, which translate to minimal S/N requirements (e.g., S/N $\gtrsim$ 40 pixel^{-1} at λ5000). This is illustrated in Figure 3.5, which presents the median S/N of each individual echelle order analyzed in this study.

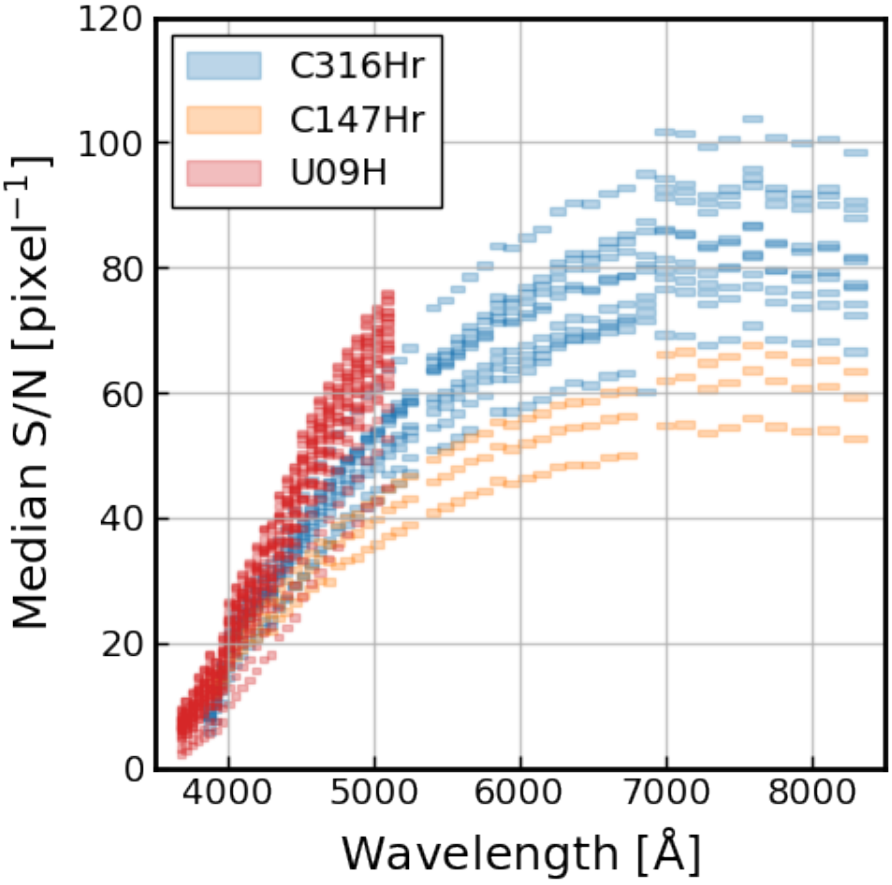


Figure 3.5: Median S/N per pixel of each echelle order in each exposure analyzed in this study before the quality of the data is degraded. The width of the horizontal bars represent the wavelength coverage spanned by the order. The colors denote the observing programs outlined in Table 3.2.

In order to test the robustness of stellar label recovery as a function of S/N, we add artificial white noise to the reduced spectra in order to decrease the median S/N by factors of 2 down to S/N $\sim$ 5 pixel^{-1}. For this analysis, we consider only spectra convolved to

$R \sim$ 10,000 as we expect the results at moderately lower and higher resolutions to be similar.

For a reduced spectrum, D_0, with flux errors, σ_{D_0}, reported from the `PypeIt` reduction pipeline, we add Gaussian noise to the spectrum as follows:

$$D = D_0 + \mathcal{N}(D_0,\ \sigma), \tag{3.2}$$

where σ satisfies the condition that the resulting flux uncertainties,

$$\sigma_D = \sqrt{(\sigma)^2 + (\sigma_{D_0})^2}, \tag{3.3}$$

yield the desired median S/N,

$$\mathrm{Med(S/N)} = \mathrm{Med}\left(\frac{D}{\sigma_D}\right). \tag{3.4}$$

Figure 3.6 illustrates an example spectral order degraded to lower S/N values.

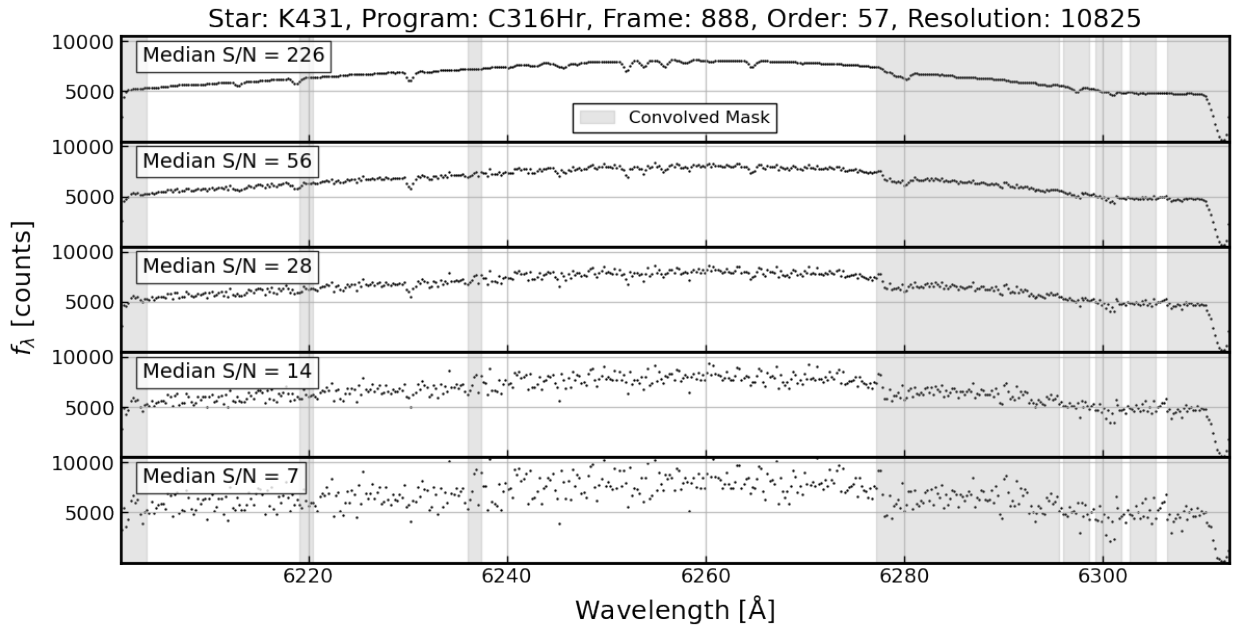


Figure 3.6: The same sample order and observational mask from Figure 3.3 convolved to $R \sim$ 10,000 (top). The lower panels depict the observed order noised up by factors of 4, 8, 16, and 32 respectively. While very little information appears to remain at the lowest S/N, this is only a small portion of the full stellar spectrum.

3.4 Spectral Fitting Analysis

In this section, we describe our framework for fitting stellar spectra. The overarching structure of our analysis (and this section) is as follows. We begin in Section 3.4.1 by generating a normalized synthetic spectrum from a set of stellar labels using `the Payne`, a fast neural-network spectral emulator. Then in Section 3.4.2, this model spectrum is forward-modelled into the observational domain given additional parameters describing various spectral broadening effects, the stellar radial velocity, and the spectral continuum. Lastly in Section 3.4.3, the model spectrum is compared directly to the observed spectrum on a pixel-by-pixel level and a posterior probability is calculated. The best-fit stellar (and nuisance) parameters are found by maximizing the posterior using both optimization techniques and Markov chain Monte Carlo (MCMC) sampling. Throughout this section, we borrow much of our notation from Section 2 of Czekala et al. (2015), which we found to be a clear, illustrative, and mathematically rigorous presentation of forward-modelling stellar spectra. The code used to perform the described spectral analysis is made public in the `PayneOptuna` Github repository[9].

3.4.1 Generating Model Spectra with The Payne

At the core of most full-spectrum fitting techniques is a model that can generate a realistic stellar spectrum, f_λ, from a set of stellar parameters or labels, θ_*. Because generating $f_\lambda(\theta_*)$ on the fly from stellar atmosphere and radiative transfer codes is computationally prohibitive, we employ `the Payne` (Ting, Conroy, Rix, and Cargile 2019), a powerful tool for spectral emulation that has been successfully used in a number of spectroscopic studies (e.g., El-Badry, Rix, et al. 2018; Kovalev et al. 2019; Ting, Conroy, Rix, and Cargile 2019; Xiang, Ting, et al. 2019; Straumit et al. 2022; Xiang, Rix, et al. 2022). At its core, `the Payne` is a fully-connected neural network that is trained to efficiently and accurately interpolate a high-dimensional grid of *ab initio* stellar spectra. Because `the Payne` is trained on synthetic spectra, it avoids confusing astrophysical correlation between elemental abundances (like bulk α-enhancements) with real spectroscopic abundance information (e.g., Ting, Rix, et al. 2017; Xiang, Ting, et al. 2019).

In short, we generate a grid of $\mathcal{O}(10^4)$ stellar labels, $\theta_* = \{T_{\rm eff},\ \log g,\ v_{\rm micro},\ [{\rm X/H}]\}$, where X includes 36 elements (C, N, O, Na, Mg, Al, Si, K, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Sr, Y, Zr, Ba, La, Ce, Pr, Nd, Sm, Eu, Gd, Dy, Ho, Er, Os, and Th).[10] For each θ_*, we compute a continuum-normalized $R = 300{,}000$ *ab initio* spectrum with the 1D LTE stellar atmosphere and radiative transfer codes, `ATLAS12` and `SYNTHE` (Kurucz 1970; Kurucz and Avrett 1981; Kurucz 1993, 2013, 2017). These spectra are convolved and sub-sampled down to the highest spectral resolution and wavelength sampling present in our archival data ($R = 86{,}600$; ${\rm d}v = 1.17\,{\rm km\,s^{-1}}$ pixel^{-1}). `The Payne` is then trained on this grid of convolved

[9] https://github.com/NathanSandford/PayneOptuna

[10] However, we ultimately exclude Ga, Ho, and Os from our analysis due to lack of constraining spectral information.

spectra. A detailed technical description of `the Payne` architecture, training, and accuracy is provided in Appendix 3.8.1.

3.4.1.1 Model Uncertainties

In addition to the flux uncertainty of the observations, we also incorporate the flux uncertainty of our models. Specifically, we include three sources of model uncertainty: interpolation errors of `the Payne`, NLTE effects, and saturated lines. These are illustrated in Figure 3.7.

The first source of uncertainty captures how well our spectral model, $f_\lambda(\theta_*)$, can reproduce the *ab initio* spectra generated directly with `ATLAS12` and `SYNTHE`. Even a well-trained model has non-zero interpolation errors, which can vary as a function of wavelength and stellar labels. We adopt the median interpolation error (MIE), $\sigma_{\rm MIE}$, as the fundamental flux uncertainty of our model spectra (gray line in Figure 3.7). On the whole, $\sigma_{\rm MIE}$ is small—the median value across the entire spectrum is $\sim 4 \times 10^{-4}$. There are portions of the spectrum, however, that exhibit larger interpolation errors—roughly 1% of the model spectrum has $\sigma_{\rm MIE} \gtrsim 10^{-2}$. This is predominantly the case for strong lines and complicated molecular features like the CH molecular band at $\lambda 4300$ seen in Figure 3.7. For simplicity, we assume that $\sigma_{\rm MIE}$ is independent of stellar labels, though we find it to be larger for spectra with [Fe/H] > -2. Fortuately, the stars considered in this study are all found to have [Fe/H] $\lesssim -2.4$. For more details on the MIE, see App 3.8.1.4.

The second source of uncertainty is introduced by the 1D LTE assumptions of our model atmosphere and radiative transfer codes. Many stellar absorption lines are known to be sensitive to NLTE effects, which will be poorly modelled by $f_\lambda(\theta_*)$ (e.g., Asplund 2005, and references therein). Instead of simply masking out NLTE lines as is standard in 1D LTE analyses, we attempt to mitigate the impact of our 1D LTE assumptions by including an additional source of uncertainty, $\sigma_{\rm NLTE}$. We define this to be the difference in normalized flux expected from LTE and NLTE treatments:

$$\sigma_{\rm NLTE} = |f_{\lambda,\rm LTE} - f_{\lambda,\rm NLTE}| \tag{3.5}$$

(blue line in Figure 3.7). To calculate $\sigma_{\rm NLTE}$, we use the NLTE Abundance Correction tool[11] developed and maintained by M. Kovalëv, which includes NLTE effects for lines of O, Mg, Si, Ca, Ti, Cr, Mn, Fe, and Co as calculated by Mashonkina, Korn, et al. (2007), Bergemann and Gehren (2008), Bergemann and Cescutti (2010), Bergemann, Pickering, et al. (2010), Bergemann (2011), Bergemann, Lind, et al. (2012), Bergemann, Kudritzki, et al. (2013), and Sitnova et al. (2013), and Bergemann, Collet, et al. (2017). This is, of course, a far from complete accounting of the NLTE effects present in real spectra, but should nevertheless substantially reduce the impact of the LTE assumptions made throughout this study.

Third and finally, a few strong spectral features, notably the Ca H&K and the hydrogen Balmer lines, in our observations are strongly saturated and thus poorly modelled by $f_\lambda(\theta_*)$.

[11] `http://nlte.mpia.de/gui-siuAC_secE.php`

We mask these lines out with

$$\sigma_{\rm sat} = \begin{cases} 1, & |\lambda - \lambda_0| < \delta\lambda \\ 0, & \text{otherwise} \end{cases}, \tag{3.6}$$

where λ_0 is the line center of the saturated feature and $\delta\lambda$ is chosen to generously encompass the width of the line (yellow region in Figure 3.7). We provide $\lambda_{\rm center}$ and $\delta\lambda$ for these lines in Table 3.3.

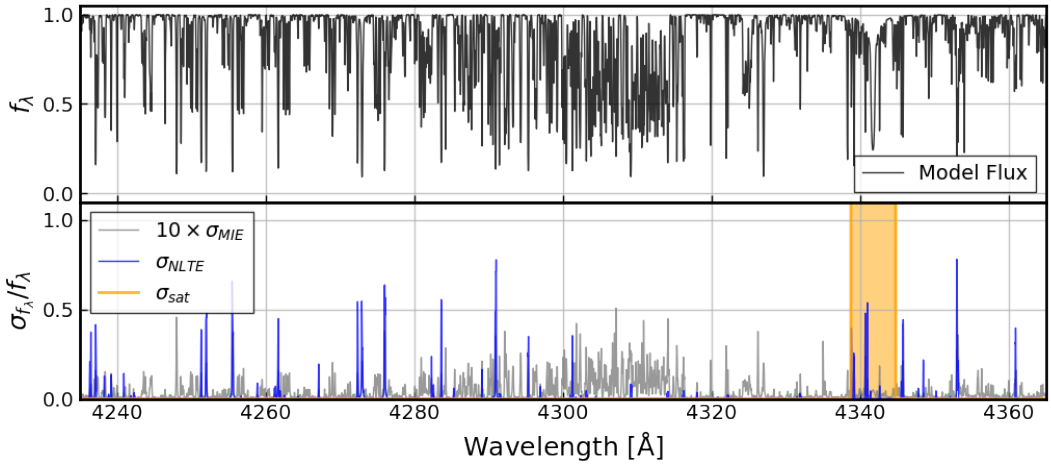


Figure 3.7: A portion of a synthetic spectrum generated with `the Payne` (top) and its fractional flux uncertainty (bottom). The total model uncertainty is the quadrature sum of the three components displayed here: the MIE of `the Payne` (gray), NLTE effects (blue), and saturated lines (yellow). For visibility, the MIE has been inflated by a factor of 10 in this figure. The saturated line masked from this portion of the spectrum is the Hγ line at λ4341.7.

Under the reasonable assumption that these three sources of uncertainty are largely uncorrelated, the total model uncertainty is then their quadrature sum,

$$\sigma_{f_\lambda} = \sqrt{\sigma^2_{\rm MIE} + \sigma^2_{\rm NLTE} + \sigma^2_{\rm sat}}. \tag{3.7}$$

3.4.2 Forward Modelling

By construction, `the Payne` emulates the normalized spectra generated by the `ATLAS12` and `SYNTHE` models and, as is, omits important observational and instrumental effects. As a result, it is necessary to incorporate these effects via forward modelling of the synthetic

Line	λ_0 [Å]	$\delta\lambda$ [Å]
Ca H	3969.6	20
Ca K	3934.8	20
Hα	6564.6	3
Hβ	4862.7	3
Hγ	4341.7	3
Hδ	4102.9	3
Hϵ	3971.2	3
Hζ	3890.2	3
Hη	3836.5	3

Table 3.3: List of saturated lines that are masked in our analysis. All line centers are given in vacuum wavelengths.

spectra before it can be compared directly to real data. This forward modelling is done in three steps, which are described below. In each step, the model flux and the model flux uncertainties are operated on identically.

3.4.2.1 Radial Velocity and Broadening Kernels

In the first forward modelling step, we account for observational and instrumental effects that alter the stellar spectrum along its wavelength dimension. We implement broadening from two sources: the instrumental line spread function (LSF) and macroturbulent motion in the stellar photosphere. We also Doppler shift the spectrum according to the radial velocity of the star. Each of these can be characterized by a kernel, $\mathcal{F}_v$, that modifies the line-of-sight velocity distribution function of $f_\lambda(\theta_*)$.

For the instrumental broadening kernel, $\mathcal{F}_v^{\rm inst}$, we adopt a zero-mean Gaussian with constant-width parameterized by the instrumental resolving power, R, as previously described in Section 3.3.4.1. For computational efficiency, we also adopt a zero-mean Gaussian for the macroturbulent broadening kernel, $\mathcal{F}_v^{\rm turb}$, which we parameterize with the macroturbulent velocity, $v_{\rm macro}$[12]. Lastly, the Doppler shift is implemented with a delta function kernel, $\mathcal{F}_v^{\rm dop} = \delta(v - v_r)$, centered at the radial velocity, v_r, of the star.

$f_\lambda(\theta_*)$ and σ_{f_λ} are modified via a convolution with these kernels in velocity space, i.e.,

$$f_\lambda(\theta_*, \theta_v) = f_\lambda(\theta_*) * \mathcal{F}_v^{\rm dop} * \mathcal{F}_v^{\rm inst} * \mathcal{F}_v^{\rm turb} \tag{3.8}$$

and

$$\sigma_{f_\lambda}(\theta_v) = \sigma_{f_\lambda} * \mathcal{F}_v^{\rm dop} * \mathcal{F}_v^{\rm inst} * \mathcal{F}_v^{\rm turb} \tag{3.9}$$

[12]The "radial-tangential" model described in Equation 17.15 of Gray (2021) would be more accurate, but adopting a Gaussian kernel for both the instrumental and macroturbulent broadening kernels allows the two broadening steps to be easily combined.

respectively, where $\theta_v = \{R,\ v_{\rm macro},\ v_r\}$ includes the additional model parameters characterizing each kernel. These convolutions are performed by multiplying the spectra with the kernels in Fourier space.

We note two velocity-related convolutions that are excluded from this post-processing: microturbulent broadening and rotational broadening. Microturbulent broadening is excluded here because it is already incorporated into the model spectra generation as part of θ_* input to `SYNTHE`. Rotational velocity is excluded because the stars in our sample are most likely slow-rotating low-mass giant stars, whose spectra do not typically exhibit substantial rotational broadening (Carlberg et al. 2011).

In practice, we hold R fixed as we expect it to be very degenerate with measurements of $v_{\rm macro}$ and other stellar parameters, especially at low resolution. Moreover, R is typically a well-known characteristic of the spectroscopic observing configuration.

3.4.2.2 Wavelength Interpolation

At this point in the post-processing, the convolved and Doppler shifted model spectrum is highly oversampled compared to real observations. It is thus necessary to resample the model flux and its uncertainties onto the discrete wavelengths corresponding to each pixel of each order, o, in the observed spectrum, i.e.,

$$f_\lambda(\theta_*, \theta_v) \mapsto M_o(\theta_*, \theta_v) \tag{3.10}$$

and

$$\sigma_{f_\lambda}(\theta_*, \theta_v) \mapsto \sigma_{M_o}(\theta_*, \theta_v). \tag{3.11}$$

This resampling is performed via linear interpolation of $f_\lambda(\theta_*, \theta_v)$ and σ_{f_λ}.

3.4.2.3 Stellar Continuum and Detector Response

This forward modelling step addresses the fact that the model we have established thus far, $M_o(\theta_*, \theta_v)$, generates a normalized stellar spectrum. However, the shape of the observed spectra is that of the stellar continuum modulated by the instrumental response function. To incorporate a realistic continuum into the normalized model spectra, we apply a two-part continuum scaling. The first operation captures the spectrograph response function both within and across spectral orders as well as the large-scale spectral energy distribution of the star. To do this, we multiply each order of the model spectrum by the blaze function of the order, which we have extracted from the combined flat-field calibration frame and scaled to the observations (see Section 3.3.2).

To account for any deviations that remain, we multiply each order of the model spectrum by a low-order nth degree polynomial function, P_o. This polynomial function can be described by a set of $n+1$ coefficients for each order, $\phi_{\rm P} = \{c_{o,n}\}$. To improve the stability of this correction while fitting, we evaluate each polynomial not as a function of λ but of a scaled wavelength

$$\lambda'_o = \frac{2a}{\lambda_{o,\max} - \lambda_{o,\min}} \left(\lambda_o - \lambda_{o,\rm mean}\right), \tag{3.12}$$

where $\lambda_{o,\max}$, $\lambda_{o,\min}$, and $\lambda_{o,\mathrm{mean}}$ are the maximum, minimum, and mean wavelengths of each order respectively, and $-a < \lambda'_o < a$.

The resulting continuum-corrected and fully post-processed spectrum can then be written as:

$$M(\Theta) = \{M_o(\theta_*, \theta_v) \times B_o P_o\} \tag{3.13}$$

$$= \left\{ M_o(\theta_*, \theta_v) \times B_o \sum_{n=0}^{N_{\mathrm{deg}}} c_{o,n} (\lambda'_o)^n \right\}, \tag{3.14}$$

where $\Theta = \{\theta_*, \theta_v, \phi_{\mathrm{P}}\}$ represents all physical and nuisance parameters of the model.

In summary, each model spectrum is described by 39 stellar labels (3 atmospheric parameters and 36 elemental abundances), 3 labels describing spectral broadening and Doppler shift (R, v_{macro}, and v_r), and $N_{\mathrm{ord}} \times (N_{\mathrm{deg}} + 1)$ continuum coefficients, where N_{ord} is the number of orders in the spectrum and N_{deg} is the degree of the continuum correction polynomial. We find $N_{\mathrm{deg}} = 4$ is suitable for most HIRES observations.

3.4.3 Model Evaluation and Spectral Fitting

With spectral model $M(\Theta)$ now defined (Equation 3.14), we can infer the physical (and nuisance) parameters, Θ, that best reproduce an observed spectrum, D, by maximizing the posterior probability

$$\ln P(\Theta|D) = \ln L(D|\Theta) + \ln \Pi(\Theta), \tag{3.15}$$

where $\ln L(D|\Theta)$ is the log-likelihood of the data given the model parameters and $\ln \Pi(\Theta)$ is the log-prior on the model parameter. For each observed spectrum, we first use an optimization algorithm to recover the maximum a posteriori value of Θ. Then we use MCMC to sample directly from $P(\Theta|D)$, validating the results of the optimizer and providing uncertainties and covariances for the recovered labels. A technical description of each method is provided in Appendix 3.8.2.

For both optimization and MCMC sampling, we adopt a Gaussian log-likelihood function for $\ln L(D|\Theta)$ in Equation 3.15:

$$\ln L(D|M) = -\frac{1}{2} \sum^{N_{\mathrm{ord}}} \sum^{N_{\mathrm{pix}}} \left[\ln(2\pi\sigma_{\mathrm{tot}}^2) + (R/2\sigma_{\mathrm{tot}})^2 \right], \tag{3.16}$$

where

$$R \equiv R(\Theta) \equiv D - M(\Theta) \tag{3.17}$$

is the residual spectrum and

$$\sigma_{\mathrm{tot}} = \sqrt{\sigma_M^2 + \sigma_D^2} \tag{3.18}$$

is the combined flux uncertainty of the model and the data. The total log-likelihood is the summation of the individual pixel log-likelihoods over all spectral orders excluding those pixels ignored by the observational masks.

3.4.3.1 Fitting $T_{\rm eff}$ and $\log g$

In practice, $T_{\rm eff}$ and $\log g$ are often determined independent of the spectral analysis using photometry. In most cases these photometrically-determined values are held fixed or coarsely iterated over during the abundance determination (e.g., Kirby, Guhathakurta, Simon, et al. 2010). This approach is frequently taken for 1D LTE analysis of low-metallicity RGB stars where "overionization" departures from LTE become increasingly important (e.g., Asplund 2005, and references therein). We find, as have previous studies (e.g., Sneden, Johnson, et al. 2000; Sobeck, Ivans, et al. 2006), that attempting to fit $T_{\rm eff}$ and $\log g$ from spectroscopy alone frequently results in surface gravities that are >0.3 dex too small or stars occupying completely unphysical parts of the Kiel diagram.

Here, we recover $T_{\rm eff}$ and $\log g$ deterministically and simultaneously with the spectral analysis by interpolating MIST isochrones using the extinction-corrected Gaia photometry and [Fe/H] abundance of the star. That is

$$T_{\rm eff}, \log g = f_{\rm Iso}(m_{\rm G,0}, G_{\rm BP,0} - G_{\rm RP,0}, [{\rm Fe/H}]), \tag{3.19}$$

where $f_{\rm Iso}$ is the interpolation function for the MIST isochrone. To convert from apparent to absolute magnitudes, we adopt a distance modulus to M15 of $\mu_{\rm M15} = 10.71$ from Baumgardt et al. (2021). Because [Fe/H] is itself a free parameter, $T_{\rm eff}$ and $\log g$ are updated iteratively with each step of the optimizer and MCMC walker. This is similar to, though less sophisticated than the techniques employed in the `MINESWEEPER` spectral fitting code (Cargile et al. 2020).

3.4.3.2 Priors

With a few exceptions, we adopt the same priors when optimizing and sampling $P(\Theta|D)$. These priors are specified below. The total log-prior is the sum of the log-priors for each label, $\Pi(\Theta) = \Pi(T_{\rm eff}) + \Pi(\log g) + ... + \Pi(c_{n,o})$.

As described in Section 3.4.3.1, Gaia photometry is used to essentially impose a delta-function prior on $T_{\rm eff}$ and $\log g$ given $\mu_{\rm M15}$, $m_{\rm G,0}$, $\rm G_{BP,0} - G_{RP,0}$, and [Fe/H]. For the remaining stellar labels ($v_t extmicro$ and all elemental abundances), we adopt uniform priors over the range of values included in the spectral training grid (see Table 3.7):

$$\begin{aligned}
v_{\rm micro} &\sim \mathcal{U}(1.2 \text{ km s}^{-1},\ 2.5 \text{ km s}^{-1}) \\
[{\rm Fe/H}] &\sim \mathcal{U}(-4.00,\ -1.00) \\
[{\rm X_1/Fe}] &\sim \mathcal{U}(-1.00,\ 1.00) \\
[{\rm X_2/Fe}] &\sim \mathcal{U}(-0.50,\ 0.50) \\
[{\rm X_3/Fe}] &\sim \mathcal{U}(-0.25,\ 1.00),
\end{aligned}$$

where $\rm X_1 =$ C, N, and O; $\rm X_2 =$ Na, Sc, V, Cr, Mn, Co, Ni, Cu, Zn, Ga, Sr, Y, Zr, Ba, and La; and $\rm X_3 =$ Mg, Al, Si, K, Ca, Ti, Ce, Pr, Nd, Sm, Eu, Gd, Dy, Ho, Er, Os, and Th.

Though the resolving power, R, is a parameter in our model, we simply adopt the observatory-provided resolutions (and subsequent artificial reductions). This effectively imposes a delta-function prior on R,

$$R_{\rm inst} \sim \delta(R_{\rm obs}).$$

We impose a uniform prior on the log macroturbulent velocity, $\log_{10} v_{\rm macro}$, of from -1.0 to 1.3

$$\log_{10} v_{\rm macro} \sim \mathcal{U}(-1.0,\ 1.3),$$

which is equivalent to bounding $v_{\rm macro}$ in linear units from 0.1 to $20\,{\rm km\,s^{-1}}$. We adopt a broad uniform prior on the radial velocity from -300 to $300\,{\rm km\,s^{-1}}$:

$$\mathcal{U}(-300\ {\rm km\ s^{-1}},\ 300\ {\rm km\ s^{-1}}).$$

Because it is difficult to predict the appropriate range of values for the continuum polynomial coefficients, $c_{n,o}$, a priori, we adopt infinitely broad uniform priors on $c_{n,o}$ during optimization. Unfortunately, the large number of coefficients ($n \times N_{\rm ord}$) make including all $c_{n,o}$ as free parameters in the MCMC sampling computationally infeasible. Future work with Hamiltonian Monte Carlo or nested sampling methods may eventually make this tractable, but for now we fix all $c_{n,o}$ to the best fit optimization values with a delta function prior:

$$c_{o,n} \sim \begin{cases} \mathcal{U}(-\infty, \infty) & \text{Optimizer} \\ \delta\left(c_{o,n}^{(\rm Opt)}\right) & \text{MCMC} \end{cases}.$$

3.4.3.3 Reparameterization

To aid in the optimization and sampling of our posteriors, we find it advantageous to reparameterize a subset of our model parameters so that they share a similar dynamic range. Instead of fitting $v_{\rm macro}$ in linear units, we fit for $\log v_{\rm macro}$. The radial velocity, v_r, is scaled by a factor of 100 so that it has units of $100\,{\rm km\,s^{-1}}$. The stellar labels, θ_*, are scaled in the same manner as they are for the training of the `the Payne` to be between -0.5 and 0.5 (see Appendix 3.8.1). The priors for these reparameterized labels are transformed accordingly.

3.4.3.4 Fitting to Multiple Exposures

There are several approaches to handling the extra constraining power enabled by multiple exposures of the same star. The simplest and most common approach involves co-adding the spectra from individual exposures to create a "stacked" spectrum with a higher S/N than from the individual exposures. This approach is limited, however, in that it hides potential observational systematics introduced at the inter-exposure level—it is impossible to say how each exposure impacts the stacked fit.

A second approach and the one we adopt in this study is to treat each exposure of the same star as an independent observation of that star. The joint log-likelihood for the $N_{\rm exp}$ exposures is then just the sum of the log-likelihood for each individual exposure,

$$\ln L(D|\Theta) = \sum^{N_{\rm exp}} \ln L(D_i|\Theta), \tag{3.20}$$

and the "stacked" posterior of the multiple exposures is

$$\ln P(\Theta|D) = \ln \Pi(\Theta) + \sum^{N_{\rm exp}} \ln L(D_i|\Theta), \tag{3.21}$$

where $\ln \Pi(\Theta)$ are the log-priors described in Section 3.4.3.2.

While we can calculate the joint likelihood by fitting all exposures simultaneously, we choose to construct it after first sampling the posteriors of the individual exposure fits. We then fit these marginalized posteriors assuming they are well-described by 1-dimensional Gaussian distributions truncated at the bounds of the uniform priors. With functional forms of the posterior distributions in hand, we convert them into likelihood functions (a task made trivial by the use of uniform priors), and combine them into a joint likelihood function. Re-introducing priors results in the stacked posterior distribution function given in Equation 3.21, which we also fit assuming 1-dimensional truncated Gaussian distributions. We take the mean and standard deviation of these distributions as the best-fit value and 1σ statistical uncertainty of the stellar label except when the best-fit value is $< 1\sigma$ from the uniform priors bounds adopted in Section 3.4.3.2. In these instances, we instead adopt the 95% upper/lower limit in lieu of the mean and standard deviation.

The left panel of Figure 3.8 illustrates an example stacked posterior for [Fe/H] (black curve) that is recovered when the five individual exposure posteriors (colored curves) are combined. The right panel illustrates the same for [N/Fe] and demonstrates the importance of using truncated distributions.

3.5 Results

In this section, we present the results of our spectral fitting. We begin in Section 3.5.1 with the recovery of stellar labels as a function of resolving power and conclude in Section 3.5.2 with the recovery of stellar labels as a function of S/N. For a comparison of the stellar labels we measure from un-altered (i.e., default resolution and S/N) spectra to literature values, see Appendix 3.8.3.

3.5.1 Label Recovery as a Function of Resolution

At each resolution and for each star in our sample, we calculate the change in stellar labels, $\delta\theta$, relative to the recovered labels at the highest available resolution for that star

$$\delta\theta_R = \theta_R - \theta_{R_0}. \tag{3.22}$$

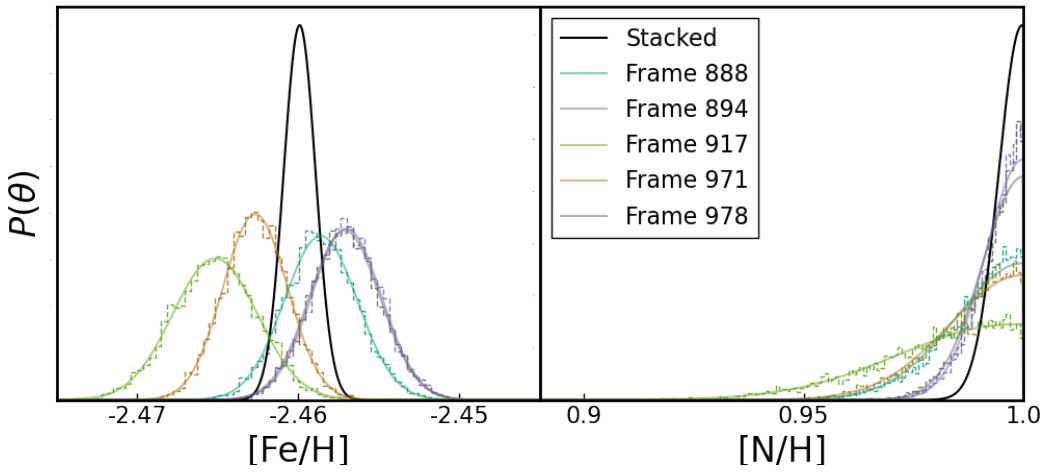


Figure 3.8: Marginalized posteriors for [Fe/H] (left) and [N/Fe] (right) for K431 observed in the C316Hr program at the degraded resolution of $R \sim 20{,}000$. Posterior samples and the best fit truncated normal distribution for the 5 individual exposures are plotted in the thin colored dashed histograms and solid curves respectively. The stacked posterior recovered when combining the individual likelihoods is plotted in the thick black line. In the case of [N/Fe], the best fit value is at the boundary of our priors (set by the extent of our training grid), necessitating the use of a truncated distribution.

Taken together, the trends of $\delta\theta$ versus resolution for our sample provide a coarse marginalization over the spectroscopic configurations (e.g., wavelength coverage) and stellar parameters (e.g., $T_{\rm eff}$, $\log g$, [Fe/H]) presented in this work. We summarize the collective trend for each stellar label with two quantities: a resolution-dependent systematic bias, $\Delta\theta$, and a resolution-dependent systematic uncertainty, $\sigma_{\rm syst}$. The systematic bias captures how much a stellar label is likely to be over/underestimated when measured at a lower resolution, while the systematic uncertainty captures the dispersion in $\delta\theta$ found across the programs and stars analyzed. We define $\Delta\theta$ to be the median and $\pm\sigma_{\rm syst}$ to be the 16th and 84th percentiles of $\delta\theta$ at each resolution. We omit from these calculations any poorly constrained fits for which the statistical uncertainty >0.5 dex.

As described in Section 3.4.3.4, some stellar label fits result in the recovery of upper or lower limits. While $\Delta\theta$ and $\sigma_{\rm syst}$ are robust to the presence of a few upper and lower limits, if a large enough fraction of the measurements of a stellar label at a given resolution are upper/lower limits, the 16th and 84th percentiles of $\delta\theta$—and thus $\pm\sigma_{\rm syst}$—may correspond to a limit. In these instances, the systematic uncertainty will be underestimated. In rare cases, $\Delta\theta$ may also correspond to a limit and be similarly underestimated.

For three elements, Ga, Ho, and Os, we struggle to measure reliable abundances even

at the highest default resolution due to their lack of informative absorption features in the observations. We exclude these three elements from our analysis moving forward.

In Figures 3.9–3.20, we present these systematic biases (solid black line) and uncertainties (gray shaded region) as a function of resolution. Solid red lines at the edge of the gray shaded region denote regions where the bias and/or uncertainty may be underestimated due to the limitations imposed by our training set and priors. For a few elements, measurements for each of the individual stars in the sample are included, color-coded by their observing program, to highlight instances where substantially different trends are exhibited from one archival dataset to the next. In these cases the U09H, C147Hr, and C316Hr programs are indicated with red squares, orange triangles, and blue circles respectively. Individual star measurements for all elements can be made available upon request. In Table 3.4, we provide $\Delta\theta$ and $\sigma_{\rm syst}$ for each element at resolutions of $R \sim 2500$, 5000, 10,000, 20,000, and 40,000. In the remainder of this section, we discuss the resolution-dependent recovery of each stellar label individually. For clarity, we organize our discussion of each element in groups loosely based on shared nucleosynthetic origin.

	$R \sim 2500$	$R \sim 5000$	$R \sim 10{,}000$	$R \sim 20{,}000$	$R \sim 40{,}000$
θ	$\Delta\theta \pm^{\sigma_{84\rm th}}_{\sigma_{16\rm th}}$	$\Delta\theta \pm^{\sigma_{84\rm th}}_{\sigma_{16\rm th}}$	$\Delta\theta \pm^{\sigma_{84\rm th}}_{\sigma_{16\rm th}}$	$\Delta\theta \pm^{\sigma_{84\rm th}}_{\sigma_{16\rm th}}$	$\Delta\theta \pm^{\sigma_{84\rm th}}_{\sigma_{16\rm th}}$
$T_{\rm eff}$ [K]	$0.78 \pm^{0.80}_{0.65}$	$0.68 \pm^{0.33}_{0.64}$	$0.31 \pm^{0.60}_{0.30}$	$0.16 \pm^{0.17}_{0.18}$	$0.17 \pm^{0.03}_{0.03}$
$\log g$	$-0.01 \pm^{0.00}_{0.00}$	$-0.00 \pm^{0.00}_{0.00}$	$-0.00 \pm^{0.00}_{0.00}$	$-0.00 \pm^{0.00}_{0.00}$	$-0.00 \pm^{0.00}_{0.00}$
$v_{\rm micro}$ [km s^{-1}]	$0.15 \pm^{0.08}_{0.07}$	$0.26 \pm^{0.08}_{0.07}$	$0.25 \pm^{0.11}_{0.10}$	$0.14 \pm^{0.03}_{0.02}$	$0.08 \pm^{0.00}_{0.00}$
$v_{\rm macro}$ [km s^{-1}]	$8.26^* \pm^{11.23^*}_{7.50}$	$5.96 \pm^{8.67}_{5.42}$	$2.52 \pm^{5.02}_{2.18}$	$0.97 \pm^{2.59}_{0.82}$	$0.63 \pm^{0.11}_{0.11}$
v_r [km s^{-1}]	$0.34 \pm^{0.37}_{1.03}$	$-0.09 \pm^{0.28}_{0.17}$	$-0.07 \pm^{0.21}_{0.06}$	$-0.03 \pm^{0.14}_{0.05}$	$-0.03 \pm^{0.01}_{0.01}$
[C/H]	$0.06 \pm^{0.02}_{0.10}$	$0.06 \pm^{0.03}_{0.10}$	$0.05 \pm^{0.01}_{0.01}$	$0.02 \pm^{0.01}_{0.01}$	$0.01 \pm^{0.00}_{0.00}$
[N/H]	$-0.00 \pm^{0.01^*}_{0.01}$	$0.00^* \pm^{0.01^*}_{0.00}$	$0.01^* \pm^{0.01^*}_{0.01}$	$0.00^* \pm^{0.02^*}_{0.00}$	$0.00^* \pm^{0.00^*}_{0.00}$
[O/H]	$-0.17 \pm^{0.17^*}_{0.51}$	$-0.03 \pm^{0.03^*}_{0.37}$	$-0.00^* \pm^{0.00^*}_{0.05}$	$-0.00^* \pm^{0.00^*}_{0.02}$	$0.00^* \pm^{0.00^*}_{0.00}$
[Na/H]	$-0.34 \pm^{0.22^*}_{0.26^*}$	$-0.24 \pm^{0.23^*}_{0.69^*}$	$-0.17 \pm^{0.17^*}_{0.63^*}$	$-0.07 \pm^{0.06^*}_{0.25}$	$-0.15 \pm^{0.02}_{0.02}$
[Mg/H]	$0.02 \pm^{0.16}_{0.10^*}$	$0.05 \pm^{0.05}_{0.11}$	$0.00 \pm^{0.06}_{0.08}$	$-0.01 \pm^{0.02}_{0.02^*}$	$-0.03 \pm^{0.03}_{0.03}$
[Al/H]	$0.13^* \pm^{0.38}_{0.36^*}$	$0.14 \pm^{0.43}_{0.14^*}$	$0.00 \pm^{0.48}_{0.18^*}$	$0.00 \pm^{0.12}_{0.32^*}$	$0.13^* \pm^{0.09}_{0.09^*}$
[Si/H]	$0.14 \pm^{0.22}_{0.05}$	$0.11 \pm^{0.19}_{0.02}$	$0.08 \pm^{0.22}_{0.04}$	$0.02 \pm^{0.19}_{0.04}$	$0.12 \pm^{0.06}_{0.06}$
[K/H]	...	...	...	$-0.00^* \pm^{0.00^*}_{0.22^*}$	$-0.54^* \pm^{0.37^*}_{0.37}$
[Ca/H]	$-0.02 \pm^{0.10}_{0.12}$	$0.02 \pm^{0.04}_{0.08}$	$-0.00 \pm^{0.05}_{0.03}$	$0.00 \pm^{0.02}_{0.02}$	$-0.00 \pm^{0.00}_{0.00}$
[Sc/H]	$-0.05 \pm^{0.13}_{0.10}$	$-0.04 \pm^{0.13}_{0.01}$	$0.01 \pm^{0.08}_{0.03}$	$0.02 \pm^{0.03}_{0.03}$	$0.03 \pm^{0.00}_{0.00}$
[Ti/H]	$-0.03 \pm^{0.09}_{0.03}$	$-0.01 \pm^{0.04}_{0.01}$	$0.00 \pm^{0.04}_{0.02}$	$0.01 \pm^{0.01}_{0.03}$	$0.01 \pm^{0.00}_{0.00}$
[V/H]	$0.07 \pm^{0.10}_{0.22}$	$0.06 \pm^{0.04}_{0.13}$	$0.02 \pm^{0.03}_{0.03}$	$0.02 \pm^{0.01}_{0.02}$	$0.01 \pm^{0.00}_{0.00}$
[Cr/H]	$0.13 \pm^{0.13}_{0.41^*}$	$0.07 \pm^{0.04}_{0.30}$	$0.04 \pm^{0.02}_{0.12}$	$0.02 \pm^{0.01}_{0.04}$	$0.02 \pm^{0.01}_{0.01}$
[Mn/H]	$0.00^* \pm^{0.00}_{0.00^*}$	$0.00^* \pm^{0.00}_{0.00^*}$	$0.04 \pm^{0.08}_{0.04^*}$	$0.00^* \pm^{0.02}_{0.00^*}$	$0.14^* \pm^{0.09}_{0.09^*}$
[Fe/H]	$-0.02 \pm^{0.01}_{0.01}$	$-0.02 \pm^{0.01}_{0.01}$	$-0.01 \pm^{0.01}_{0.01}$	$-0.00 \pm^{0.01}_{0.01}$	$-0.00 \pm^{0.00}_{0.00}$
[Co/H]	$-0.03 \pm^{0.07}_{0.31}$	$0.06 \pm^{0.02}_{0.19}$	$0.04 \pm^{0.02}_{0.10}$	$-0.00 \pm^{0.01}_{0.03}$	$-0.02 \pm^{0.01}_{0.01}$
[Ni/H]	$0.01 \pm^{0.09}_{0.07}$	$-0.00 \pm^{0.04}_{0.02}$	$-0.02 \pm^{0.04}_{0.04}$	$0.00 \pm^{0.01}_{0.02}$	$-0.01 \pm^{0.01}_{0.01}$
[Cu/H]	$0.28^* \pm^{0.57^*}_{0.28^*}$	$0.26^* \pm^{0.59^*}_{0.26^*}$	$0.10^* \pm^{0.26}_{0.10^*}$	$0.07^* \pm^{0.12}_{0.07^*}$	$0.02 \pm^{0.03}_{0.03}$
[Zn/H]	$0.32 \pm^{0.44^*}_{0.53}$	$0.30 \pm^{0.32}_{0.29}$	$0.14 \pm^{0.20}_{0.06}$	$0.07 \pm^{0.05}_{0.02}$	$0.06 \pm^{0.02}_{0.02}$
[Sr/H]	$-0.03 \pm^{0.06}_{0.19}$	$0.11 \pm^{0.06}_{0.11}$	$0.10 \pm^{0.07}_{0.14}$	$0.02 \pm^{0.09}_{0.07}$	$0.07 \pm^{0.03}_{0.03}$

[Y/H]	$0.07\pm^{0.19}_{0.13}$	$-0.01\pm^{0.14}_{0.03}$	$-0.04\pm^{0.07}_{0.02}$	$-0.02\pm^{0.03}_{0.01}$	$-0.01\pm^{0.01}_{0.01}$
[Zr/H]	$-0.12\pm^{0.08}_{0.37}$	$-0.02\pm^{0.06}_{0.15}$	$0.02\pm^{0.03}_{0.08}$	$0.02\pm^{0.02}_{0.04}$	$0.02\pm^{0.00}_{0.00}$
[Ba/H]	$-0.00^{*}\pm^{0.00^{*}}_{0.54}$	$-0.13\pm^{0.13^{*}}_{0.12}$	$-0.04^{*}\pm^{0.04^{*}}_{0.17}$	$-0.01^{*}\pm^{0.01^{*}}_{0.07}$	$-0.01^{*}\pm^{0.01^{*}}_{0.01}$
[La/H]	$0.01^{*}\pm^{0.17^{*}}_{0.01}$	$0.02^{*}\pm^{0.16^{*}}_{0.02}$	$0.03\pm^{0.06^{*}}_{0.03}$	$0.02\pm^{0.02^{*}}_{0.02}$	$0.02^{*}\pm^{0.01^{*}}_{0.01}$
[Ce/H]	$0.25\pm^{0.07}_{0.09}$	$0.08\pm^{0.11}_{0.01}$	$0.05\pm^{0.06}_{0.04}$	$0.02\pm^{0.04}_{0.02}$	$0.04\pm^{0.02}_{0.02}$
[Pr/H]	$0.08\pm^{0.03}_{0.06}$	$0.11\pm^{0.03}_{0.01}$	$0.08\pm^{0.04}_{0.01}$	$0.05\pm^{0.02}_{0.01}$	$0.03\pm^{0.01}_{0.01}$
[Nd/H]	$-0.04\pm^{0.07}_{0.13}$	$0.02\pm^{0.06}_{0.08}$	$0.02\pm^{0.04}_{0.04}$	$0.02\pm^{0.01}_{0.02}$	$0.02\pm^{0.00}_{0.00}$
[Sm/H]	$-0.28\pm^{0.13}_{0.14}$	$-0.07\pm^{0.06}_{0.02}$	$0.03\pm^{0.03^{*}}_{0.03}$	$0.02\pm^{0.02}_{0.02}$	$0.02\pm^{0.00}_{0.00}$
[Eu/H]	$-0.12\pm^{0.07}_{0.43}$	$-0.00\pm^{0.01^{*}}_{0.10}$	$0.01^{*}\pm^{0.03^{*}}_{0.01}$	$0.00^{*}\pm^{0.01^{*}}_{0.00}$	$0.00^{*}\pm^{0.00^{*}}_{0.00}$
[Gd/H]	$0.14\pm^{0.29}_{0.51}$	$0.10\pm^{0.05}_{0.16}$	$0.04\pm^{0.03}_{0.05}$	$0.01\pm^{0.02}_{0.03}$	$0.01\pm^{0.01}_{0.01}$
[Dy/H]	$0.10\pm^{0.14^{*}}_{0.47}$	$0.03\pm^{0.15}_{0.23}$	$0.05\pm^{0.07}_{0.07}$	$0.05\pm^{0.01}_{0.01}$	$0.04\pm^{0.01}_{0.01}$
[Er/H]	$0.33^{*}\pm^{0.27^{*}}_{0.19}$	$0.17\pm^{0.11^{*}}_{0.02}$	$0.13\pm^{0.14}_{0.06}$	$0.05\pm^{0.05}_{0.02}$	$0.10\pm^{0.04}_{0.04}$
[Th/H]	-0.41	$-0.18\pm^{0.04}_{0.04}$	$-0.02\pm^{0.02}_{0.07}$	$0.00\pm^{0.00}_{0.05}$	-0.04

Table 3.4: Trends in stellar label recovery with resolution. Asterisks denote instances where the reported quantities are impacted by the imposed boundaries of the training set. If no systematic uncertainty is provided, there was only one measurement for which the statistical uncertainty was <0.5 dex.

3.5.1.1 Atmospheric Parameters

Effective Temperature, Surface Gravity, and [Fe/H] In Figure 3.9, we present the change in recovered atmospheric parameters $T_{\rm eff}$, $\log g$, and [Fe/H] as a function of resolution. Only very minimal differences are found between high-resolution and low-resolution measurements. At $R \sim 2500$, $T_{\rm eff}$, $\log g$, and [Fe/H] only differ by approximately $+1$ K, -0.01 dex, and -0.02 dex respectively from the measurements made at $R \sim 40{,}000$–$80{,}000$. The systematic uncertainties are similarly small.

The robust recovery of [Fe/H] at all resolutions is reassuring, albeit unsurprising given the abundance of well calibrated Fe absorption lines and a long history of reliable low-resolution (and photometric) stellar metallicity measurements. The wealth of well-modelled Fe lines minimizes the impact of blending with imperfectly modelled lines at low resolution. The similarity in trend between $T_{\rm eff}$, $\log g$, and [Fe/H] is a direct result of the strong covariance between these labels, which is introduced by the determination of $T_{\rm eff}$ and $\log g$ from isochrones dependent on the stellar photometry and [Fe/H].

Radial, Macroturbulent, and Microturbulent Velocities Figure 3.10 shows the changes in the recovered velocity-related parameters v_r, $v_{\rm macro}$, and $v_{\rm micro}$. As expected, the recovery of radial velocity across the sample shows little trend with resolving power. The systematic uncertainty in v_r is $\lesssim$0.25 km s^{-1} for $R \gtrsim 5000$ and $\sim$1 km s^{-1} for $R \sim 2500$. These spreads are large compared to the formal measurement uncertainties (0.1–0.6 km s^{-1}), but on par with expectations for low- and medium-resolution surveys (e.g., Xiong et al. 2021).

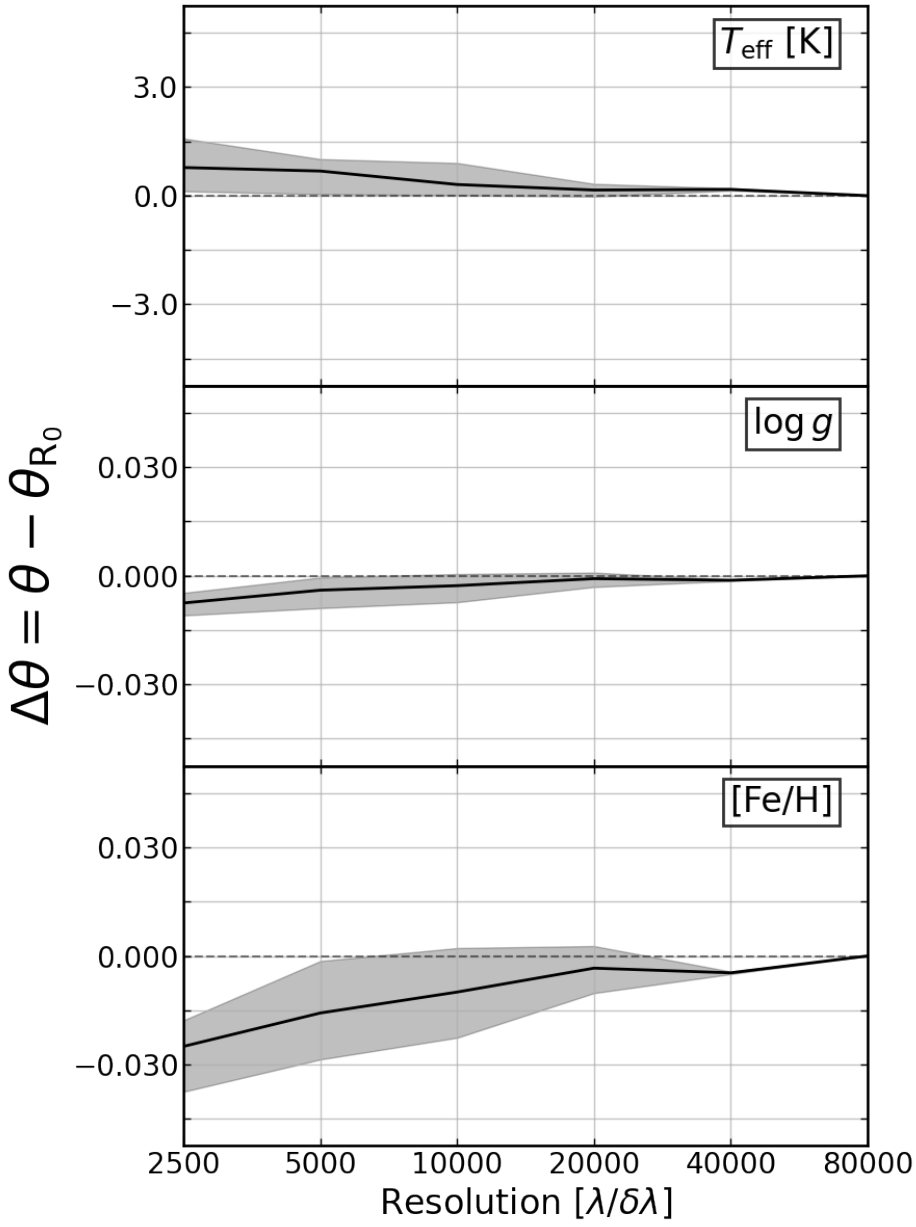


Figure 3.9: Systematic biases (solid black lines) and 1-σ systematic uncertainties (gray shaded regions) in the recovery of $T_{\rm eff}$ (top), $\log g$ (middle), and [Fe/H] (bottom) as a function of resolution. All three labels are recovered with only very minimal differences (+1 K, −0.01 dex, and −0.03 dex) across the entire range of resolutions analyzed. Systematic uncertainties are similarly small.

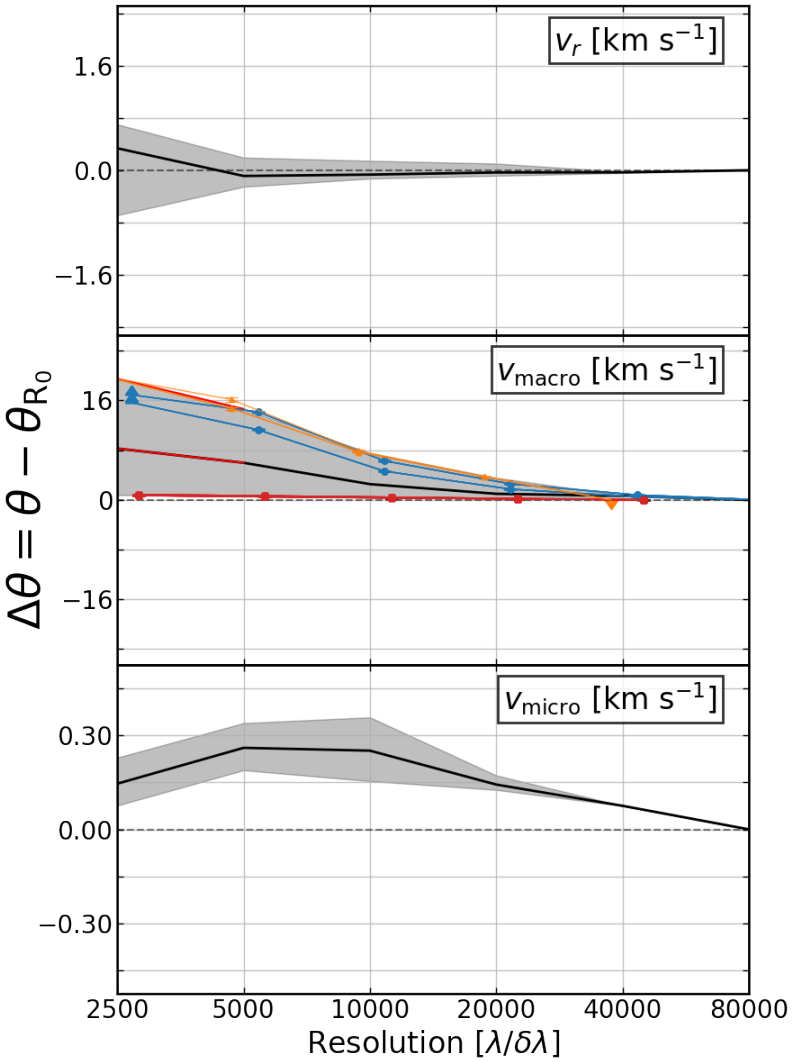


Figure 3.10: Same as Figure 3.9 except for velocity-based atmospheric parameters v_r (top), $v_{\rm macro}$ (middle), and $v_{\rm micro}$ (bottom). v_r is recovered consistently at all resolutions with small systematic uncertainties ($\lesssim$0.5 km s^{-1} for $R \gtrsim 5000$ and $\sim$1.5 km s^{-1} for $R \sim 2500$). $v_{\rm macro}$ exhibits distinct trends between the U09H observations (red squares) and the 147Hr and C316Hr observations (orange triangles and blue circles respectively). The large systematic offsets seen at low resolution for the latter observations is attributed to incorrectly specified instrumental broadening. We recover $v_{\rm micro} \lesssim 0.5$ km s^{-1} higher at lower resolutions consistent with its correlation with [Fe/H] and the small trend seen for [Fe/H] in Figure 3.9.

For $v_{\rm macro}$, we find two distinct trends, one for the older U09H observations (red squares) and one for the post-upgrade C147Hr and C316Hr observations (orange triangles and blue

circles respectively). For the newer observations, the measured value of $v_{\rm macro}$ increases by up to $16\,{\rm km\,s^{-1}}$ as the resolution is decreased to $R \sim 2500$, while for the older observations the resolution dependence is much weaker with an offset of only $\sim 1\,{\rm km\,s^{-1}}$ at $R \sim 2500$. This suggests that the observed trend is driven by an observational systematic present in the C147Hr and C316Hr data, most likely a mismatch between the assumed and true default spectroscopic resolution. Because both macroturbulent and instrumental broadening are implemented with Gaussian kernels, $v_{\rm macro}$ and R are entirely degenerate. As a result of not fitting for R, $v_{\rm macro}$ compensates for this mismatch. We do not find evidence that $v_{\rm macro}$ is correlated in any meaningful way with stellar chemical abundances, which are the primary concern of this study. As such, we simply treat $v_{\rm macro}$ as a nuisance parameter that characterizes the instrumental LSF.

For $v_{\rm micro}$, a more moderate trend with resolution is seen with measurements $\sim$0.1–$0.3\,{\rm km\,s^{-1}}$ larger at $R \lesssim 20{,}000$ than the measurements made at $R \gtrsim 40{,}000$. Most, if not all, of this offset can be attributed to the correlation of $v_{\rm micro}$ and [Fe/H], which we find to be the two most highly correlated stellar labels in our analysis (with the exception of $T_{\rm eff}$ and $\log g$). The $-0.03\,{\rm dex}$ Δ[Fe/H] seen in Figure 3.9 alone can explain $\Delta v_{\rm micro} \sim 0.15$. The growth of spectral masks with decreasing resolution (see Section 3.4.1.1) may also impact the fitting of extended line profiles, which could introduce systematics into the measurement of $v_{\rm micro}$.

3.5.1.2 Iron-Peak Element Abundances

In Figures 3.11–3.13, we present the systematic bias and uncertainty of iron-peak elements (V, Cr, Mn, Fe, Co, Ni, Cu, and Zn) as a function of resolution. In summary, we find that [V/H], [Cr/H], [Fe/H], [Co/H], and [Ni/H] are recovered consistently at all resolutions, though [V/H] and [Cr/H] display small biases towards higher abundances at low resolution. [Cu/H] and [Zn/H] exhibit large systematic biases and uncertainties at low resolution. For [Mn/H] the boundary of the training grid limit a complete picture of the bias and spread in abundance measurement as a function of resolution. We describe the results for each element in detail below.

Vanadium We recover [V/H] consistently and with small systematic uncertainties ($<$0.05 dex) at all resolutions higher than $R \gtrsim 10{,}000$. At lower resolutions, a small bias towards higher [V/H] develops and grows to 0.07 dex, but increasing systematic uncertainties maintain 1σ consistency with the $R \sim 40{,}000$–$80{,}000$ measurements. The increasing systematic uncertainty is driven by diverging trends between the older blue-only U09H observations (red squares), which trend higher as the resolution is decreased, and the newer full-optical C147Hr and C316Hr observations (orange triangles and blue circles respectively), which trend lower as the resolution is decreased. These trends are driven by heavy blending of mis-modeled lines in the blue ($\lambda < 4500\,$Å) and mis-fit continuum regions coupled with weak lines in the red ($\lambda \sim 6000\,$Å) respectively.

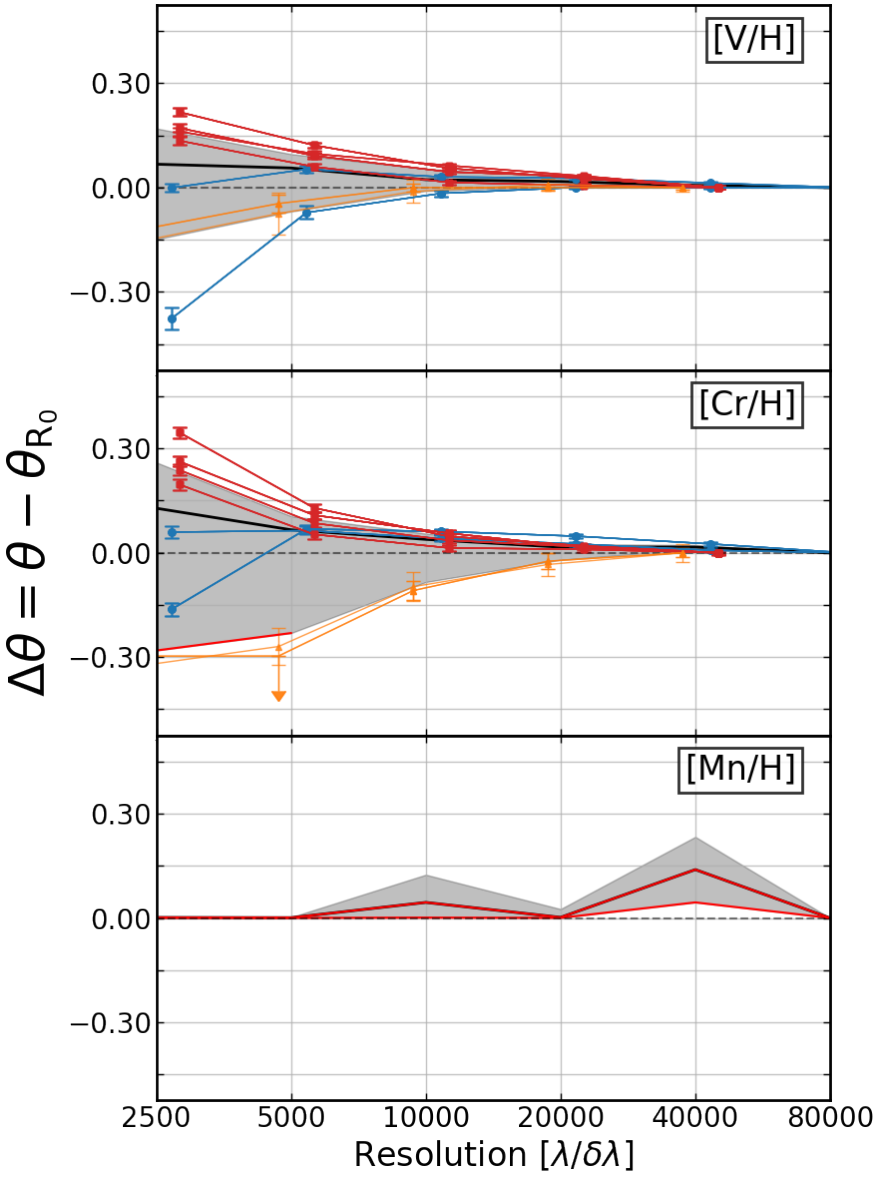


Figure 3.11: Same as Figure 3.9 except for iron-peak elements V (top), Cr (middle), and Mn (bottom). [V/H] and [Cr/H] are recovered consistently down to $R \sim 5000$ with gradually increasing systematic uncertainties and a slight bias towards higher values at the lowest resolution. These systematic trends are driven by a combination of blending of imperfectly modeled lines in the blue and imperfectly modeled continuum regions in the red. Upper limits on [Mn/H] are recovered at all resolutions.

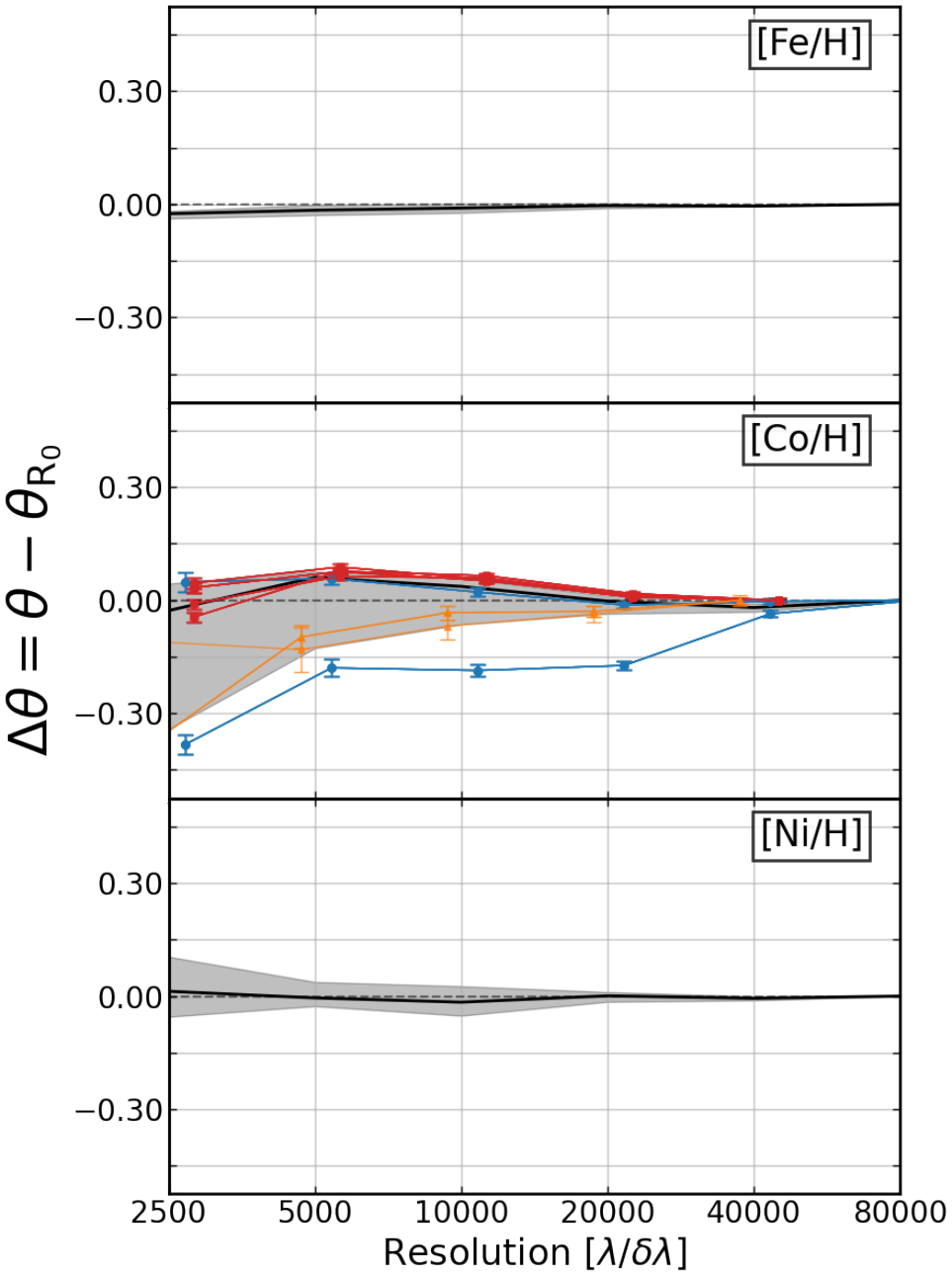


Figure 3.12: Same as Figure 3.9 except for iron-peak elements Fe (top), Co (middle), and Ni (bottom). We find very small ($\lesssim 0.03$ dex) systematic effects for [Fe/H]. [Co/H] and [Ni/H] are also recovered consistently at all resolutions, though [Co/H] exhibits a substantial $\sim$0.3 dex systematic uncertainty at the lowest resolutions.

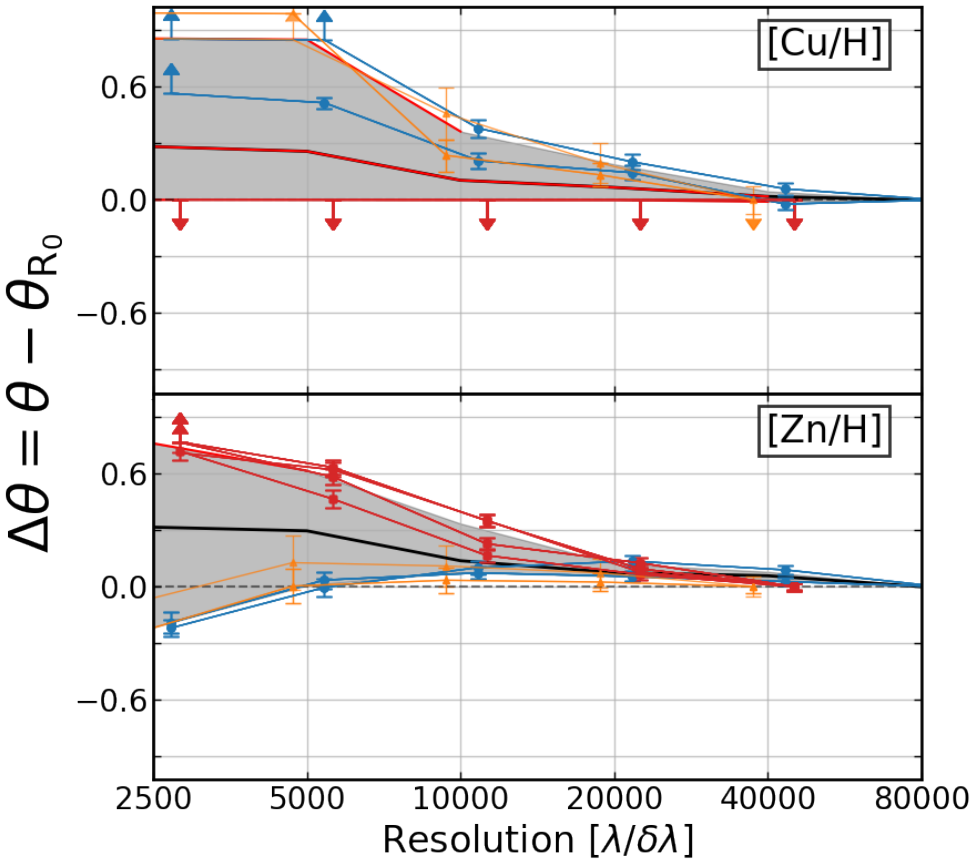


Figure 3.13: Same as Figure 3.9 except for iron-peak elements Cu (top) and Zn (bottom). We find the recovery of [Cu/H] and [Zn/H] to be strongly resolution dependent with both large biases and uncertainties. These systematic effects are driven by the reliance on only a handful of lines which are influenced by the modeling of neighboring lines at lower resolutions.

Chromium The recovery of [Cr/H] as a function of resolution resembles that of [V/H] discussed above. As the resolution is decreased to $R \sim 2500$, the systematic bias and uncertainty increase to $\sim$0.13 and $\sim$0.3 dex respectively. As with [V/H], the increasing systematic uncertainty is driven by diverging behavior between the U09H, C147Hr, and C316Hr observations. The same underlying causes can be attributed as well. At $R < 5000$, only upper limits on [Cr/H] are recovered for the C147Hr observations, leading to the lower uncertainty interval being underestimated.

Manganese For nearly all observations and resolutions we recover [Mn/H] that are near or at the lower bound of the training set ([Mn/Fe] $= -0.5$), which precludes any robust quantification of the systematic bias and uncertainty. As we discuss in Appendix 3.8.3.1,

this is in general agreement with previous LTE measurements from Sobeck, Ivans, et al. (2006) and Sobeck, Kraft, et al. (2011) who measure [Mn/Fe] for these stars in the range of -0.3 to -0.6 dex. Fits to the $\sim$50 Mn lines in the spectra appear qualitatively reasonable, suggesting that the [Mn/Fe] value we would recover with an extended training set is not too far beyond the currently imposed limits.

Iron As discussed previously in Section 3.5.1.1, [Fe/H] is recovered with only small ($\lesssim$0.02 dex) systematic biases and uncertainties across the full range in resolutions analyzed.

Cobalt We find that [Co/H] is generally recovered consistently from $R \sim$ 2500–80,000. [Co/H] recovery exhibits a small positive bias of 0.04–0.06 dex at $5000 \lesssim R \lesssim 10{,}000$, but none at $R \sim 2500$. A negatively skewed systematic uncertainty increases gradually and grows to $\sim$0.3 dex at the lowest resolution, primarily driven by large negative biases in the measurements from C147Hr and C316Hr observations.

Upon deeper investigation, we determine that these biases can be traced to two sources: the poorly-modeled CN band at λ3883 in K731, K934, and K969 and the reliance on weak red-optical Co I lines, which are biased by the presence of correlated noise in the low-resolution, high-S/N regime. With their bluer wavelength coverage, the U09H observations contain approximately three times as many Co lines, leading to more consistent [Co/H] recovery.

Nickel We recover [Ni/H] consistently across the full range of resolutions analyzed. Systematic uncertainties increase gradually with decreasing resolution to 0.08 dex at $R \sim 2500$.

Copper We find the resolution-dependent recovery of [Cu/H] to be strongly dependent on the observational setup. For U09H observations, we recover only upper bounds ([Cu/Fe] $<$ -0.5 dex), while for C147Hr and C316Hr observations, we measure [Cu/H] values that steadily rise by nearly 1 dex and become lower limits ([Cu/Fe] > 0.5 dex) as the resolution is decreased to $R \sim 2500$. The result is a very large (0.3–0.8 dex) systematic uncertainty that is likely still underestimated due to the limits imposed by our priors. The systematic bias, nominally 0.3 dex at $R \sim 2500$, is likely also underestimated.

For the U09H observations, constraints on [Cu/H] come predominantly from two weak Cu I lines ($\lambda\lambda$063.8, 5107.0), which are both underestimated. The C316Hr and C147Hr observations also include the λ5783.7 Cu I line, which is located next to the edge of an order making it quite sensitive to the continuum determination. Indeed, we find that the trend towards higher [Cu/Fe] with decreasing resolution is caused by an increasingly poor fit to the continuum in the region of this line.

Zinc The recovery of [Zn/H] as a function of resolution resembles a more extreme case of the systematic biases and uncertainties seen for [V/H] and [Cr/H]. As we decrease the resolution below $R \lesssim 10{,}000$, a $\sim$0.3 dex positive bias develops and the systematic uncertainty grows

to ∼0.5 dex. We find that this trend is predominantly driven by the recovery of [Zn/H] from U09H measurements, which are ∼0.7 dex larger at $R \sim 2500$ than at the default resolution of $R \sim 40{,}000$. [Zn/H] measurements from C147Hr and C316Hr observations are largely consistent across all resolutions.

Because the measurement of [Zn/H] relies on only three Zn I lines ($\lambda\lambda$4681.4, 4723.5, 4811.9), the measurement is quite sensitive to systematics. In the U09H observations, the λ4723.5 line falls near the edge of the detector and is partially lost as the resolution is decreased. This further increases the impact of blending with poorly-modeled lines on the remaining two lines.

3.5.1.3 α Element Abundances

In Figure 3.14, we present the change in the recovered abundance of α elements (Mg, Si, Ca, and Ti) as a function of resolution. In summary, we find that [Mg/H], [Ca/H], and [Ti/H] are recovered consistently with small to modest systematic uncertainties across all resolutions, while [Si/H] displays a substantial bias towards higher abundances at nearly all resolutions. We describe the results for each element in detail below.

Magnesium We recover [Mg/H] consistently at all resolutions and find that the systematic uncertainty gradually increases to ∼0.15 dex as the resolution is decreased to $R \sim 2500$. Systematic uncertainties on [Mg/H] are slightly underestimated for $R < 40{,}000$ measurements, because we recover only upper bounds ([Mg/H] < -0.25) for stars K731 and K969. We find sizeable scatter (∼0.1–0.2 dex) between the [Mg/H] measured from repeat observation, which contributes to the systematic uncertainty seen for the stacked measurements. This is due to the fact that most of the strong Mg lines in the spectrum exhibit substantial NLTE effects and are masked or down-weighted in the fit. As a result, the measurement of [Mg/H] relies more heavily on weaker Mg lines and indirect information scattered throughout the spectrum.

Silicon We recover [Si/H] to be 0.1–0.15 dex larger at nearly all resolutions smaller than the default resolution. Similarly sized systematic uncertainties are present as well, though they are positively skewed to even higher [Si/H] abundances. Combined with the mixed agreement to literature [Si/H] measurements (see Appendix 3.8.3.1), this suggests that substantial model inaccuracies exist. Indeed, much of the spectral information for Si is indirectly accessible through absorption lines of other elements, which are not modeled sufficiently accurately in this work. As with [Mg/H], this reliance on indirect spectral features also adds ∼0.1–0.2 dex scatter between repeat observations of the same star.

Calcium We recover [Ca/H] consistently across the full range of resolutions analyzed. Systematic uncertainties increase gradually with decreasing resolution to ∼0.1 dex at $R \sim 2500$.

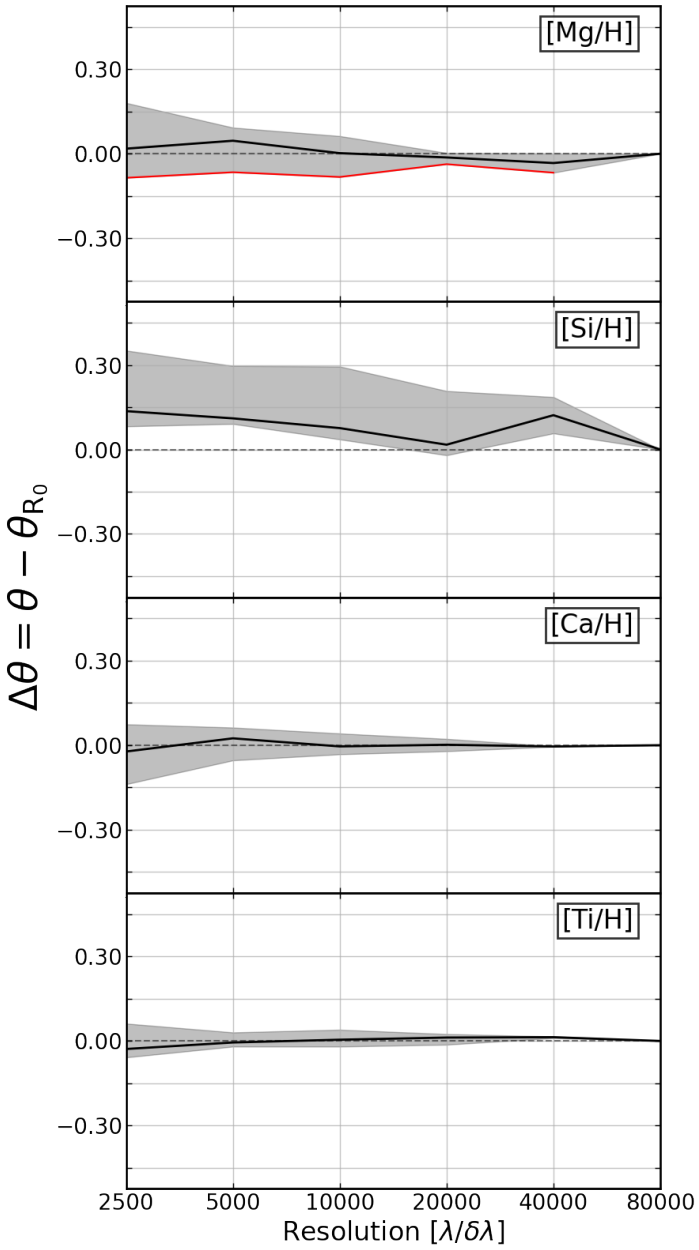


Figure 3.14: Same as Figure 3.9 except for α elements Mg, Si, Ca, and Ti (from top to bottom). We find the recovery of [Mg/H], [Ca/H], and [Ti/H] to be consistent as a function of resolution down to $R \sim 2500$. Larger uncertainties on [Mg/H] are due to the masking of NLTE-sensitive lines. [Si/H] displays a substantial bias with resolution and large systematic uncertainties as a result of its strong dependence on the stellar atmospheric structure.

Titanium We recover [Ti/H] consistently across the full range of resolutions analyzed. Systematic uncertainties increase gradually with decreasing resolution to $\sim$0.05 dex at $R \sim$ 2500.

3.5.1.4 C, N, O Abundances

In Figure 3.15, we present the change in the recovered abundance of the light elements C, N, and O as a function of resolution. In summary, we find [C/H] and [N/H] to be recovered robustly and consistently at all resolutions, while the consistent recovery of [O/H] is more challenging. For [N/H] and [O/H] the boundaries of the training grid limit a complete picture of the bias and spread in abundance as a function of resolution. The recovery of each element is described in more detail below.

Carbon While [C/H] recovery exhibits a small $\lesssim$0.05 dex positive bias for $R \lesssim$ 10,000, it is largely consistent across the full range of resolutions. Systematic uncertainties increase gradually with decreasing resolution to $\sim$0.05 dex at $R \sim$ 2500. The small positive bias may be related to the strong negative correlation we find between [C/H] and [Fe/H] (see Section 3.6.3). Despite their complicated nature, the C molecular features are fit well at all resolutions. This is reassuring given the large number of low-resolution searches for C-enhanced metal poor stars (e.g., Arentsen et al. 2022, and references therein).

Nitrogen Because we recover lower limits on [N/H] ([N/H $>$ 1.0) for most of the stars in our sample, it is difficult to robustly quantify any resolution-dependent systematic effects. For 4 stars with blue-optical U09H observations, K341, K386, K462, and K934, we do obtain constraints on [N/H] (i.e., not lower limits) at all resolutions. For all of these but K462, we recover [N/H] consistently (to better than $<$ 0.05 dex). In the case of K462, we recover [N/H] to be $\sim$0.15 dex lower at $R \sim$ 2500 than at the default resolution of $R \sim$ 45,000, though the cause of this bias is challenging to diagnose. Given the presence of so many lower limits, we cannot rule out the presence of a positive bias, nor can we quantify a positive systematic uncertainty.

Oxygen Similar to [N/H], we recover only lower limits on [O/H] ([O/H $>$ 1.0) for the majority of stars, and thus cannot fully quantify the nature of resolution-dependent systematics on the measurement of [O/H]. Large scatter in the U09H observations towards lower [O/H] at low resolution lead to large ($>$0.3 dex) negative systematic uncertainties and a $\sim$0.15 dex bias below $R \sim$ 10,000. This is likely because the vast majority of the O information is present only through indirect effects on C molecular features and changes to the atmospheric structure (see Ting, Conroy, Rix, and Asplund 2018). In the C147Hr and C316Hr observations, two O I lines are accessible at $\lambda\lambda$6302.0, 6365.5, but the former falls in a telluric mask and the later is very weak. Our inability to make conclusive statements regarding the recovery of [O/H] speaks to the challenge of measuring oxygen abundances from optical spectra—even at high resolution.

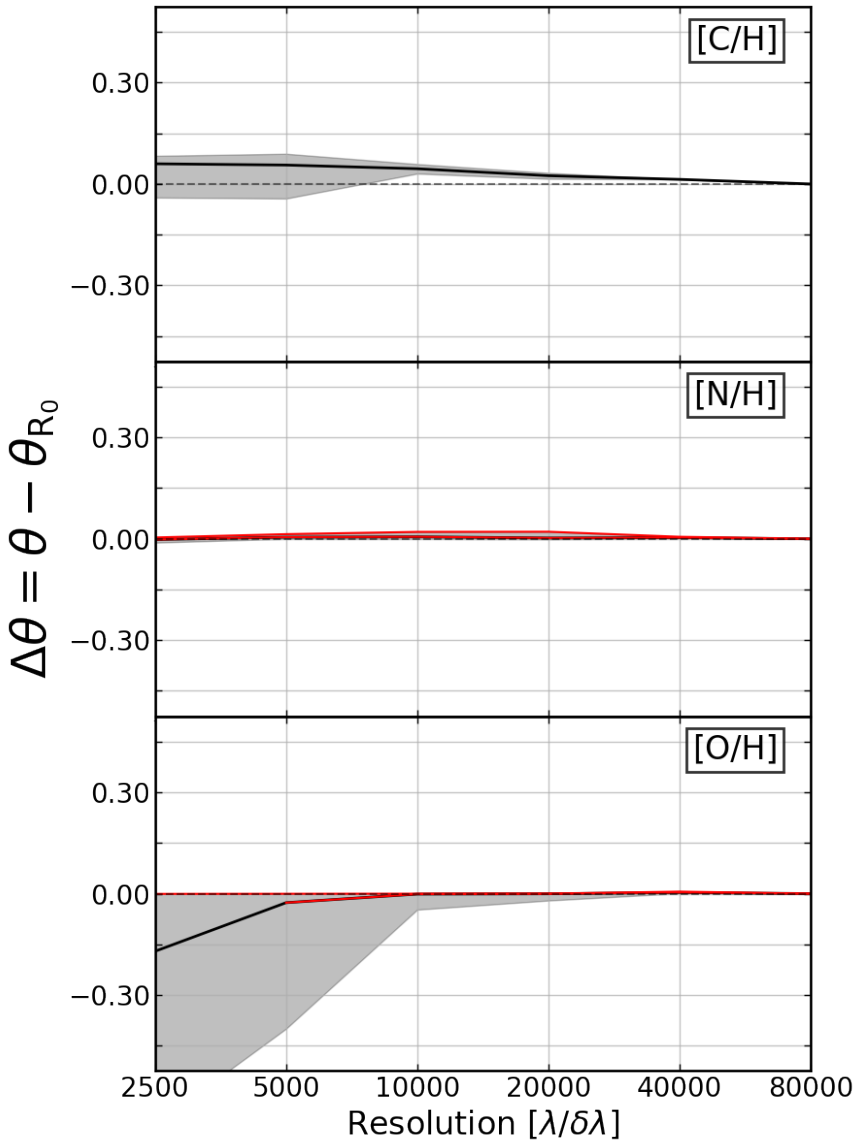


Figure 3.15: Same as Figure 3.9 except for C (top), N (middle), and O (bottom). We recover [C/H] consistently with small uncertainties at all resolutions. Resolution-dependent systematics are challenging to quantify for [N/H] and [O/H] due to the measurement of lower limits. For the U09H observations, the measurement of [N/H] appears consistent as a function of resolution. The measurement of [O/H] from these spectra is particularly challenging as most of the information comes indirectly from the impact of O on the atmospheric structure.

3.5.1.5 Light-Odd Element Abundances

In Figure 3.16, we present the change in the recovered abundance of light-odd elements (Na, Al, K, and Sc) as a function of resolution. With the exception of [Sc/H], we find that light-

odd elements are recovered quite poorly and inconsistently at nearly all resolutions. The recovery of each element is described in more detail below.

Sodium We struggle to recover [Na/H] consistently at nearly all resolutions. The systematic bias towards lower values of [Na/H] brings some measurements into better agreement with the literature (e.g., K341 and K431) but also worsens the agreement of others (e.g., K462; see Appendix 3.8.3.1). The $\gtrsim 0.4$ dex systematic uncertainty seen at $R \lesssim 10{,}000$ is characteristic of the large scatter seen in literature measurements for [Na/H], but the presence of both lower and upper limits in our sample mean that these already large systematic uncertainties are likely underestimated. The challenge in recovering consistent [Na/H] is driven largely by the lack of good Na lines in the spectrum. The two strongest Na feature, the Na doublet at $\lambda\lambda$5891.6, 5897.6, falls entirely within telluric masks, and the three next-strongest lines at $\lambda\lambda$4979.9, 4984.2, 5684.2, 5689.8 are all fairly weak. Two of these lines, those at $\lambda\lambda$4979.9, 4984.2, contribute strongly to the negative bias as they are in close proximity to a handful of poorly-fit Fe I lines.

Aluminum As with [Na/H], we find [Al/H] to be challenging to measure consistently for nearly all observations at all resolutions. The presence of both lower and upper limits in our sample mean that both the large systematic uncertainties ($\gtrsim$0.3 dex) and the $\sim$0.1 dex systematic bias are likely underestimated. Like Na, the challenge in recovering consistent [Al/H] is due to the lack of good Al lines in the data. The two strongest Al features, the Al I lines at $\lambda\lambda$3945.1, 3962.6, are lost to the Ca H&K mask. The remaining Al information is either indirect (mainly through the CN bands) or a handful of very weak Al I lines.

Potassium The recovery of [K/H], like the recovery of [Na/H] and [Al/H], is challenging at all resolutions. Very few measurements fall within the bounds of our training set ($-0.25 \leq$ [K/Fe] ≤ 1.0), making it impossible to quantify the true impact of resolution-dependent systematics. This is due to the extreme paucity of K lines in the observed spectra. The most prominent potassium feature, the K I line at λ7701.1, is a part of the telluric mask. The remaining K features at $\lambda\lambda$4045.3, 4048.4 are very weak, and are further down-weighted by NLTE masks. This lack of K lines prevents any measurement of [K/H] to better precision than 0.5 dex below $R < 10{,}000$.

Scandium Unlike for the other light-odd elements, we recover [Sc/H] consistently down to $R \sim 10{,}000$ and with only a small ($\sim$0.05 dex) systematic negative bias at lower resolutions. The systematic uncertainty gradually increases with decreasing resolution to $\sim$0.1 dex at $R \sim 2500$. In contrast to the dearth of Na, Al, and K lines, there are $\sim$40 Sc lines contained in the archival spectra, enabling robust [Sc/H] measurements at all resolutions. Blending with neighboring imperfectly modelled lines is responsible for the small systematic uncertainty and bias at low resolution.

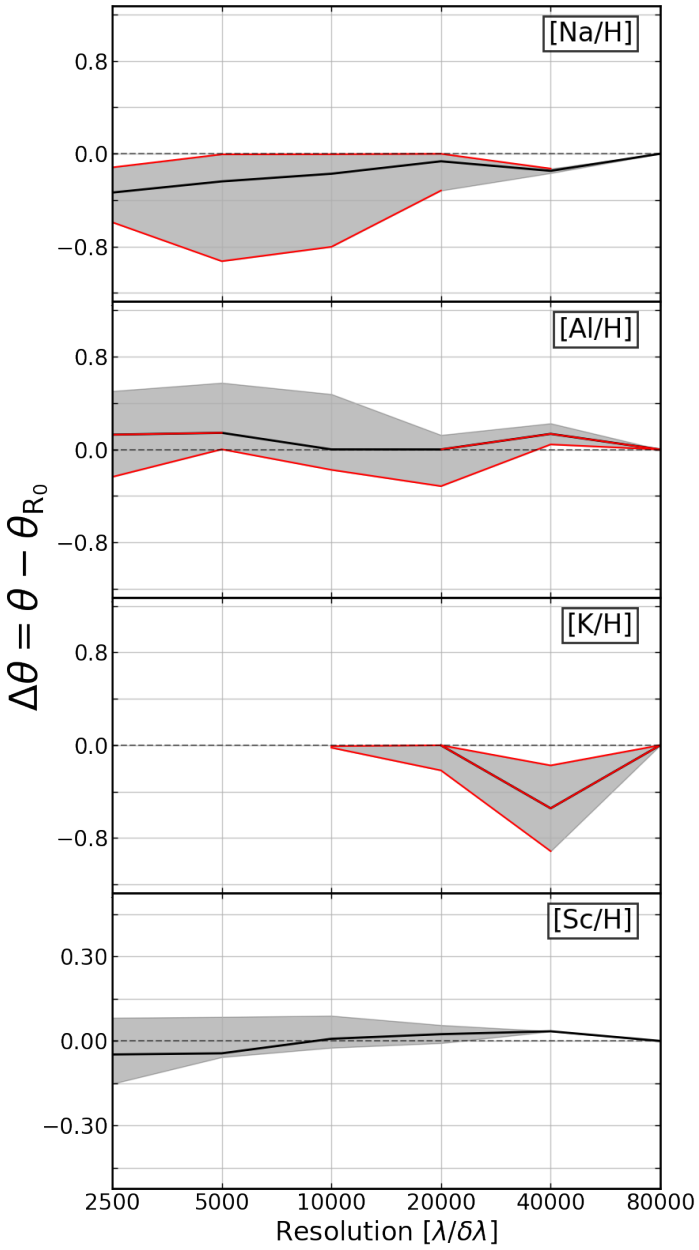


Figure 3.16: Same as Figure 3.9 except for Na, Al, K, and Sc (from top to bottom). We find poor and inconsistent recovery of [Na/H], [Al/H], and [K/H] at nearly all resolutions owing to the sparsity of absorption features. [Sc/H] is recovered consistently as a function of resolution with modest systematic uncertainties.

3.5.1.6 Neutron-Capture Element Abundances

In Figures 3.17–3.19, we present the change in the recovered abundance of neutron-capture elements (Sr, Y, Zr, Ba, La, Ce, Pr, Nd, Sm, Eu, Gd, Dy, Er, and Th) as a function of resolution. In summary, we find [Sr/H], [Y/H], [Zr/H], [Nd/H], [Sm/H], [Gd/H], [Dy/H], and [Th/H] to be recovered consistently down to at least $R \sim 10000$, though a subset show large uncertainties or noticeable biases at the lowest resolutions. We find substantial biases with resolution for [Ce/H], [Pr/H], and [Er/H]. For [Ba/H], [La/H], and [Eu/H], the boundaries of the training grid limit a complete picture of the resolution-dependent systematic bias and uncertainties. The recovery of each element is described in more detail below.

Strontium While the recovery of [Sr/H] is slightly biased by $\sim$0.1 dex at a few individual resolutions, it it consistent to within the 0.1–0.2 dex systematic uncertainties for $R \lesssim 20{,}000$. Constraints on [Sr/H] come primarily from the two strong Sr II resonance lines at $\lambda\lambda$4078.9, 4216.7 and secondarily from two weak lines at $\lambda\lambda$4163.0, 4608.6. Blending, saturation, and NLTE effects in the strong lines all contribute to sizeable ($\sim$0.2 dex) scatter in the measured [Sr/H] from repeat observations at all resolutions.

Yttrium We find the recovery of [Y/H] to be largely consistent as a function of resolution. A small $\lesssim$0.05 (0.10) dex negative (positive) bias is seen for $R \sim$ 10,000 (2500), but this is within the systematic uncertainty, which grows gradually as the resolution is decreased to $\sim$0.15 dex at $R \sim 2500$. Roughly 40 Y lines between 4100 and 5500 Å contribute to the robust measurement of [Y/H]. Blending of these lines with neighboring lines with small errors is responsible for the small systematic uncertainty and bias at low resolution.

Zirconium We recover [Zr/H] consistently for all resolutions $R \gtrsim 5000$ and biased by $\sim$0.1 dex lower values at $R \sim 2500$. Positive systematic uncertainties grow gradually to $\sim$0.1 dex at $R \sim 2500$, while negative systematic uncertainties grow gradually to $\sim$0.3 dex. Both the negative bias and substantially larger negative systematic uncertainties are driven by the measurements from the C147Hr observations, for which we recover [Zr/H] to be as much as $\sim$0.6 dex smaller at $R \sim 2500$ than at the default resolution. The bias in the C147Hr measurements appears to be driven by a combination of blending and a poorly approximated continuum shape.

Barium For over half of the stars in our sample, we recover lower limits on [Ba/H] ([Ba/Fe] > 0.5), which obfuscate the complete picture of resolution-dependent systematics. As we decrease the resolution, we do see increasingly large negative uncertainties—up to 0.45 dex at $R \sim 2500$. This result, along with the poor agreement with literature [Ba/H] measurements (see Appendix 3.8.3.1) suggest that there are some quite substantial inaccuracies in the Ba features of our model. Indeed, line saturation, NLTE effects, and hyperfine splitting are all at play in the strongest optical Ba lines (e.g., Eitner et al. 2019).

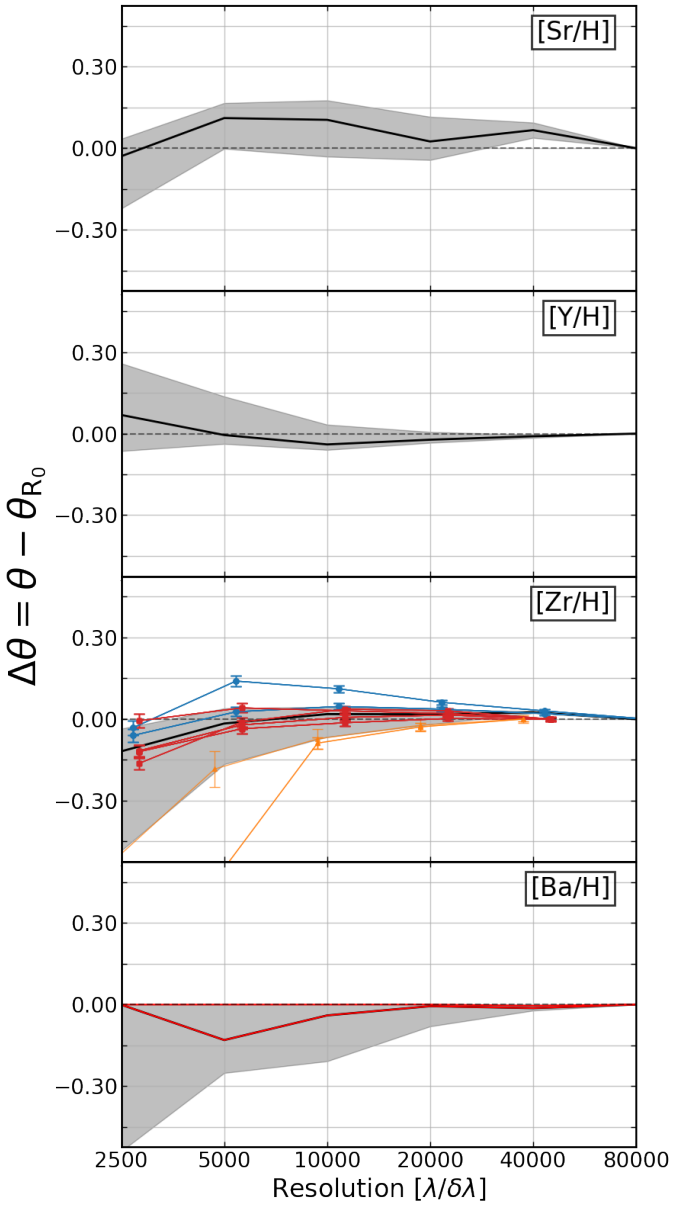


Figure 3.17: Same as Figure 3.9 except for neutron-capture elements Sr, Y, Zr, and Ba (from top to bottom). We recover [Y/H] and [Zr/H] consistently down to $R \sim 5000$ and with small positive and negative biases respectively at $R \sim 2500$. We recover [Sr/H] with a slight positive bias at lower resolution. The presence of lower limits prevents the robust quantification of systematics for [Ba/H]. The measurement of both [Sr/H] and [Ba/H] suffer from substantial NLTE effects and hyperfine splitting in their strong resonance lines.

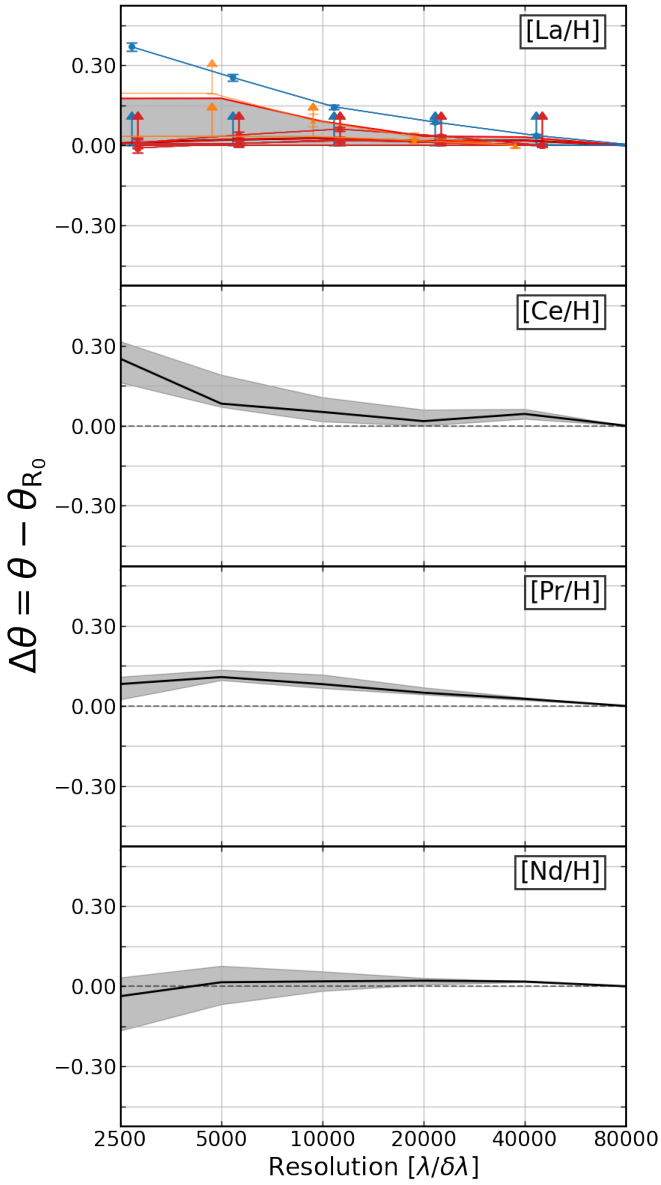


Figure 3.18: Same as Figure 3.9 except for neutron-capture elements La, Ce, Pr, and Nd (from top to bottom). We find a strong resolution dependence for the recovery of [Ce/H] and [Pr/H] as a result of blending in the feature-dense region around 4000 Å. The recovery of [La/H] appears similarly biased though the presence of lower limits prevents robust quantification of the systematic effects. [Nd/H] is recovered consistently as a function of resolution with modest systematic uncertainties.

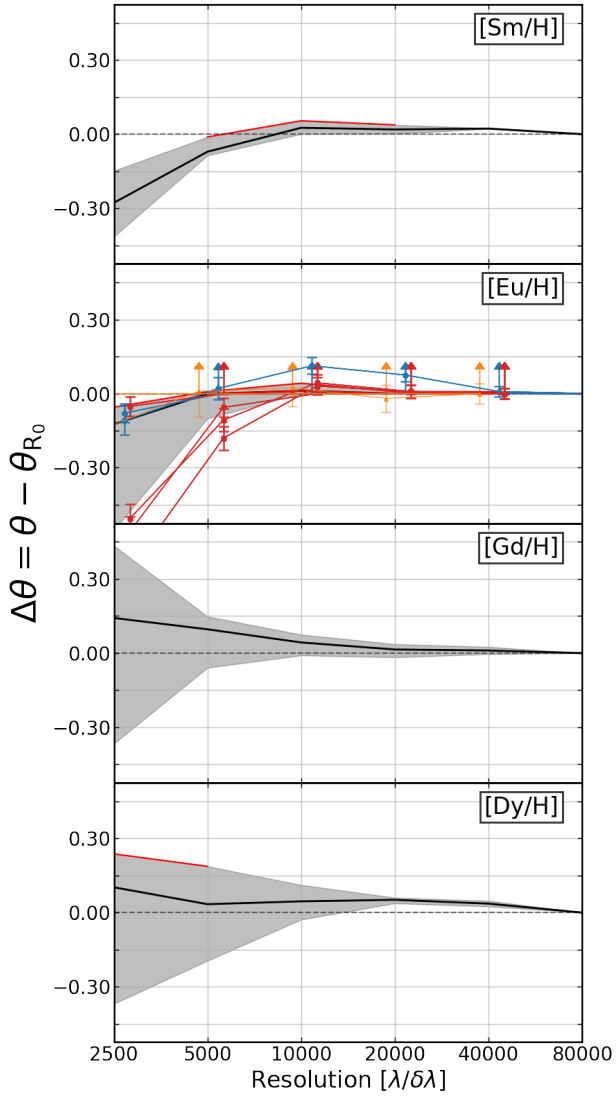


Figure 3.19: Same as Figure 3.9 except for neutron-capture elements Sm, Eu, Gd, and Dy (from top to bottom). The recovery of [Gd/H] and [Dy/H] appears consistent, albeit with rapidly increasing uncertainties down to $R \sim 2500$. We find that [Sm/H] is well recovered down to $R \sim 10{,}000$, but exhibits increasing bias at lower resolutions. The measurement of [Gd/H], [Dy/H], and [Sm/H] are all characterized by fitting many very weak lines in crowded regions of the stellar spectrum. The presence of lower limits for [Eu/H] prevents robust quantification of the resolution-dependent systematic effects on its recovery. [Eu/H] also suffers from substantial NLTE effects and hyperfine splitting in their strong resonance lines.

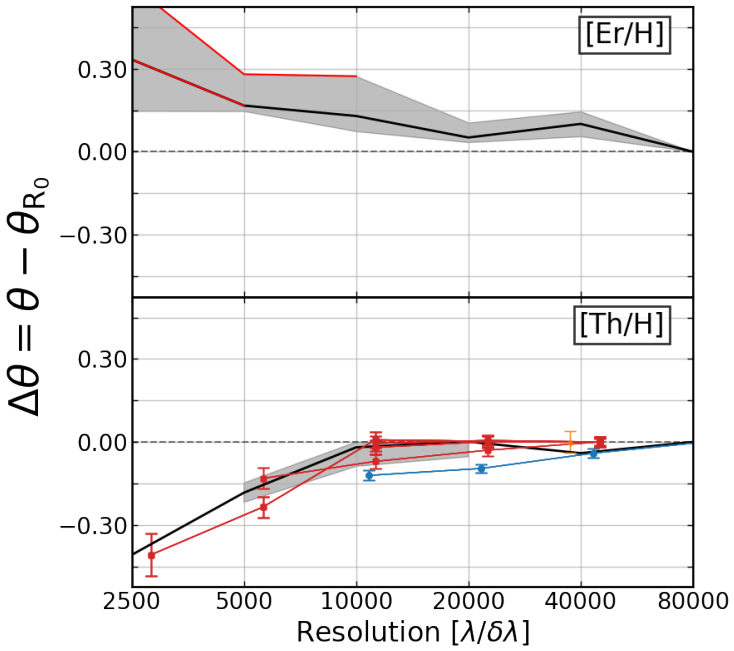


Figure 3.20: Same as Figure 3.9 except for neutron-capture elements Er (top) and Th (bottom). The recovery of [Er/H] and [Th/H] are all made difficult due to a paucity of absorption lines in the observed spectra. The recovery of [Er/H] exhibits a substantial bias towards larger values as the resolution is decreased. A negative bias is apparent for the few instances in which [Th/H] can be recovered below $R \sim 10{,}000$.

Lanthanum The recovery of [La/H] displays a biased towards larger values at lower resolutions, though the extent of this bias is unknown due to the boundary of our model grid. Lower limits ([La/Fe] > 0.5 dex) are recovered for the majority of stars. Similarly, the systematic uncertainty on [La/H] grows to at least $\sim$0.15 dex as we decrease the resolution to $R \sim 2500$. This bias is predominantly driven by measurements made with the C147Hr and C316Hr observations, which are biased by as much 0.35 dex as a result of blending in a few important lines at longer wavelengths ($\lambda\lambda$4663.8, 4922.4, 4923.2, 5124.4, 6392.3). The U09H observations, on the other, yield quite consistent (to $\lesssim$0.05 dex) [La/H] measurements across all resolutions.

Cerium We find a growing systematic bias towards higher [Ce/H] as the resolution is decreased. At $R \sim 2500$ ($R \sim 5000$), we measure [Ce/H] $\sim$0.25 (0.08) dex higher than we do at the default resolutions. Despite the substantial bias, systematic uncertainties remain small ($\lesssim$ 0.05 dex for $R \gtrsim 10{,}000$ and $\sim$0.1 dex for $R \lesssim 5000$). While one would expect the large number (100's–1000's) of Ce lines present in these spectra to yield robust [Ce/H]

measurements regardless of resolution, a closer inspection of the spectra reveals that the majority of these Ce lines reside between 3800 and 4600 Å among a high density of other lines, including complex molecular absorption bands. As a result, the impact of blending in this portion of the spectrum is especially large. When the resolution decreases, [Ce/H] increases to compensate for missing and underestimated lines.

Praseodymium We find a similar, albeit smaller, systematic bias with resolution for [Pr/H] recovery as we do for [Ce/H]. At $R \sim 20{,}000$, we recover [Pr/H] to be 0.05 dex larger than at higher resolutions. At lower resolutions, this bias increases slightly to $\sim$0.1 dex. Systematic uncertainties remain small ($\lesssim$0.05 dex) across all resolutions. As with Ce, blending, especially in the region between 4000 and 4100 Å, is source of the systematic bias.

Neodymium We recover [Nd/H] consistently across the full range of resolutions analyzed. Systematic uncertainties increase gradually with decreasing resolution to $\sim$0.1 dex at $R \sim 2500$. While Nd has a similar number of lines as Ce and Pr, these lines are more broadly distributed throughout the spectrum. As a result, the recovery of [Nd/H] is less susceptible to the impact of blending in the blue-optical.

Samarium We recover [Sm/H] consistently at resolutions above $R \gtrsim 10{,}000$. At lower resolutions, a systematic negative bias develops and grows to 0.25 dex at $R \sim 2500$. Systematic uncertainties grow to as large as $\sim$0.2 dex, though these may be underestimated due to lower limits recovered at our model grid boundary ([Sm/Fe] > 1). Upon visual inspection, the ability of our spectral model to fit the many Sm lines present in the data is quite mixed—some are fit well, others are overestimated, and others still are underestimated. We believe the source of the bias at the lowest resolutions is due to the dominance of a few of the stronger lines, namely at $\lambda\lambda$4069.5, 4108.4, 4156.4, 4204.2, 4468.6, which are overestimated at the default resolution.

Europium The recovery of [Eu/H] as a function of resolution, like that of [Sm/H], is obscured by lower limits at the boundary of our model grid ([Eu/Fe] > 1). At the lowest resolution, we find a systematic $\sim$0.15 dex bias towards lower values of [Eu/H] and a large $\sim$0.4 dex systematic uncertainty. This bias is most pronounced for measurements of [Eu/H] from U09H observations. This result, along with the poor agreement with literature [Eu/H] measurements (see Appendix 3.8.3.1) suggest that there are some quite substantial inaccuracies in the Eu features of our model. As with Ba, line saturation, NLTE effects, and hyperfine splitting are all at play in the strongest optical Eu lines (Mashonkina and Gehren 2000).

Gadolinium We recover [Gd/H] consistently at all resolutions when the systematic uncertainty is taken into account. As the resolution is decreased to $R \sim 5000$ (2500), a systematic bias towards higher [Gd/H] grows to $\sim$0.1 (0.25) dex, while the systematic uncertainty grows

at a larger rate, reaching ∼0.15 (0.55) dex. The increasing systematic uncertainty is driven by blending of complicated and imperfectly modeled absorption features with the ∼100 weak Gd lines present bluewards of 4500 Å.

Dysprosium The recovery of [Dy/H] is largely consistent across all resolutions, though a small positive bias of ∼0.05 dex is seen at intermediate resolutions ($5000 \lesssim R \lesssim 40{,}000$). The systematic uncertainty in [Dy/H] recovery grows steadily with decreasing resolution to ∼0.4 dex at $R \sim 2500$. As for Gd, the increasing systematic uncertainty is driven by blending of complicated and imperfectly modeled absorption features with the ∼100 weak Dy lines present bluewards of 4500 Å.

Erbium We find that the recovery of [Er/H] is biased high by >0.1 dex at all resolutions lower than the default. As the resolution is decreased to $R \sim 2500$, the bias increases steadily to at least 0.3 dex. The positively skewed systematic uncertainty increases with decreasing resolution to ≳0.2 dex. Due to the recovery of several lower limits ([Er/Fe] > 1.0), the bias and uncertainties for $R \lesssim 5000$ may be underestimated. Compared to other neutron-capture elements, Er has far fewer lines in the observed spectra, and the lines that do exist are quite weak and blended. As a result, the recovery of [Er/H] is quite sensitive to model fidelity at all but the highest resolutions.

Thorium We find [Th/H] to be consistently recovered for $R \gtrsim 10{,}000$. At lower resolutions, we measure [Th/H] to be ∼0.4 dex smaller. Systematic uncertainties are ∼0.05 dex, although the limited number of stars for which we can measure [Th/H], especially at low resolution, adds makes the uncertainty difficult to quantify across the full resolution range. Our ability to recover Th is limited by the small handful of Th II lines detectable in the observed spectra. Of the three strongest lines, λλ3676.6, 3742.2, 4020.3, only the last is contained within spectral coverage of the C147Hr and C316Hr observations. Nearly all Th II lines are substantially impacted by blends at lower resolutions, leading to the observed bias.

3.5.2 Label Recovery as a Function of Signal/Noise

Here we present the change in the recovered stellar parameters as a function of S/N for our sample of stars fit at $R \sim 10{,}000$. Similar to the presentation in Section 3.5.1, the change in stellar parameters, $\Delta\theta$, is reported relative to a fiducial measurement, in this case, the value recovered at the native S/N with the same resolution ($\Delta\theta = \theta_\sigma - \theta_{\sigma_0}$). In this analysis, we consider stellar label measurements from individual exposures rather than from the stacked posteriors so that they can be more easily mapped to a median S/N. The results of this analysis for each element are presented in Table 3.5.

Figure 3.21, illustrates the trends in recovery as a function of S/N for the 20 elements (C, Mg, Ca, Sc, Ti, V, Cr, Fe, Co, Ni, Sr, Y, Zr, Ce, Pr, Nd, Sm, Gd, Dy, and Th) that we found

θ	S/N $\sim$ 5 $\Delta\theta \pm^{\sigma_{84\text{th}}}_{\sigma_{16\text{th}}} \pm\sigma_{\text{stat}}$	S/N $\sim$ 10 $\Delta\theta \pm^{\sigma_{84\text{th}}}_{\sigma_{16\text{th}}} \pm\sigma_{\text{stat}}$	S/N $\sim$ 20 $\Delta\theta \pm^{\sigma_{84\text{th}}}_{\sigma_{16\text{th}}} \pm\sigma_{\text{stat}}$	S/N $\sim$ 40 $\Delta\theta \pm^{\sigma_{84\text{th}}}_{\sigma_{16\text{th}}} \pm\sigma_{\text{stat}}$	S/N $\sim$ 75 $\Delta\theta \pm^{\sigma_{84\text{th}}}_{\sigma_{16\text{th}}} \pm\sigma_{\text{stat}}$
T_{eff} [K]	$0.38 \pm^{1.04}_{0.79} \pm1.44$	$-0.16 \pm^{0.51}_{0.55} \pm0.99$	$-0.41 \pm^{0.19}_{0.51} \pm0.55$	$-0.09 \pm^{0.14}_{0.45} \pm0.30$	$0.00 \pm^{0.03}_{0.00} \pm0.20$
$\log g$	$-0.00 \pm^{0.01}_{0.02} \pm0.01$	$0.00 \pm^{0.01}_{0.01} \pm0.01$	$0.00 \pm^{0.00}_{0.00} \pm0.00$	$0.00 \pm^{0.00}_{0.00} \pm0.00$	$0.00 \pm^{0.00}_{0.00} \pm0.00$
v_{micro} [km s^{-1}]	$0.16^* \pm^{0.28^*}_{0.11} \pm0.22$	$0.11 \pm^{0.18^*}_{0.06} \pm0.13$	$0.06 \pm^{0.06^*}_{0.07} \pm0.07$	$0.01 \pm^{0.05}_{0.03} \pm0.04$	$0.00 \pm^{0.00}_{0.00} \pm0.02$
v_{macro} [km s^{-1}]	$4.67 \pm^{0.43}_{4.73} \pm0.83$	$4.42 \pm^{0.36}_{4.08} \pm0.42$	$4.30 \pm^{0.23}_{4.07} \pm0.22$	$4.11 \pm^{0.21}_{4.11} \pm0.12$	$0.00 \pm^{0.00}_{0.00} \pm0.01$
v_r [km s^{-1}]	$-0.22 \pm^{1.17}_{0.12} \pm0.51$	$-0.09 \pm^{0.58}_{0.07} \pm0.26$	$-0.05 \pm^{0.25}_{0.03} \pm0.14$	$-0.01 \pm^{0.09}_{0.03} \pm0.08$	$0.00 \pm^{0.03}_{0.00} \pm0.06$
[C/H]	$-0.11 \pm^{0.09}_{0.09} \pm0.10$	$-0.10 \pm^{0.07}_{0.13} \pm0.07$	$-0.07 \pm^{0.06}_{0.11} \pm0.05$	$-0.02 \pm^{0.02}_{0.09} \pm0.03$	$0.00 \pm^{0.01}_{0.00} \pm0.01$
[N/H]	$0.05^* \pm^{0.02^*}_{0.15} \pm0.71$	$-0.00^* \pm^{0.00^*}_{0.09} \pm0.19$	$-0.03^* \pm^{0.03^*}_{0.07} \pm0.08$	$-0.01^* \pm^{0.01^*}_{0.01} \pm0.04$	$0.00^* \pm^{0.00^*}_{0.00} \pm0.02$
[O/H]	...	$-0.07 \pm^{0.01}_{0.41} \pm0.14$	$-0.09 \pm^{0.12^*}_{0.47} \pm0.22$	$-0.00 \pm^{0.01^*}_{0.29} \pm0.06$	$0.00^* \pm^{0.00^*}_{0.00} \pm0.04$
[Na/H]	...	...	$-0.11^* \pm^{0.01^*}_{0.01} \pm0.36$	$0.00 \pm^{0.13^*}_{0.08^*} \pm0.31$	$0.00 \pm^{0.14^*}_{0.00^*} \pm0.20$
[Mg/H]	$0.45^* \pm^{0.19^*}_{0.14} \pm0.35$	$0.14 \pm^{0.15}_{0.15} \pm0.20$	$0.01 \pm^{0.22}_{0.10} \pm0.13$	$0.01 \pm^{0.12}_{0.05} \pm0.08$	$0.00 \pm^{0.04}_{0.00} \pm0.05$
[Al/H]	...	...	$0.04^* \pm^{0.10}_{0.11^*} \pm0.42$	$0.00^* \pm^{0.07}_{0.30^*} \pm0.31$	$0.00 \pm^{0.00}_{0.04^*} \pm0.17$
[Si/H]	$0.13^* \pm^{0.09^*}_{0.08} \pm0.35$	$0.03 \pm^{0.06^*}_{0.06} \pm0.12$	$-0.02 \pm^{0.04}_{0.04} \pm0.06$	$-0.01 \pm^{0.02}_{0.03} \pm0.03$	$0.00 \pm^{0.00}_{0.00} \pm0.02$
[K/H]	...	...	...	...	$0.00^* \pm^{0.00^*}_{0.00^*} \pm0.30$
[Ca/H]	$-0.28^* \pm^{0.05}_{0.07^*} \pm0.07$	$-0.28^* \pm^{0.06}_{0.07^*} \pm0.05$	$-0.24^* \pm^{0.14}_{0.08^*} \pm0.05$	$-0.13 \pm^{0.15}_{0.07} \pm0.05$	$0.00 \pm^{0.03}_{0.00} \pm0.03$
[Sc/H]	$-0.06 \pm^{0.04}_{0.22^*} \pm0.31$	$-0.08 \pm^{0.03}_{0.12} \pm0.17$	$-0.06 \pm^{0.02}_{0.04} \pm0.08$	$-0.03 \pm^{0.03}_{0.02} \pm0.05$	$0.00 \pm^{0.00}_{0.02} \pm0.03$
[Ti/H]	$0.10 \pm^{0.03}_{0.16} \pm0.08$	$0.03 \pm^{0.02}_{0.06} \pm0.05$	$0.01 \pm^{0.01}_{0.02} \pm0.03$	$0.00 \pm^{0.01}_{0.01} \pm0.02$	$0.00 \pm^{0.00}_{0.00} \pm0.01$
[V/H]	$0.13 \pm^{0.03}_{0.04} \pm0.20$	$0.05 \pm^{0.06}_{0.11} \pm0.11$	$0.01 \pm^{0.04}_{0.06} \pm0.06$	$-0.00 \pm^{0.01}_{0.02} \pm0.03$	$0.00 \pm^{0.00}_{0.01} \pm0.02$
[Cr/H]	$0.20 \pm^{0.05}_{0.04} \pm0.21$	$0.09 \pm^{0.03}_{0.03} \pm0.12$	$0.03 \pm^{0.02}_{0.02} \pm0.07$	$0.01 \pm^{0.01}_{0.02} \pm0.04$	$0.00 \pm^{0.00}_{0.00} \pm0.03$
[Mn/H]	$0.17^* \pm^{0.11}_{0.14^*} \pm0.25$	$0.04^* \pm^{0.06}_{0.10^*} \pm0.15$	$0.00^* \pm^{0.02}_{0.10^*} \pm0.10$	$-0.00^* \pm^{0.01}_{0.12^*} \pm0.09$	$0.00^* \pm^{0.00}_{0.05^*} \pm0.08$
[Fe/H]	$-0.01 \pm^{0.02}_{0.04} \pm0.04$	$0.00 \pm^{0.02}_{0.03} \pm0.03$	$0.01 \pm^{0.01}_{0.01} \pm0.02$	$0.00 \pm^{0.01}_{0.01} \pm0.01$	$0.00 \pm^{0.00}_{0.00} \pm0.01$
[Co/H]	$-0.01 \pm^{0.02^*}_{0.02} \pm0.31$	$0.00 \pm^{0.02}_{0.04} \pm0.14$	$-0.01 \pm^{0.02}_{0.13} \pm0.07$	$-0.01 \pm^{0.02}_{0.10} \pm0.04$	$0.00 \pm^{0.00}_{0.06} \pm0.02$
[Ni/H]	$-0.05 \pm^{0.05}_{0.04^*} \pm0.19$	$-0.06 \pm^{0.03}_{0.03} \pm0.12$	$-0.05 \pm^{0.03}_{0.02} \pm0.07$	$-0.01 \pm^{0.01}_{0.01} \pm0.04$	$0.00 \pm^{0.01}_{0.00} \pm0.03$
[Cu/H]	...	$0.00^* \pm^{0.00}_{0.00^*} \pm0.12$	$0.00^* \pm^{0.00}_{0.00^*} \pm0.14$	$-0.00^* \pm^{0.00}_{0.30^*} \pm0.06$	$0.00^* \pm^{0.00}_{0.23^*} \pm0.05$
[Zn/H]	...	$-0.06 \pm^{0.16}_{0.05^*} \pm0.38$	$-0.06 \pm^{0.14}_{0.16} \pm0.27$	$0.00 \pm^{0.11^*}_{0.12} \pm0.17$	$0.00 \pm^{0.08}_{0.00} \pm0.12$
[Sr/H]	...	$-0.10^* \pm^{0.03^*}_{0.14^*} \pm0.22$	$-0.18 \pm^{0.18^*}_{0.17^*} \pm0.20$	$-0.03 \pm^{0.03^*}_{0.12} \pm0.13$	$0.00 \pm^{0.00^*}_{0.00} \pm0.09$
[Y/H]	$0.09 \pm^{0.36^*}_{0.10} \pm0.24$	$0.02 \pm^{0.18}_{0.04} \pm0.15$	$-0.01 \pm^{0.10}_{0.03} \pm0.08$	$-0.01 \pm^{0.03}_{0.02} \pm0.05$	$0.00 \pm^{0.02}_{0.00} \pm0.04$
[Zr/H]	$-0.05 \pm^{0.12}_{0.14} \pm0.28$	$0.02 \pm^{0.09^*}_{0.14} \pm0.18$	$0.01 \pm^{0.06^*}_{0.10} \pm0.10$	$0.00 \pm^{0.06}_{0.05} \pm0.06$	$0.00 \pm^{0.05}_{0.00} \pm0.04$
[Ba/H]	...	$-0.14^* \pm^{0.10^*}_{0.53^*} \pm0.25$	$-0.16 \pm^{0.16^*}_{0.52^*} \pm0.19$	$-0.00 \pm^{0.00^*}_{0.39} \pm0.11$	$0.00^* \pm^{0.00^*}_{0.00} \pm0.05$
[La/H]	$-0.17^* \pm^{0.17^*}_{0.19} \pm0.35$	$-0.10^* \pm^{0.12^*}_{0.22} \pm0.15$	$-0.01^* \pm^{0.01^*}_{0.17} \pm0.07$	$-0.00 \pm^{0.00^*}_{0.09} \pm0.04$	$0.00 \pm^{0.00^*}_{0.00} \pm0.03$
[Ce/H]	$-0.16 \pm^{0.07}_{0.09} \pm0.21$	$-0.10 \pm^{0.05}_{0.06^*} \pm0.12$	$-0.06 \pm^{0.03}_{0.05^*} \pm0.09$	$-0.02 \pm^{0.02}_{0.05^*} \pm0.05$	$0.00 \pm^{0.01}_{0.00} \pm0.02$
[Pr/H]	$-0.25 \pm^{0.09}_{0.11} \pm0.31$	$-0.16 \pm^{0.06}_{0.03} \pm0.16$	$-0.07 \pm^{0.05}_{0.11} \pm0.09$	$-0.04 \pm^{0.07}_{0.12} \pm0.05$	$0.00 \pm^{0.00}_{0.04} \pm0.04$
[Nd/H]	$0.13 \pm^{0.07}_{0.33^*} \pm0.19$	$0.04 \pm^{0.05}_{0.27} \pm0.11$	$0.01 \pm^{0.02}_{0.17} \pm0.05$	$-0.00 \pm^{0.01}_{0.04} \pm0.03$	$0.00 \pm^{0.00}_{0.01} \pm0.02$
[Sm/H]	$-0.00^* \pm^{0.12^*}_{0.14} \pm0.35$	$-0.00^* \pm^{0.09^*}_{0.04} \pm0.14$	$0.03^* \pm^{0.02^*}_{0.03} \pm0.07$	$0.02 \pm^{0.05^*}_{0.02} \pm0.04$	$0.00 \pm^{0.03^*}_{0.00} \pm0.03$
[Eu/H]	...	$-0.00^* \pm^{0.00^*}_{0.02} \pm0.23$	$-0.00^* \pm^{0.02^*}_{0.03} \pm0.22$	$0.00^* \pm^{0.05^*}_{0.00} \pm0.13$	$0.00^* \pm^{0.00^*}_{0.00} \pm0.08$
[Gd/H]	$0.10^* \pm^{0.07^*}_{0.07} \pm0.38$	$0.01 \pm^{0.08}_{0.05} \pm0.22$	$-0.03 \pm^{0.06}_{0.05} \pm0.15$	$-0.01 \pm^{0.02}_{0.04} \pm0.08$	$0.00 \pm^{0.00}_{0.00} \pm0.05$
[Dy/H]	$-0.69 \pm^{0.01}_{0.01} \pm0.32$	$-0.29 \pm^{0.29^*}_{0.20} \pm0.28$	$-0.09 \pm^{0.13^*}_{0.18} \pm0.17$	$-0.02 \pm^{0.08^*}_{0.09} \pm0.11$	$0.00 \pm^{0.05}_{0.00} \pm0.07$
[Er/H]	$-0.36^* \pm^{\text{NaN}}_{\text{NaN}} \pm0.49$	$-0.00 \pm^{0.06^*}_{0.30} \pm0.41$	$0.01 \pm^{0.03^*}_{0.18} \pm0.24$	$0.01 \pm^{0.08^*}_{0.01} \pm0.13$	$0.00 \pm^{0.01^*}_{0.00} \pm0.06$
[Th/H]	...	$0.24^* \pm^{0.00^*}_{0.00} \pm0.26$	$0.18^* \pm^{0.06^*}_{0.17} \pm0.24$	$0.05 \pm^{0.11^*}_{0.05} \pm0.15$	$0.00 \pm^{0.00}_{0.00} \pm0.09$

Table 3.5: Trends in stellar label recovery with S/N. Asterisks denote instances where the reported quantities are impacted by the imposed boundaries of the training set.

to have minimal resolution-dependent systematic bias in Section 3.5.1. The presentation of these results follows the same conventions as Figures 3.9–3.19 except that we also include the 1σ statistical uncertainties inferred from MCMC sampling as blue shaded regions for reference.

We find that most of these elements show little to no dependence on the S/N down to S/N ~ 5 pixel^{-1}. While we find small differences between high and low S/N measurements, they are typically smaller than the 1σ statistical uncertainties inferred from the posteriors. The scatter found in the trends between individual exposures is generally consistent with the statistical uncertainty.

The recovery of upper/lower limits at the model grid boundary impede robust characterization of the S/N-dependence for several elements across the full S/N range, including: Sm and Th below S/N < 40 pixel^{-1}, Sr below S/N < 20 pixel^{-1}, and Gd below S/N < 10 pixel^{-1}.

For two elements, Mg and Dy, we find that the low-S/N measurements become inconsistent with the high-S/N measurements below S/N $\lesssim 10$ pixel^{-1} at which point the measurement precision is already quite poor ($\gtrsim$0.3 dex). For two other elements, C and Ca, we find more substantial trends as the S/N is decreased. For C, we find a negative bias that increases to $\sim$0.15 dex and a systematic uncertainty of $\sim$0.1 dex S/N $\lesssim 40$ pixel^{-1}. For Ca, we find a much more striking trend with S/N. Below S/N $\lesssim 40$ pixel^{-1}, [Ca/H] is recovered to be at least 0.3 dex lower than at the default S/N. The origin of these S/N-dependent systematics is challenging to ascertain and is worthy of future investigation.

3.6 Discussion

3.6.1 Fidelity of Low-Resolution Abundance Measurements

The primary motivation of this work is to identify which elements can and cannot be robustly measured from low-resolution low-metallicity stellar spectra and quantify the trends in abundance recovery exist as a function of resolution so that low- and high-resolution measurements can be integrated. Here we summarize our findings, referring to the resolution-dependent systematic biases and uncertainties for each element reported in Table 3.4. We recommend the usage of these values for 22 of the 36 elements considered in this work: C, Mg, Ca, Sc, Ti, V, Cr, Fe, Co, Ni, Zn, Sr, Y, Zr, Ce, Pr, Nd, Sm, Gd, Dy, Er, and Th. We urge caution in the adoption of these factors for the remaining elements due to inconsistent agreement with literature measurements (O, Na, Al, Si, and K; see Appendix 3.8.3.1) or the limitations set by the extent of our training grid (N, Mn, Cu, Ba, La, and Eu).

For the broad optical wavelength coverage considered in this work, we highlight $R \sim$ 10,000 as an inflection point in the trends of $\Delta\theta$ and $\sigma_{\rm syst}$, below which both quantities increase more sharply. At $R \sim$ 10,000, 20 elements (C, Mg, Ca, Sc, Ti, V, Cr, Fe, Co, Ni, Sr, Y, Zr, Ce, Pr, Nd, Sm, Gd, Dy, and Th) are recovered with $\Delta\theta \lesssim 0.1$ dex and $\sigma_{\rm syst} \lesssim 0.15$ dex. This decreases to 14 elements (C, Mg, Ca, Sc, Ti, V, Fe, Ni, Y, Zr, Ce,

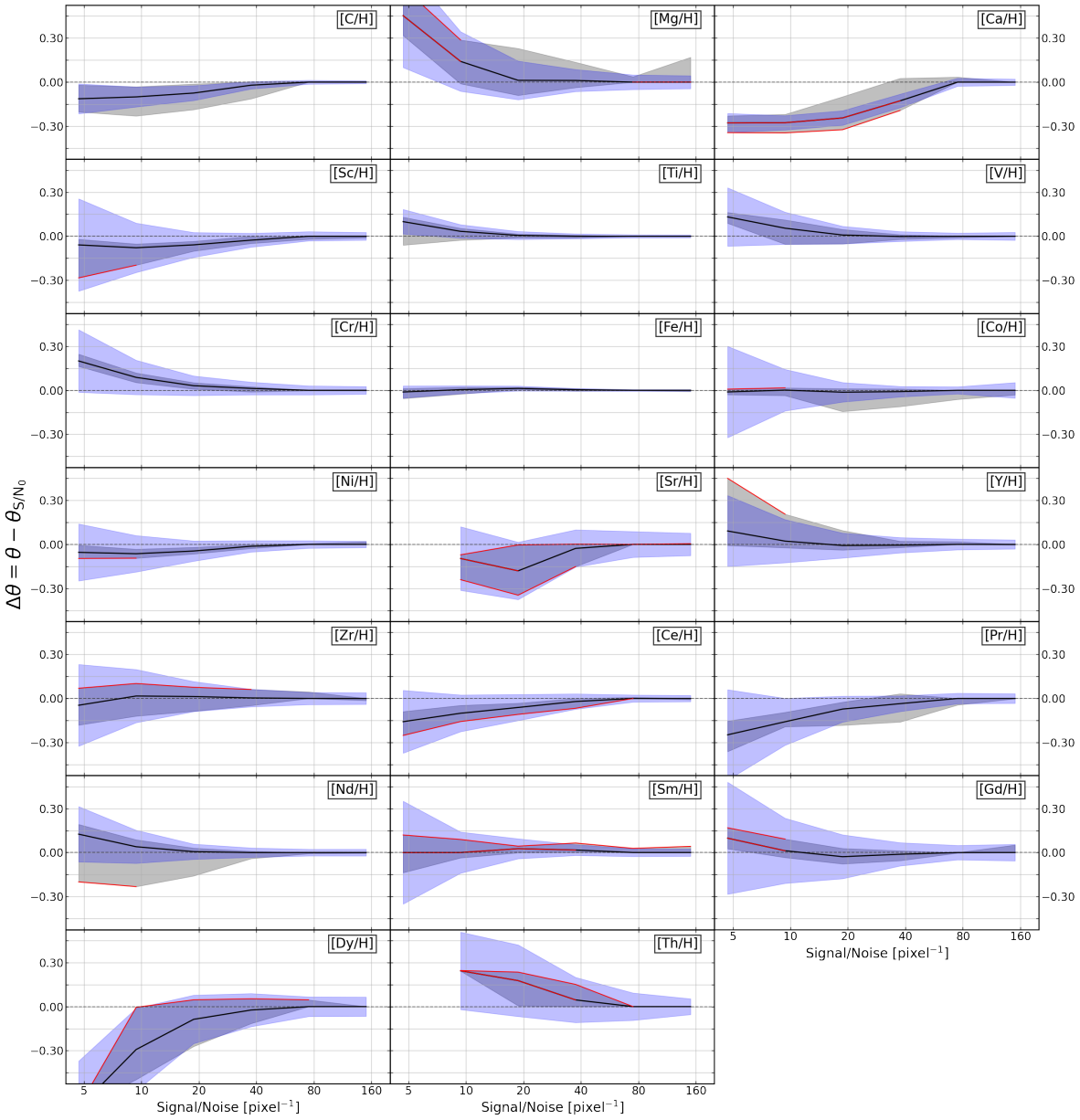


Figure 3.21: Systematic biases (solid black lines) and uncertainties (gray shaded regions) in the recovery of elements at $R \sim 10{,}000$ as a function of S/N. The median formal statistical uncertainties (blue shaded regions) are included for reference. Only elements that were found to have minimal resolution-dependent systematics in Section 3.5.1 are included. Most of these elements display S/N-dependent systematic effects that are small compared to the statistical uncertainties, though a few (C, Mg, Ca, Dy) are biased at very low S/N.

Nd, Sm, and Gd) at $R \sim 5000$ and 9 elements at $R \sim 2500$ (C, Mg, Ca, Sc, Ti, Fe, Ni, Y, and Nd). With that said, the fact that multiple individual elements—including at least one from each broad nucleosynthetic grouping no-less—can be robustly measured at $R \sim 2500$ is very promising for low-resolution surveys in the MW and LG (e.g., LAMOST, DESI, and PFS, and the low-resolution modes of SDSS-V, 4MOST, and WEAVE).

Generally speaking, the fidelity of an abundance measurement as a function of resolution is related primarily to the number (and secondarily to the strength) of absorption features for that element. Elements with many absorption lines spread across the entire spectrum tend to show the least sensitivity to model-data mismatch at low-resolution, while elements with only a few lines—especially a few weak lines—exhibit the strongest trends with resolution. This makes sense intuitively as the presence of many additional lines anchors the measurement even if some of the lines are contaminated by poorly-modeled neighboring features. When only a few lines are present, contamination of any single line can substantial bias the measurement. Similarly, we find that elements that are primarily constrained by their indirect and subtle effects on the lines of other elements (i.e., through changes to the atmospheric structure) are also sensitive to model fidelity. While these elements may still be able to be measured from low-resolution spectroscopy, much more careful treatment of the spectral features and the regions around these features is necessary.

The qualitative conclusions (e.g., which elements are more/less robustly recovered as a function of resolution) of this analysis should be broadly applicable, though the exact systematic uncertainties and biases that are reported are likely to be a strong function of the observed wavelength coverage, the observed stellar parameters, and the adopted stellar models. For example, it is not immediately clear how these results might change for a more metal-rich stellar population. On one hand, absorption features are stronger and more numerous at higher metallicity, providing additional spectral information to better anchor low-resolution measurements. By the same token, however, the increased strength and number of absorption features will lead to increased blending in the information-dense regions of the spectrum. How these two effects interact and which effect is dominant for which elements is worthy of additional investigation. Before any low-resolution "corrections" from this work are naively applied to drastically different observations (e.g., solar-metallicity dwarf stars or NIR observations), our analysis must first be extended to larger and broader datasets and spectroscopic configurations. Similarly, this study should be repeated for additional stellar models if models other than `ATLAS12` and `SYNTHE` are used.

3.6.2 Fidelity of Low-S/N Abundance Measurements

A secondary motivation of this work was to evaluate the prospect of accurately measuring multi-element abundances from low-S/N data at $R \sim 10{,}000$. Here we summarize our findings, referring to the S/N-dependent systematic biases and uncertainties for each element reported in Table 3.5. We recommend the usage of these values for the 20 elements identified in Section 3.6.1, which show only small to modest bias and uncertainties at $R \sim 10{,}000$ ($\Delta\theta \lesssim 0.1\,\mathrm{dex}$ and $\sigma_{\mathrm{syst}} \lesssim 0.15\,\mathrm{dex}$): C, Mg, Ca, Sc, Ti, V, Cr, Fe, Co, Ni, Sr, Y, Zr, Ce,

Pr, Nd, Sm, Gd, Dy, and Th. For nearly all of these elements, robust, albeit less precise, measurements can be made at S/N as low as 5 pixel^{-1} without the need to invoke additional systematic uncertainty. For four elements, C, Mg, Ca, and Dy, we find biases at very low S/N in excess of the statistical and systematic uncertainties. The origin of these trends is difficult to identify and warrants additional investigation. We recommend a minimum S/N of $\sim$10 $^{-1}$ for Mg and Dy and $\sim$40 $^{-1}$ for C and Ca.

3.6.3 Stellar Label Uncertainties and Correlations

The use of MCMC methods in our spectroscopic analysis enables us to robustly quantify the formal statistical uncertainties on measurements of [X/H] as well as the element-to-element measurement correlations. In Figure 3.22, we present the median pairwise correlations found between the 33 measured elemental abundances, v_r, $v_{\rm macro}$, and $v_{\rm micro}$ at the convolved resolution of $R \sim$ 10,000. $T_{\rm eff}$ and $\log g$ are omitted as they are 1-to-1 correlated with Fe. Each panel depicts the correlation of a label pair as measured from the MCMC posterior samples. To guide the eye, panels are shaded according to their Pearson correlation coefficient, r.

Figure 3.22 shows that the majority of stellar labels are not strongly correlated ($r \lesssim 0.05$) at $R \sim$ 10,000. The strongest correlations belong to elements which have many absorption features across the observed wavelength range. Most obvious among these is Fe, which is strongly anti-correlated ($r \sim -0.1$ to -0.6) with $\sim$20 other stellar labels. C, and to a lesser extent Mg, Si, and Ti, also exhibit correlations of $r \gtrsim 0.05$ with roughly a dozen other elements as a result of their contributions to stellar atmospheric structure. We also compare the correlations we infer at $R \sim$ 10,000 with those that we infer at both lower and higher resolutions. We find that the pairwise correlation between elements at $R \sim$ 40,000 is very similar to what we find at $R \sim$ 10,000. At $R \sim$ 2500, the pairwise correlation increases in magnitude for most elements and for $v_{\rm macro}$, though it still remains below $r \lesssim 0.2$ for most element pairs.

3.6.3.1 Comparison of Uncertainties to CRLBs

Calculating uncertainties and correlations using MCMC sampling is a computationally expensive undertaking, especially given the high-dimensionality of abundance measurements explored here. As a result, its application to the datasets of large spectroscopic surveys (e.g., APOGEE, GALAH, LAMOST) are intractable. Recently, the use of Cramér-Rao Lower Bounds (CRLBs; Fréchet 1943; Darmois 1945; Rao 1945; Cramer 1946), the maximum precision predicted by a Fisher Information analysis, has been proposed as a fast and easy method to forecast the chemical abundance precision achievable from a given stellar spectral dataset (e.g., Ting, Conroy, Rix, and Cargile 2019; Sandford, Weisz, et al. 2020). Here we take the opportunity to compare the statistical uncertainties we measure from our MCMC fitting technique to those forecasted by the CRLBs.

To calculate the CRLBs of our observations, we employ the `Chem-I-Calc`[13] Python pack-

[13]https://chem-i-calc.readthedocs.io/en/latest/

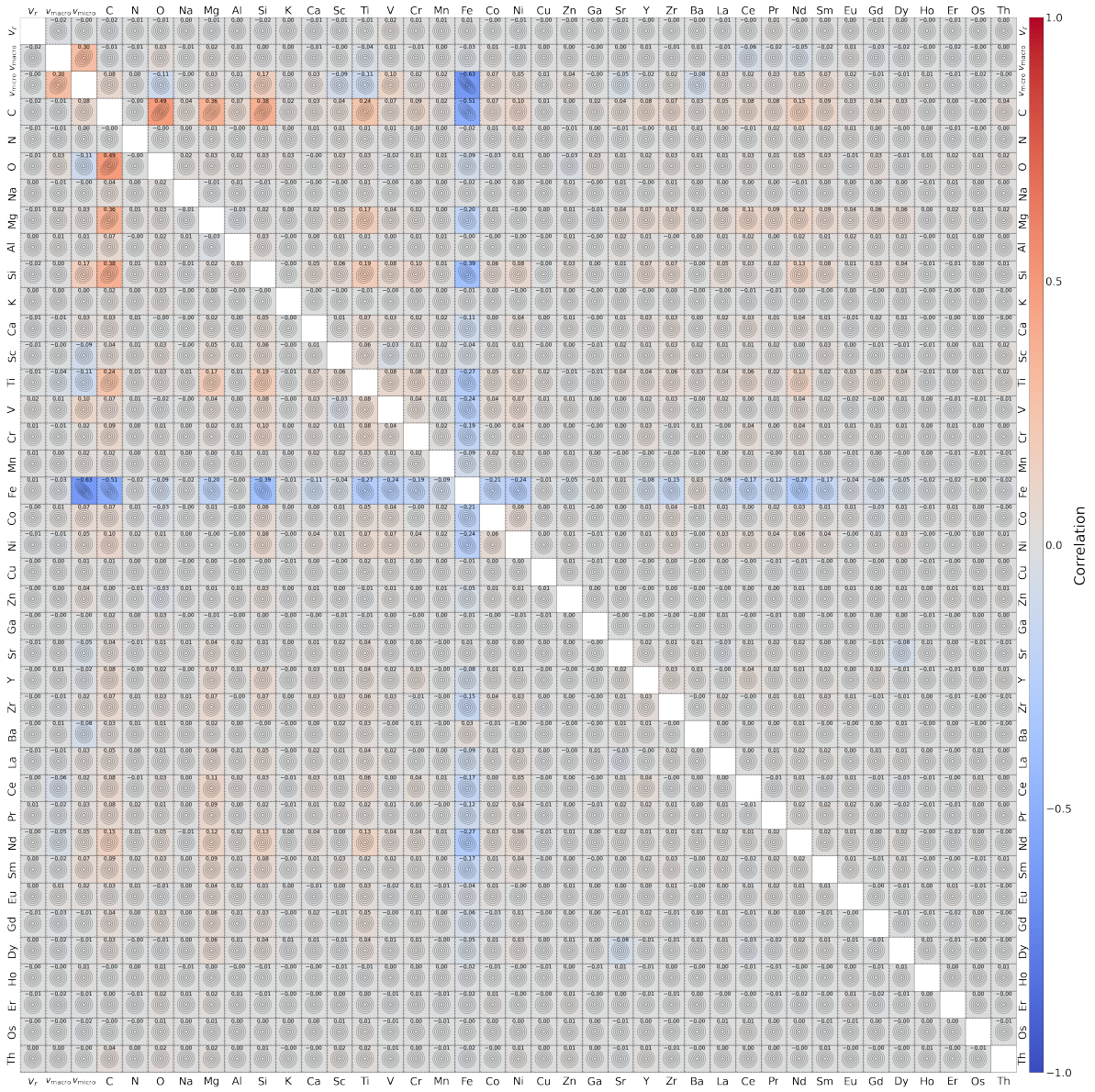


Figure 3.22: Median correlations in the measurements of all 33 elemental abundances, v_r, $v_{\rm macro}$, and $v_{\rm micro}$ at $R \sim 10{,}000$. Each panel depicts the correlation of a different pair of labels with the color of the panel indicating the strength and direction of the correlation. While most labels are not strongly correlated with one other, labels that contribute to a large number of pixels across the observed wavelength range like Fe, C, Mg, Si, and Ti exhibit modest correlations.

age with a few minor adjustments (Sandford 2020; Sandford, Weisz, et al. 2020). For each star in our sample, we generate gradient spectra using `the Payne` and adopt the total S/N with both model errors and observational masks included. Because $T_{\rm eff}$ and $\log g$ are inferred deterministically from stellar photometry and [Fe/H], we treat them as fixed parameters in the CRLB calculation.

In Figure 3.23, we present a comparison of the statistical uncertainty found through MCMC sampling, $\sigma_{\rm MCMC}$, and the uncertainty forecasted by the CRLB, $\sigma_{\rm CRLB}$, for each element. Points and error bars represent the median and 16th and 84th percentiles across all individual exposure measurements performed in this study, omitting measurements which are within 2σ of the uniform prior bounds and measurements for which $\sigma_{\rm CRLB} > 0.5$. For 28 (24) of the 33 chemical abundances, $\sigma_{\rm CRLB}$ is within 20% (10%) of $\sigma_{\rm CRLB}$. We find no trend in the (dis)agreement as a function of the resolution or S/N of the observation, nor as a function of the expected precision.

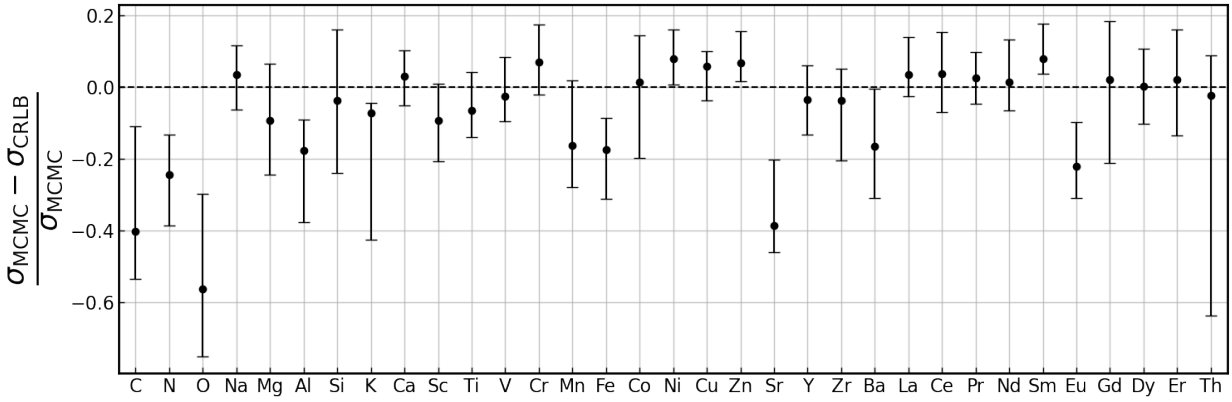


Figure 3.23: Fractional difference in the formal statistical uncertainty on [X/H] and the precision forecasted by the CRLB. Points and error bars represent the median and 16th and 84th percentiles across all individual exposure measurements performed in this study. For ∼85% of the elements considered here, the uncertainties are in general agreement. Large deviations from zero are found in instances of non-Gaussian posteriors (e.g., C, N, and O) and/or substantial model-data mismatches (e.g., Sr and Eu).

Because the CRLB represents the maximum theoretically achievable precision, we would expect $\sigma_{\rm MCMC} \gtrsim \sigma_{\rm CRLB}$. While this is the case for many elements and measurements, it is not universally true. For example, the statistical uncertainties on Fe and Eu are consistently ∼20% smaller than forecasted by the CRLBs. C, N, O, Al, K, Sr, and Ba are also recovered to better precision than the CRLBs predict—in some cases by large margins. These deviations from the forecasted precision are driven by 1) non-Gaussian posteriors for which $\sigma_{\rm MCMC}$ underestimates the true uncertainty and/or 2) mismatches between the model and observed spectra that invalidate the assumption of an un-biased estimator in the CRLB calculation

(e.g., Ting, Conroy, Rix, and Cargile 2017; Sandford, Weisz, et al. 2020). In the case of C, N, and O, we believe that the better-than-expected precision is due in part to non-Gaussian posteriors and in part to over-estimation of the correlation between these three elements in the CRLBs. Indeed, if the correlation in CNO spectral features is ignored in the CRLB calculation, the agreement between the forecasted and realized statistical uncertainties is much better (though large variance remains).

Given the general agreement found in this comparison, and no instances of the CRLB drastically over-predicting the expected precision, we suggest that, going forward, CRLBs can be safely adopted as conservative forecasts of the statistical precision to the 10–20% level.

3.6.4 Implications for Chemical Evolution Studies

A primary use of stellar chemical abundance measurements is to constrain stellar and galactic physics by fitting models to the chemical enrichment history of a system. Recovering meaningful constraints, however, requires accurate abundances with well-characterized uncertainties. Anything less will lead to biased or misleading conclusions. High-precision, inaccurate measurements are perhaps the most disastrous combination as they will strongly influence a chemical evolution model from the true solution. Less catastrophic, but still undesirable are accurate measurements with uncertainties that are under-predicted as these will lead to accurate model predictions but will overstate the constraining power of the data. Moreover, mischaracterized uncertainties will bias studies concerned with the intrinsic dispersion of stellar chemical abundances (e.g., to understand stochasticity in nucleosynthetic pathways or inhomogenous mixing of the ISM; see Griffith et al. (2023) and Ting and Weinberg (2022)).

For these reasons, it is important to fold in accurate estimates of the systematic uncertainties like those presented in this work. For most low-resolution stellar spectroscopic observations, this precludes the $\lesssim$0.1 dex precision on many elements that higher resolution surveys can achieve. As a result, the vast majority of chemical abundance measurements in the next decade, especially for stars outside the MW, will be systematics limited in precision. Nevertheless, low-precision (0.2–0.3 dex) measurements can still be incredibly informative as long as they are accurate and there are sufficient numbers of stars (e.g., Kirby, Cohen, Smith, et al. 2011; Sandford, Weinberg, et al. 2022). Even after appropriately accounting for the systematic uncertainties quantified in this work, the highly-multiplexed, low-resolution spectrographs of the next decade have the potential to reveal transformative new insight into galactic chemical evolution in the MW and throughout the LG and beyond.

3.7 Summary

We perform a completely self-consistent analysis of 40 Keck/HIRES observations of 8 metal-poor RGB stars in M15, which have been degraded to a range of lower resolutions and S/N.

We fit for 36 stellar labels (including 33 elemental abundances) and ∼100–200 nuisance parameters (mostly continuum coefficients) using full-spectrum fitting techniques and quantify the systematic biases and uncertainties that are introduced as the quality of the data is degraded. Our primary conclusions are as follows:

1. Observations at resolutions down to $R \sim 10{,}000$ can measure 20 elements (C, Mg, Ca, Sc, Ti, V, Cr, Fe, Co, Ni, Sr, Y, Zr, Ce, Pr, Nd, Sm, Gd, Dy, and Th) to within $\lesssim$0.1 dex of high-resolution observations with $\lesssim 0.15$ dex systematic uncertainties.

2. Nine elements (C, Mg, Ca, Sc, Ti, Fe, Ni, Y, and Nd) can be measured to this same level of consistency down to $R \sim 2500$.

3. Only four elements (C, Mg, Ca, and Dy), exhibit substantial S/N-dependent bias at $R \sim 10{,}000$ in excess of statistical uncertainties below S/N ~ 10 pixel^{-1}.

4. For ∼85% of elements, the precision forecasted by the CRLBs provides a good estimate of the formal uncertainties computed with MCMC sampling.

5. The predominant source of systematic bias and uncertainty at low-resolution is blending of poorly-modelled absorption features, which impacts elements with few and/or weak lines most strongly.

6. Applying this methodology to a larger sample of stars is necessary to extend these conclusions more broadly to a wider range of stellar types and metallicities.

We conclude with an optimistic outlook. In this work we find that even with imperfect models, low-resolution measurements that are consistent with their high-resolution counterparts are possible for a representative sample of elements. As such, the next decade of highly-multiplexed low-resolution spectroscopic surveys and instruments are poised to dramatically increase our understanding of the MW and LG's chemical evolution. Furthermore, because we have adopted 1D-LTE stellar models in this analysis, the systematic effects we report represent a conservative estimate. Ongoing improvements to stellar models (e.g., 3D-NLTE physics, updated atomic line data), will continue to alleviate these systematics and further increase the viability of high-precision accurate low-resolution spectroscopic chemical abundance measurements.

3.8 Appendices

3.8.1 Appendix: `The Payne` - Technical Details and Training

3.8.1.1 Appendix: Neural Network Architecture

As in previous implementations of `the Payne`, we adopt a fully-connected neural network with two hidden layers of $N_1 = N_2 = 300$ neurons each. The first hidden layer expects as

input an array of $N_\theta = 39$ stellar labels ($T_{\rm eff}$, $\log g$, $v_{\rm micro}$, and [X/H], where X includes the elements C, N, O, Na, Mg, Al, Si, K, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Sr, Y, Zr, Ba, La, Ce, Pr, Nd, Sm, Eu, Gd, Dy, Ho, Er, Os, and Th).[14] The output of the model is an array of normalized flux values corresponding to each wavelength pixel of the *ab initio* spectra it is trained on. Employing a leaky ReLU activation function,

$$\mathrm{LReLU}(x) = \begin{cases} x, & \text{if } x > 0 \\ 0.01x, & \text{otherwise,} \end{cases} \tag{3.23}$$

the model architecture can be represented by the following equations:

$$f_j^{(1)}(\theta_*) = \mathrm{LReLU}\left(\sum_{i=1}^{N_\theta}\left[w_{i,j}^{(1)}\theta_{*,i}\right] + b_j^{(1)}\right) \tag{3.24}$$

$$f_k^{(2)}(\theta_*) = \mathrm{LReLU}\left(\sum_{j=1}^{N_1}\left[w_{j,k}^{(2)} f_j^{(1)}(\theta_*)\right] + b_k^{(2)}\right) \tag{3.25}$$

$$f_\lambda^{(\mathrm{out})}(\theta_*) = \sum_{k=1}^{N_2}\left[w_{k,\lambda}^{(\mathrm{out})} f_k^{(2)}(\theta_*)\right] + b_\lambda^{(\mathrm{out})}, \tag{3.26}$$

where $w^{(1)}$, $b^{(1)}$, $w^{(2)}$, $b^{(2)}$, $b^{(\mathrm{out})}$, and $w^{(\mathrm{out})}$ are the weights and biases of the neurons in the first hidden layer, second hidden layer, and the output layer. Like later implementations of `the Payne` (e.g., Kovalev et al. 2019; Straumit et al. 2022; Xiang, Rix, et al. 2022), this architecture capitalizes on the continuity of the spectrum in the wavelength dimension and utilizes the information contained in adjacent pixels to better predict the flux of each pixel—in contrast to the architecture used originally in Ting, Conroy, Rix, and Cargile (2019), which used an independent model for each pixel.

The total number of model parameters in a neural network with this architecture is given by

$$N_{\rm par} = (N_\theta + 1) \times N_1 + (N_1 + 1) \times N_2 + (N_2 + 1) \times N_{\rm pix}, \tag{3.27}$$

where $N_{\rm pix}$ is the number of pixels in the model spectrum. Adopting $N_1 = N_2 = 300$ and training on *ab initio* spectra with $N_{\rm pix} = 262{,}144$ with $N_\theta = 39$ as we do, requires a model with $N_{\rm par} \sim 7.9 \times 10^7$ parameters. Despite the large number of parameters to optimize, such a model can be optimized in a reasonable ∼150 hours on a NVIDIA A40 GPU.

3.8.1.2 Appendix: Training Set

Training `the Payne` requires a set of stellar spectra with known labels that span the parameter space of the observed stars. Because the stars considered in this work have been well studied, we could generate a dense training set around the literature values for these

[14] Due to lack of spectral information, we omit Ga, Ho, and Os from further analysis in this paper.

MIST version	1.2
Initial $v/v_{\rm crit}$	0.4
$t_{\rm age}$	10 to 14 Gyr
$\Delta \log t_{\rm age}$	0.01
[Fe/H]	-4.0 to -1.0
Δ[Fe/H]	0.1
[α/H]	0.0

Table 3.6: Characteristics of the MIST isochrone set from which $T_{\rm eff}$, $\log g$, and [Fe/H] are initially drawn.

stars ([Fe/H] ~ -2.5 at the tip of the RGB). However, we choose to generate a much more ambitious training set that covers the entire RGB over a large range of metallicities. The reasons for this are twofold: 1) to avoid simply reproducing literature results by construction and 2) to generate a training set with applications beyond the RGB of M15.

We begin the construction of our training set by randomly drawing 25,000 sets of $T_{\rm eff}$, $\log g$, and [Fe/H] values from MIST isochrones (Paxton, Bildsten, et al. 2011; Paxton, Cantiello, et al. 2013; Paxton, Marchant, et al. 2015; Choi et al. 2016; Dotter 2016; Paxton, Schwab, et al. 2018) with $3500 \leq T_{\rm eff}$ K ≤ 6000, $0.0 \leq \log g \leq 4.0$, $-4.0 \leq$ [Fe/H] ≤ -1.0, and $10 \leq t_{\rm age}$ [Gyr] ≤ 14. Only RGB stars are included in this sample. For each sample, $v_{\rm micro}$ is determined from the empirical relation found in Holtzman et al. (2015),

$$v_{\rm micro} = 2.478 - 0.325 \log g. \tag{3.28}$$

To smooth over the discrete isochrone tracks and allow for $v_{\rm micro}$ offset from the empirical relation, we add zero-mean Gaussian scatter to each of these labels with $\sigma_{T_{\rm eff}} = 250$ K, $\sigma_{\log g} = 0.25$, $\sigma_{\rm [Fe/H]} = 0.25$, and $\sigma_{v_{\rm micro}} = 0.25$ km s^{-1}. Lastly, for each sample, we draw elemental abundances [X/H] from a uniform distribution with the condition $-1.0 \leq$ [X$_1$/Fe] ≤ 1.0 for X$_1$ = C, N, and O; $-0.5 \leq$ [X$_2$/Fe] ≤ 0.5 for X$_2$ = Na, Sc, V, Cr, Mn, Co, Ni, Cu, Zn, Ga, Sr, Y, Zr, Ba, and La; and $-0.25 \leq$ [X$_3$/Fe] ≤ 1.0 for X$_3$ = Mg, Al, Si, K, Ca, Ti, Ce, Pr, Nd, Sm, Eu, Gd, Dy, Ho, Er, Os, and Th. Summaries of our MIST isochrones and sampling scheme are presented in Tables 3.6 and 3.7 respectively. We note that while 25000 *ab initio* may seem like a large training set, it is still orders of magnitude smaller than would be required for grid interpolation over the broad 39-dimensional parameter space.

ab initio spectra are generated using the same method described in Ting, Conroy, Rix, and Cargile (2019), which we summarize here. For each of the 25,000 sets of stellar labels, we compute 1D LTE model atmospheres using the `ATLAS12` code maintained by R. Kurucz (Kurucz 1970; Kurucz and Avrett 1981; Kurucz 1993, 2005, 2013, 2017). We adopt Solar abundances from Asplund et al. (2009) and the standard mixing length theory with a mixing length of 1.25 and no overshoot for convection. After the model atmosphere converges, we use the `SYNTHE` radiative transfer code (also maintained by R. Kurucz) to produce its normalized spectrum at a nominal resolution of $R = 300{,}000$.

Label	Distribution
Intermediate Samples from MIST Isochrones	
$T_{\rm eff,\ iso}$	$\mathcal{U}_{\rm MIST}(3500\ {\rm K},\ 6000\ {\rm K})$
$\log g_{\rm iso}$	$\mathcal{U}_{\rm MIST}(0.0,\ 0.4)$
$[{\rm Fe/H}]_{\rm iso}$	$\mathcal{U}_{\rm MIST}(-4.0,\ -1.0)$
Final Samples with Scatter	
$T_{\rm eff}$	$\mathcal{N}(T_{\rm eff,\ iso},\ 250\ {\rm K})$
$\log g$	$\mathcal{N}(\log g_{\rm iso},\ 0.25)$
[Fe/H]	$\mathcal{N}([{\rm Fe/H}]_{\rm iso},\ 0.25)$
$v_{\rm micro}$	$\mathcal{N}(2.478 - 0.325 \log g,\ 0.25)$
$[{\rm X_1/Fe}]$	$\mathcal{U}(-1.00,\ 1.00)$
$[{\rm X_2/Fe}]$	$\mathcal{U}(-0.50,\ 0.50)$
$[{\rm X_3/Fe}]$	$\mathcal{U}(-0.25,\ 1.00)$

Table 3.7: Distributions from which the training label sets are drawn. $T_{\rm eff}$, $\log g$, and [Fe/H] are drawn initially from the MIST isochrone set described in Table 3.6 before additional scatter is applied. X_1 includes C, N, and O. X_2 includes Na, Sc, V, Cr, Mn, Co, Ni, Cu, Zn, Ga, Sr, Y, Zr, Ba, and La. X_3 includes Mg, Al, Si, K, Ca, Ti, Ce, Pr, Nd, Sm, Eu, Gd, Dy, Ho, Er, Os, and Th.

For a little less than ∼20% of the labels, the stellar atmosphere and/or spectrum fails to converge. These failed models predominantly belong to stellar atmospheres with very low metallicities ([Fe/H] $\lesssim -3.0$). It is possible that better initialization of low-metallicity atmospheres might improve convergence, but we leave this to a future study.

The ∼20,500 successfully generated spectra are then continuum normalized using the theoretical continua from `SYNTHE`. Lastly, the normalized spectra are convolved and sub-sampled down to the highest spectral resolution and wavelength sampling present in our archival data ($R = 86{,}600$ and ${\rm d}v = 1.17\,{\rm km\,s^{-1}}$ pixel^{-1}).

3.8.1.3 Appendix: Training Procedure

We implement our adoption of `the Payne` using `PyTorch`, a powerful and flexible Python machine learning framework, and `Pytorch Lightning`[15], a lightweight wrapper designed to streamline the development and training of PyTorch models.

As with many machine learning techniques, it is helpful to scale the input labels so that they all share a similar dynamic range of order unity with zero mean. To do so, we normalize all stellar labels according to

$$\theta'_{*,i} = \frac{\theta_{*,i} - \theta_{*,i,\rm min}}{\theta_{*,i,\rm max} - \theta_{*,i,\rm min}} - 0.5, \tag{3.29}$$

[15] `https://www.pytorchlightning.ai/`

where $\theta_{*,i,\min}$ and $\theta_{*,i,\max}$ are the minimum and maximum values of each label, i, included in the training set. For clarity, we drop the prime notation throughout the rest of this work and convert back to physical units when reporting results.

We train the model directly on 80% of the successfully generated spectra ($\sim$16,000) and validate with the remaining 20% ($\sim$4000). Training is performed iteratively in batches of 512 spectra using a rectified Adam optimizer (with a learning rate of 10^{-4}) to minimize the L1 loss (i.e., the mean absolute error). Though unknown to the optimization, the L1 loss is also calculated on the cross-validation dataset each epoch. Training is halted after 2000 epochs without improvement of the best L1 validation loss, at which point the model that minimized the L1 validation loss is chosen as the final model. Training of the model was completed in $\sim$150 GPU hours after $\sim$24,000 epochs.

3.8.1.4 Appendix: Accuracy

We determine the internal accuracy (i.e., the median interpolation error; MIE) of `the Payne` as is in (Ting, Conroy, Rix, and Cargile 2019). Using the trained neural networks, we generate spectra for each set of stellar labels in the cross-validation dataset and compare to the original *ab initio* spectra generated with `ATLAS12` and `SYNTHE`. The median interpolation error is thus

$$\sigma_{\rm MIE} = {\rm Med}(|f_\lambda(\theta_{*,\rm valid}) - f_{\lambda,\rm valid}|). \tag{3.30}$$

Figure 3.24 graphically presents how accurately `the Payne` interpolates the synthetic spectra. In the top left panel, we show the distribution of interpolation errors for our cross-validation set, taking the median over all wavelength pixels. We find that for $\sim$85% of spectra, the MIE is $\lesssim$0.1%, though the long tail of the distribution indicates that some spectra have errors as high as $\sim$1%. Unlike Ting, Conroy, Rix, and Cargile (2019) who find larger MIE for cooler stars, we find that this long tail towards higher errors corresponds to stars with higher metallicities (Fe/H > -2; red histogram). Because the stars analyzed in this paper all have [Fe/H] $\lesssim -2.4$, the adopted MIE is likely on the conservative end.

In the bottom panel, we show the pixel-by-pixel MIE for the entire wavelength range of the model, taking the median over all cross-validation spectra; this is the $\sigma_{\rm MIE}$ adopted in Equation 3.7. Typical pixel-by-pixel MIEs for the `the Payne` are 0.01–0.1% for $\lambda > 4500$ and 0.1–0.5% for $\lambda < 4500$. The MIE is generally larger in the blue due to the higher density of absorption lines and the presence of complicated molecular features. The MIE is also larger in the proximity of strong absorption features like the the Balmer lines. We believe this to be the reason why higher-metallicity stars in our cross-validation set have larger MIE than the lower-metallicity stars. The results over all wavelength pixels is summarized in the top right panel, which shows the cumulative number of wavelength pixels as a function of MIE. Roughly 80% of pixels have $\sigma_{\rm MIE} < 0.001$, and 95% of pixels have $\sigma_{\rm MIE} < 0.006$.

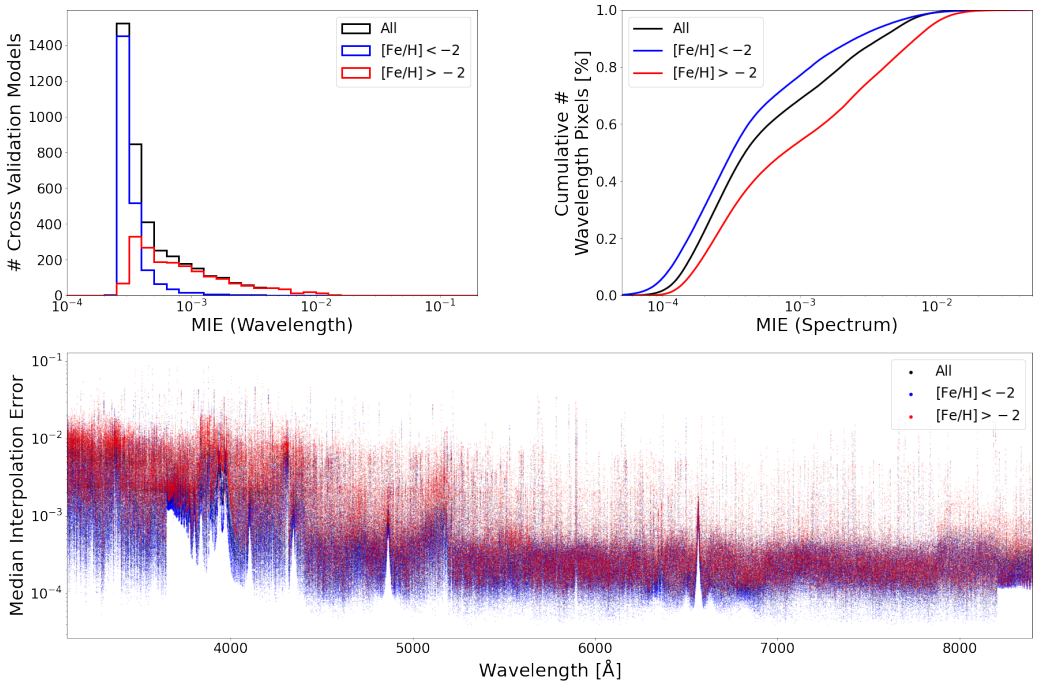


Figure 3.24: Top Left: Histogram of the median interpolation error of each model in the cross validation set. The median error is consistently larger for higher-metallicity stars ([Fe/H] > -2; red) compared to lower-metallicity stars ([Fe/H] > -2; blue). Top Right: Cumulative percentage of pixels in each spectrum as a function of the median interpolation error. Approximately 80% of pixels have $\sigma_{\rm MIE} < 0.001$, and 95% of pixels have $\sigma_{\rm MIE} < 0.006$. Bottom: The median interpolation error across the cross-validation set as a function of wavelength. Errors are largest in the proximity of strong H lines and complicated molecular features.

3.8.2 Appendix: Fitting Routines

3.8.2.1 Appendix: Optimization

Using the automatic differentiation engine of `Pytorch` and the Adam optimization algorithm, we minimize the negative log-posterior, which is equivalent to maximizing Equation 3.15. The optimization is performed 10 times with unique initializations. The parameters from the trial with the highest log-posterior value after convergence are taken as the best-fit optimization values. In the rest of this section, we present our choices of initialization,

Label	Initialization	Learning Rate	Decay	Timescale	Tolerance
θ'_*	Prior	1e-1	0.99	10	1e-5
$\log_{10} v_{\rm macro}$	Prior	1e-1	0.99	10	1e-4
v_r	Grid Search	1e-3	0.9	10	1e-4
$c_{o,n}$	`np.polyfit`	1e-1	0.99	10	5e-2

Table 3.8: Initialization procedure and optimization hyper-parameters for the stellar labels and nuisance parameters of our model.

learning rates, and convergence criteria. These choices are summarized in Table 3.8.

Initialization For each initialization, we begin by defining a fiducial model spectrum with the mean stellar labels of the training set (i.e., $\theta'_{*,i} = 0$) and the appropriate resolution for the observations. No Doppler shift, macroturbulent broadening, or continuum correction (other than the blaze function) is applied . Using this fiducial spectrum, the radial velocity is then initialized via a grid search from -300 to $300\,{\rm km\,s^{-1}}$ in steps of $2\,{\rm km\,s^{-1}}$ to the value that minimizes the negative log-posterior. The polynomial continuum coefficients are then initialized by performing a polynomial fit with `np.polyfit` to the ratio of the observed spectrum and the (now Doppler-shifted) fiducial spectrum. All other labels are initialized by randomly sampling from their priors.

Learning Rates We find learning rates of 0.1 to work well for all labels but the radial velocity, which requires a much smaller learning of 0.001 due to the sensitivity of $P(\Theta|D)$ to v_r at high resolution. To improve convergence, the learning rates are decayed every 10 step by a multiplicative factor of 0.9 for v_r and 0.99 for all other labels.

Convergence Convergence of the optimization is achieved when the change in all model parameters is below a given threshold. We define this tolerance to be 10^{-5} for the scaled stellar labels, θ'_*, and 10^4 for both $\log_{10} v_{\rm macro}$ and v_r.

We define the convergence criteria for the continuum coefficients slightly differently. Instead of imposing a threshold on the change in $c_{n,o}$, we require that at every pixel (excluding masked pixels) the value of the continuum polynomial changes by less than 5% of the observed flux uncertainty.

3.8.2.2 Appendix: MCMC Sampling

Using the affine invariant MCMC methods introduced by Goodman et al. (2010) and implemented in `emcee`[16], we sample our log-posterior distribution (Equation 3.15). As in the optimization routine, the log-posterior is evaluated using scaled stellar labels, θ'_*, which are

[16] `https://emcee.readthedocs.io/`

converted to physical units when results are reported. In the rest of this section, we present the specifics of our sampling routine.

Initialization Before sampling the posterior in earnest, we begin by initializing 128 walkers at the maximum a posteriori value of Θ found via our optimization algorithm. Gaussian scatter of 0.1 is applied to all labels except the continuum coefficients, which are held constant throughout MCMC sampling. During this burn-in phase, the walkers sample the posterior until 1) the mean value for each label of the walkers changes less than 0.5% over the previous 100 steps and 2) the mean $\log P$ of the walkers has changed by less than 10^{-5}%. After the burn-in phase is complete, 512 walkers are initialized around the location of the walker with the highest $\log P$ with a Gaussian scatter in each label equal to half the standard deviation of the burn-in walkers for that label. Now that the initialization is complete, and the walkers have had a chance to settle around the maximum a posteriori, we begin the production run of our posterior sampling.

Convergence We sample the posterior distribution until the following two convergence criteria have been met: 1) the auto-correlation time, τ, has changed by <1% over the previous 100 steps and 2) the sampler has run for $>30\tau$ steps. If these criteria have not been met after 15,000 steps, walkers are re-initialized around the location of the walker with the highest $\log P$, and the sampling is restarted. Once convergence has been reached, we discard the first 5τ samples from each walker and thin the samples of each chain by $\sim \tau/2$ to remove any residual effects of burn-in or correlated samples. Unthinned chains can be made available upon request.

Move Proposal The default move proposal of `emcee` is the "stretch move" method of Goodman et al. (2010), which is not well suited for the dimensional of our problem. Instead, we adopt a weighted mixture of 80% differential evolution proposals (`emcee.moves.DEMove`; Braak et al. 2006; Nelson et al. 2014) and 20% differential evolution snooker proposals (`emcee.moves.DESnookerMove`; ter Braak et al. 2008). We find that this combination of move proposals improves the convergence time by more than an order of magnitude over the default move proposal.

3.8.3 Appendix: Comparison with Literature Values

As a check on our fitting procedure, we compare our default high-resolution high-S/N stellar label measurements to those measured from previous stellar spectral analyses of the same stars. Our high-resolution measurements and those included in the literature comparison are presented in Table 3.9. Their chemical abundances have been adjusted to place them on the Asplund et al. (2009) Solar abundance scale adopted in our analysis. This collection of literature measurements is a non-exhaustive, but representative sample of previous spectro-

scopic studies in M15 across a wide range of spectroscopic configurations, fitting techniques, and measured elements.

In Figure 3.25 we present our measurements of $T_{\rm eff}$, $\log g$, and [Fe/H] (black stars) alongside literature measurements (colored circles) for the 8 stars in our sample. Error bars represent 1σ uncertainties where available and are too small to be visible for the statistical uncertainty of our own measurements.

Overall, our measurements fall nicely amidst the locus of literature measurements. On average, our measurements differ from the median literature value for $T_{\rm eff}$, $\log g$, and [Fe/H] by roughly 100 K, -0.1 dex, and -0.15 dex respectively. The most outlying atmospheric parameters are recovered for star K934, for which we recover values of $T_{\rm eff}$, $\log g$, and [Fe/H] to be $\sim$300 K, 0.5 dex, and 0.3 dex lower respectively than measured from APOGEE spectra by Masseron et al. (2019) and Jönsson et al. (2020). We believe that the proximity of K934 to K731 in the CMD (Figure 3.2) justify our measurements, which would place the two stars similarly close in $T_{\rm eff}$-$\log g$ space. We also note that two APOGEE studies also consistently recover [Fe/H] higher than most other studies, perhaps owing to the difference in wavelength coverage (NIR versus optical).

In Figure 3.26, we present a comparison of our measurements with measurements from the same literature studies for the remaining 32 chemical abundances. The same symbol schema is adopted as in Figure 3.25. When separate abundances are provided for neutral and ionized atomic species (e.g., [Ti I/Fe I] versus [Ti II/Fe II]), the abundances for the ionized species are represented with open circles.

In brief, we find generally good agreement with the literature for measurements of C, N, O, Mg, Ca, Ti, V, Cr, Mn, Co, Ni, Cu, Y, Zr, La, Ce, Pr, Nd, Gd, Dy, Er, and Th with a few caveats. For 5 elements, our measurements are systematically offset $\sim$0.25 dex higher (Sc, Ba, Sm, and Eu) or lower (Zn) compared to literature measurements. Mixed or poor agreement with values was found for (Na, Al, Si, and K). In general, differences between our abundance measurements and those from the literature can be attributed to differences in the stellar models, line lists, oscillator strengths, and/or wavelength coverage employed. Across all elements, our measurements agree best with literature measurements made over similar wavelength ranges (e.g., the optical; Sneden, Kraft, et al. 1997; Sobeck, Kraft, et al. 2011) and less well with those made over non-overlapping spectral coverage (e.g., the NIR; Mészáros et al. 2015; Masseron et al. 2019; Jönsson et al. 2020). As in Section 3.5, we present a detailed comparison with the literature for each element loosely grouped by nucleosynthetic origin.

3.8.3.1 Appendix: Literature Comparison by Element Group

Iron-Peak Elements In addition to Fe, we find generally good agreement with the literature for iron-peak elements V, Cr, Co, and Ni. While we recover only upper limits on Mn, these upper limits are generally consistent with literature values. The one exception to our general agreement with the literature is for K934, for which Jönsson et al. (2020) reports [Mn/Fe] $= 0.68 \pm 0.15$. Because this is the only reported measurement of [Mn/Fe] in K934,

Reference	$T_{\rm eff}$ [K]	$\log g$	[Mg/Fe]	[Ca/Fe]	[Fe/H]	...
			K341			
This Study	4415	0.78	0.37	0.04	−2.47	...
S+97	4275	0.45	0.72	0.32	−2.35	...
S+00b	4275	0.45	...	0.56	−2.45	...
S+06	4275	0.45	...	...	−2.46	...
R+09	...	...	...	...	−2.32	...
C+09b	4324	0.69	0.49	...	−2.23	...
S+11	4343	0.88	0.60	0.22	−2.53	...
W+13	4324	0.69	...	0.16	−2.32	...
K+18	4253	0.67	0.66	0.28	−2.49	...
M+19	4545	0.80	0.27	0.17	−2.08	...
J+20	4377	0.64	0.39	0.28	−2.30	...
			K386			
This Study	4390	0.54	0.24	0.01	−2.49	...
S+97	4200	0.15	...	0.19	−2.43	...
S+00b	4200	0.15	...	0.19	−2.35	...
O+06	4200	0.35	...	...	−2.40	...
S+06	4200	0.15	...	...	−2.51	...
C+09a	4313	0.65	...	...	−2.33	...
W+13	4313	0.65	...	0.10	−2.33	...
K+18	4263	0.65	0.15	0.19	−2.50	...
M+19	4548	0.81	0.28	0.51	−2.14	...
J+20	4449	0.90	...	...	−2.08	...
			K431			
This Study	4489	0.78	0.16	0.15	−2.45	...
S+97	4375	0.50	0.38	0.28	−2.43	...
L+06	4350	0.50	0.33	0.32	−2.36	...
S+06	4375	0.50	...	...	−2.50	...
W+13	4377	0.77	...	...	−2.34	...
K+18	4351	0.84	...	0.12	−2.49	...
M+19	4670	1.09	0.22	0.15	−2.20	...
J+20	4543	1.02	...	...	−2.09	...
			...			

Table 3.9: Non-exhaustive compilation of literature measurements of stellar parameters for stars in our sample. All chemical abundances have been scaled to the Asplund et al. (2009) Solar abundance scale for consistency in comparison. Reference abbreviations are as follows: S+97 = Sneden, Kraft, et al. (1997), S+00b = Sneden, Pilachowski, et al. (2000), L+06 = Letarte et al. (2006), O+06 = Otsuki et al. (2006), S+06 = Sobeck, Ivans, et al. (2006), R+09 = Roederer, Kratz, et al. (2009), C+09a = Carretta, Bragaglia, Gratton, Lucatello, et al. (2009), C+09b = Carretta, Bragaglia, Gratton, and Lucatello (2009), S+11 = Sobeck, Kraft, et al. (2011), W+13 = Worley et al. (2013), K+18 = Kirby, Xie, et al. (2018), M+19 = Masseron et al. (2019), J+20 = Jönsson et al. (2020). (This table is available in its entirety in machine-readable form online.)

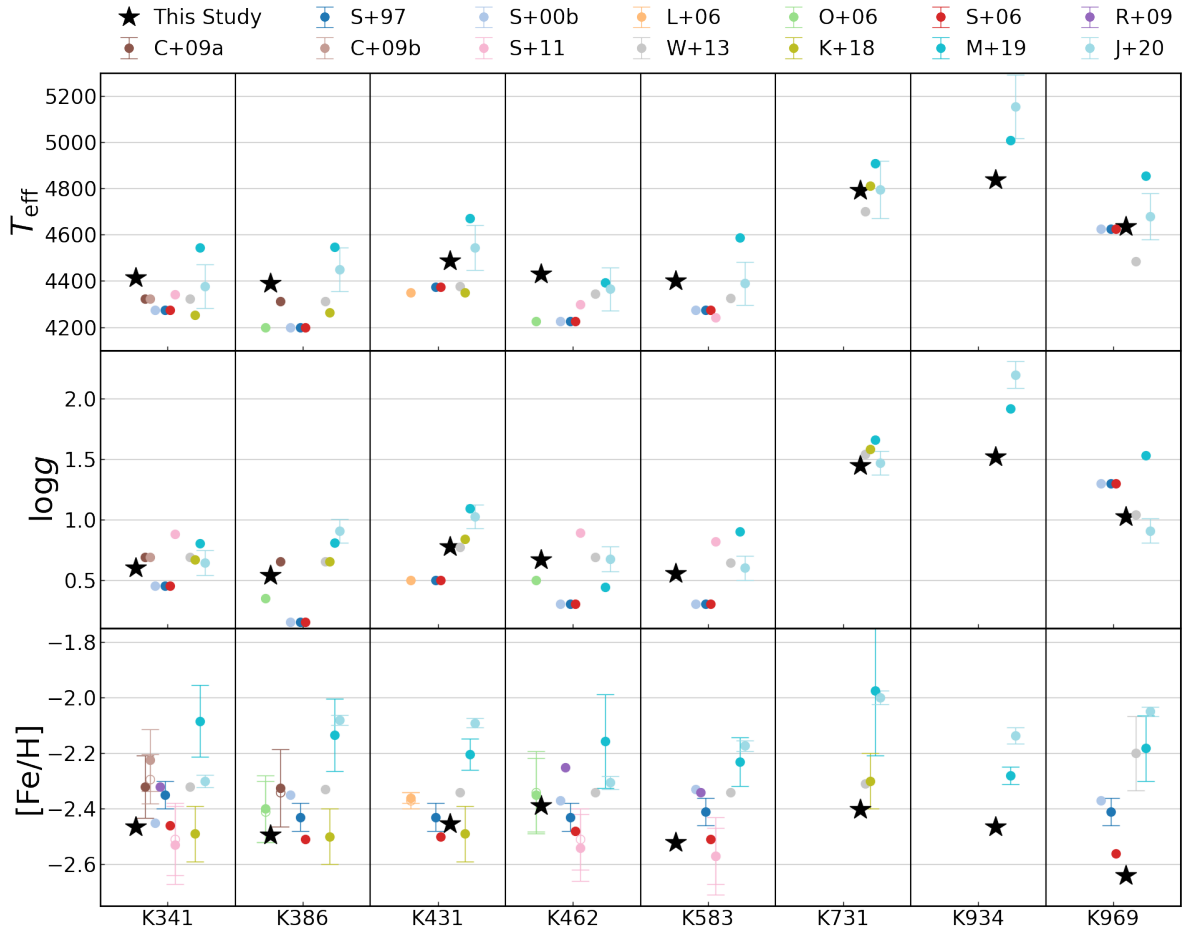


Figure 3.25: The $T_{\rm eff}$ (top), $\log g$ (middle), and [Fe/H] (bottom) measured for each of the 8 stars in our sample using the full-spectrum fitting techniques presented in this paper (black stars). For comparison, we also plot the values for $T_{\rm eff}$, $\log g$, and [Fe/H] reported in a representative sample of literature studies of the same stars (colored circles). In instances where studies report separate values for neutral and ionized atomic species, the ionized value is represented by an open circle. Error bars represent 1σ uncertainties when provided and are too small to be visible for our own measurement uncertainties. Scatter is added in the x-dimension for clarity; points are ordered from left to right in order of increasing mean observed wavelength. A key to the abbreviated references is provided in Table 3.9.

and Jönsson et al. (2020) do not report [Mn/Fe] for any of the other stars in this sample, it is difficult to identify the source of this discrepancy. Given the ubiquity of low [Mn/Fe] abundances in metal-poor stars (see Sobeck, Ivans, et al. 2006), we are inclined to trust our measurement in this instance. We note that large NLTE offsets (∼0.2–0.4 dex) have been

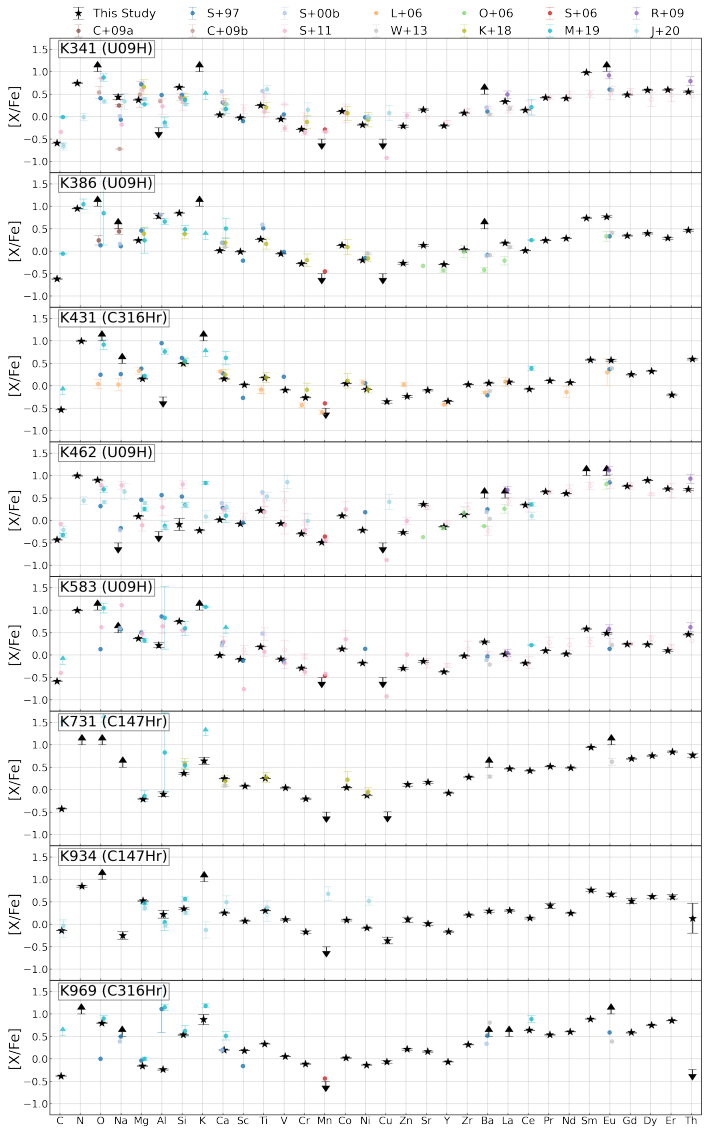


Figure 3.26: The detailed chemical composition of the 8 M15 stars as measured in this study (black stars) and as reported in the literature (colored circles; the same as in Figure 3.25). 95% upper and lower limits are plotted where appropriate (see Section 3.4.3.4) and when reported by the literature. As in Figure 3.25, when separate abundances are provided for neutral and ionized atomic species, the ionized values are represented with open circles. Scatter is added in the x-dimension for clarity; points are ordered from left to right in order of increasing mean observed wavelength.

calculated for [Mn/Fe] in low-metallicity stars (e.g., Bergemann and Gehren 2008; Larsen et al. 2022), which have not been accounted for in either this study or any of the referenced studies.

Only three stars in our sample have literature [Cu/Fe] measurements and all come from one of two studies, Sobeck, Kraft, et al. (2011) and Jönsson et al. (2020). The values reported by these two studies are discrepant by as much as 1 dex, an effect we attribute to the difference in wavelength coverage of the two surveys: Sobeck, Kraft, et al. (2011) used optical Keck/HIRES spectra (the same archival spectra, in fact, as analyzed in this paper), and Jönsson et al. (2020) used NIR APOGEE spectra. Jönsson et al. (2020) also urges caution in adopting the [Cu/Fe] measurements for metal-poor stars with [Fe/H] < -1 as they find systematically higher [Cu/Fe] at lower [Fe/H] in contrast to previous studies and the expectations of the nucleosynthetic origin of Cu (e.g., Sneden, Gratton, et al. 1991; Cayrel et al. 2004; Ishigaki et al. 2013). It is possible, however, that NLTE effects are responsible for the low [Cu/Fe] values found from optical spectroscopy (e.g., Roederer and Barklem 2018). For these three stars, we recover upper limits on [Cu/Fe] that are consistent with the lower abundances of Sobeck, Kraft, et al. (2011). This is in line with expectations given the same underlying observations and LTE assumptions.

We routinely recover [Zn/Fe] to be $\sim$0.25 dex smaller than reported in the literature by Letarte et al. (2006) and Sobeck, Kraft, et al. (2011). We believe this offset is driven by the Zn I line at λ4681.4, which is blended with a poorly modelled Fe I line at λ4681.6 and excluded from the analysis of Letarte et al. (2006) and Sobeck, Kraft, et al. (2011). The other two Zn lines at $\lambda\lambda$4723.5, 4811.9 are slightly underestimated in our fits, consistent with the $\sim$0.25 dex lower measurement of [Zn/Fe].

α Elements In general, we find good agreement with the literature for α elements Mg, Ca, and Ti, though [Ca/Fe] measurements are on the low end of reported values for a few stars. This is most likely a result of differences in the handling of NLTE effects.

We find good agreement with the literature for Si for the C147Hr and C316Hr programs, but the blue-only U09H program observations yield somewhat more mixed agreement, in some cases off by 0.25–0.5 dex. This can be traced to the role of Si as an electron donor in stellar atmospheres. When the full optical spectrum is available, Si is primarily constrained through its isolated absorption lines near λ5700 and λ7400. However, when only the blue-optical spectrum is available, Si is primarily constrained through its indirect influence on other absorption features through changes to the atmospheric structure. While indirect measurement of elements is feasible (e.g., Ting, Conroy, Rix, and Asplund 2018), it relies heavily on the accuracy of the stellar atmospheric models. We know the 1D-LTE models employed in this work to be imperfect, thus explaining the inconsistencies with the literature seen for the blue-only [Si/Fe] abundances.

C, N, O There exists large (0.5–1.0 dex) scatter in the literature values measured for C, N, and O abundances, owing to the complicated nature of their molecular features and the

varying wavelength coverage and methods of these studies (see for example the analysis of C abundances across surveys by Arentsen et al. 2022). We find good agreement with the literature for [C/Fe] with the exception of few measurements from Masseron et al. (2019), which are substantially higher than our values.

Only three of our stars have previously measured [N/Fe]. For K386, our measurement agrees with the measurement of Masseron et al. (2019), but for K341 and K462, we measure [N/Fe] to be >0.5 dex larger than reported by Jönsson et al. (2020).

For nearly all stars in our sample, literature values span a large range from [O/Fe] $\sim$ 0.0–1.0. Our measurements, either lower limits of [O/Fe] $\gtrsim$ 1.0 or in the range of [O/Fe] $\sim$ 0.75–1.0, are most consistent with the high end of the reported literature values.

Light-Odd Elements The light-odd elements Na, Al, and K, similar to C, N, and O, exhibit large (0.5–1.0 dex) scatter in the reported literature values, which is due to the combination of the limited absorption features available for these elements as well as their sensitivity to NLTE effects (e.g., Asplund 2005; Asplund et al. 2009). We recover [Na/Fe] values that either fall among the literature values or lie slightly above the literature measurements except for stars K462, for which we measure [Na/Fe] $\sim$ 0.25 lower than previously reported. For roughly half of the stars in our sample we measure [Al/Fe] in agreement with literature values, while we recover substantially lower [Al/Fe] for the others. In general, measurements of [K/Fe] (or lower limits) are in coarse agreement with measurements (or lower limits) from Masseron et al. (2019). As with Na, the exceptions to this is K462, for which we recover much lower [K/Fe].

Literature measurements of [Sc/Fe] come from Sneden, Kraft, et al. (1997) and Sobeck, Kraft, et al. (2011), which largely analyzed the same archival spectra as in this paper. Our measurements of [Sc/Fe] are in good agreement with those from Sobeck, Kraft, et al. (2011) except for the sole Sc I measurement in K583, which is itself highly discrepant from the Sc II measurement from the same study. Agreement is good with Sneden, Kraft, et al. (1997) for the 4 stars observed as part of the U09H program, while for the remaining stars our values are higher by 0.25–0.5 dex. Not coincidentally, the 4 stars for which agreement is best are the 4 stars for which the same spectra are analyzed in both studies.

Neutron-Capture Elements We find good agreement for neutron-capture elements Y, Zr, La, Pr, Nd, Gd, Dy, Er, and Th. We recover values for [Sr/Fe] that are in good agreement with the values reported by Sobeck, Kraft, et al. (2011), but are $\sim$0.5 dex higher than reported by Otsuki et al. (2006). This discrepancy was previously identified by Sobeck, Kraft, et al. (2011) and attributed to uncertainties in measuring abundances from the Sr II resonance lines.

Similarly, we find good agreement in [Ce/Fe] with Sobeck, Kraft, et al. (2011), but mixed agreement with Masseron et al. (2019). For stars K341 and K462, we recover [Ce/Fe] values that match Masseron et al. (2019) quite well, but for stars K386, K431, K583, and K969, we recover [Ce/Fe] values that are $\sim$0.25–0.50 dex smaller. This is of order the systematic

error that Masseron et al. (2019) reports between [Ce/Fe] values measured using differently derived atmospheric parameters.

Our recovered abundances for Ba, Sm, and Eu are consistently $\sim$0.25–0.50 dex larger than the values reported in the literature. We link these offsets to a combination of factors, including line saturation, NLTE effects, and hyperfine splitting that result in inaccurately modeled line profile shapes (see Roederer, Lawler, et al. 2008; Eitner et al. 2019).

Chapter 4

Strong Outflows and Inefficient Star Formation in the Reionization-era Ultra-faint Dwarf Galaxy Eridanus II

In Chapters 2 and 3, I focused on pushing extragalactic stellar archaeology farther and fainter through the measurement of stellar chemical abundances with low-resolution spectroscopy. The goal of measuring stellar chemistry in ever smaller and more distant galaxies is, of course, to investigate the galactic and stellar processes that shaped their evolution. In this final chapter, I present an analysis of the ultra-faint dwarf galaxy, Eridanus II, and what can be learned about its early evolution when only sparse stellar metallicity measurements are available.

This chapter has been adapted from the manuscript Sandford, Weinberg, et al. (2022), which has been submitted to the Astrophysical Journal and is currently undergoing peer review. It was coauthored by the following individuals, and is included in this thesis with their express permission: David Weinberg, Daniel Weisz, and Sal Fu.

4.1 Abstract

We present novel constraints on the underlying galaxy formation physics (e.g., mass loading factor, star formation history, metal retention) at $z \gtrsim 7$ for the low-mass ($M_* \sim 10^5\,\mathrm{M}_\odot$) Local Group ultra-faint dwarf galaxy (UFD) Eridanus II (Eri II). Using a hierarchical Bayesian framework, we apply a one-zone chemical evolution model to the Eri II CaHK-based photometric metallicity distribution function (MDF; [Fe/H]) and find that the evolution of Eri II is well-characterized by a short, exponentially declining star-formation history ($\tau_{\mathrm{SFH}} = 0.39\pm^{0.18}_{0.13}\,\mathrm{Gyr}$), a low star-formation efficiency ($\tau_{\mathrm{SFE}} = 27.56\pm^{25.14}_{12.92}\,\mathrm{Gyr}$), and a large mass-loading factor ($\eta = 194.53\pm^{33.37}_{42.67}$). Our results are consistent with Eri II forming the

majority of its stars before the end of reionization. The large mass-loading factor implies strong outflows in the early history of Eri II and is in good agreement with theoretical predictions for the mass-scaling of galactic winds. It also results in the ejection of >90% of the metals produced in Eri II. We make predictions for the distribution of [Mg/Fe]-[Fe/H] in Eri II as well as the prevalence of ultra metal-poor stars, both of which can be tested by future chemical abundance measurements. Spectroscopic follow-up of the highest metallicity stars in Eri II ([Fe/H] > -2) will greatly improve model constraints. Our new framework can readily be applied to all UFDs throughout the Local Group, providing new insights into the underlying physics governing the evolution of the faintest galaxies in the reionization era.

4.2 Introduction

At the faintest end of the galaxy luminosity function, ultra-faint dwarf galaxies (UFDs) are some of the oldest ($\gtrsim$13 Gyr), lowest mass ($M_* \lesssim 10^6$), most metal-poor ([Fe/H] $\lesssim -2.0$), and dark matter-dominated ($M/L \gtrsim 100$) systems in the Universe (e.g., Simon 2019, and references therein). As such, Local Group (LG) UFDs and their stellar populations provide a powerful lens through which to study a wide range of astrophysics from the nature of dark matter to star formation, stellar evolution, and chemical enrichment in the early Universe before the epoch of re-ionization.

Eridanus II (Eri II; $M_V = -7.1$), initially discovered in the Dark Energy Survey by Bechtol et al. (2015) and Koposov et al. (2015), is an ideal UFD to study low-mass galaxy evolution at early times. Its dynamical mass ($M_{1/2} = 1.2\pm^{0.4}_{0.3}\times10^7\,\mathrm{M}_\odot$) and stellar metallicity distribution function (MDF; $\langle$[Fe/H]$\rangle = -2.38\pm0.13$ and $\sigma_{\rm [Fe/H]} = 0.47\pm^{0.12}_{0.09}$) measured from calcium triplet (CaT) observations of the brightest red giant branch (RGB) stars in Eri II confirm its status as a metal-poor dark matter-dominated dwarf galaxy (Li et al. 2017). Later spectroscopic and variable star studies provided independent confirmation of the metal-poor dark matter-dominated nature of Eri II (Zoutendijk, Brinchmann, Boogaard, et al. 2020; Martínez-Vázquez et al. 2021; Zoutendijk, Brinchmann, Bouché, et al. 2021). Meanwhile, its star formation history (SFH) measured from deep broadband imaging is consistent with Eri II forming nearly all of its stellar mass ($\sim 2\times10^5\,\mathrm{M}_\odot$) in a short ($<500$ Myr) burst over 13 Gyr ago, making it a true relic of the pre-reionization era (Gallart et al. 2021; Simon, Brown, et al. 2021). Further, its current distance at $\sim$ 350 kpc (Crnojević et al. 2016; Li et al. 2017; Martínez-Vázquez et al. 2021; Simon, Brown, et al. 2021) and its orbit inferred from Gaia eDR3 proper motions place Eri II at first infall into the Milky Way (MW), indicating that it likely evolved in isolation and thus removing the need to account for ram pressure stripping or tidal interactions during its evolution (Battaglia, Taibi, et al. 2022; Fu, Weisz, Starkenburg, Martin, Ji, et al. 2022).

Recently, Fu, Weisz, Starkenburg, Martin, Ji, et al. (2022, hereafter F22) presented newly measured [Fe/H] abundances for 60 Eri II RGB stars from deep narrowband photometry of the calcium H&K doublet (CaHK) acquired with the *Hubble Space Telescope* (*HST*). These observations roughly quadrupled the number of Eri II stars with known metallicities,

substantially improving the sampling of the MDF measured from the CaT observations of Li et al. (2017). F22 found the MDF of Eri II to be characterized by a mean metallicity of $\langle[\mathrm{Fe/H}]\rangle = -2.50 \pm 0.07$ with a dispersion of $\sigma_{[\mathrm{Fe/H}]} = 0.42 \pm 0.06$. While F22 fit simple "closed box" and "leaky box" chemical evolution models to the Eri II MDF, constraints on the physical processes (e.g., star formation and outflows) governing its chemical evolution have yet to be attempted.

Here we use the analytic one-zone galactic chemical evolution models first presented in Weinberg et al. (2017, hereafter WAF17) to fit the MDF of Eri II in a hierarchical Bayesian framework that can be applied uniformly to the MDFs of all observed UFDs, present and future. The key assumption of these models is that the star-forming gas reservoir of Eri II is efficiently mixed and can therefore be approximated as chemically homogeneous at any given time. The models enable us to infer key galactic evolution parameters for Eri II, including the star formation efficiency (SFE), star formation history (SFH) timescale, and the mass loading factor for Eri II and place them in context of past observational and theoretical low-mass galaxy studies.

The structure of this paper is as follows. In Section 4.3, we summarize the data included in our analysis. In Section 4.4, we describe our chemical evolution model and fitting techniques. We present and discuss our results in Sections 4.5 and 4.6 respectively before concluding in Section 4.7. Throughout this work when converting between redshift and age, we assume the flat ΛCDM cosmology of Aghanim et al. (2020).

4.3 Data

In this study, we use the iron abundances, [Fe/H], of 60 stars in the UFD galaxy Eri II measured by F22 from *HST* CaHK narrowband photometry (WFC3/UVIS F395) in conjunction with archival broadband photometry (ACS/WFC 475W and F814W). This sample contains only RGB stars with F475 $\lesssim$ 24 in the inner 260 pc region of Eri II and excludes all stars within 2 half-light radii, r_h, of the singular star cluster in Eri II.

F22 fit the CaHK color index[1] of each star using 13 Gyr old mono-metallic α-enhanced MIST isochrones (Choi et al. 2016; Dotter 2016) to infer [Fe/H] for each star in their sample. Employing Bayesian techniques enabled them to recover the posterior distribution of [Fe/H] for each star, assuming a flat prior. Many stars in their sample exhibit non-Gaussian uncertainties in [Fe/H] with long tails towards low metallicity, which occur as a result of less distinguishable CaHK absorption features in metal-poor stars. A few stars have [Fe/H] posteriors that truncate at [Fe/H] = −4.0 due to the limited extent of the MIST model grid. To capture this non-Gaussianity in our analysis, we approximate the sampled posterior distribution of each star from the MCMC chains of F22 using bounded Gaussian kernel density estimation (KDE).

Figure 4.1 shows the MDF of the 60 Eri II RGB stars in our sample. In the top panel, we plot the MDF as a histogram using the posterior median [Fe/H] values for each star reported

[1]Defined as F395N − F475W − 1.5(F475W − F814W)

by F22. A bin width of 0.35 dex is chosen to match the median measurement uncertainty. In the bottom panel, we display the approximated CaHK [Fe/H] posterior distribution for each star. These will later be used as priors in our analysis (see Section 4.4.2).

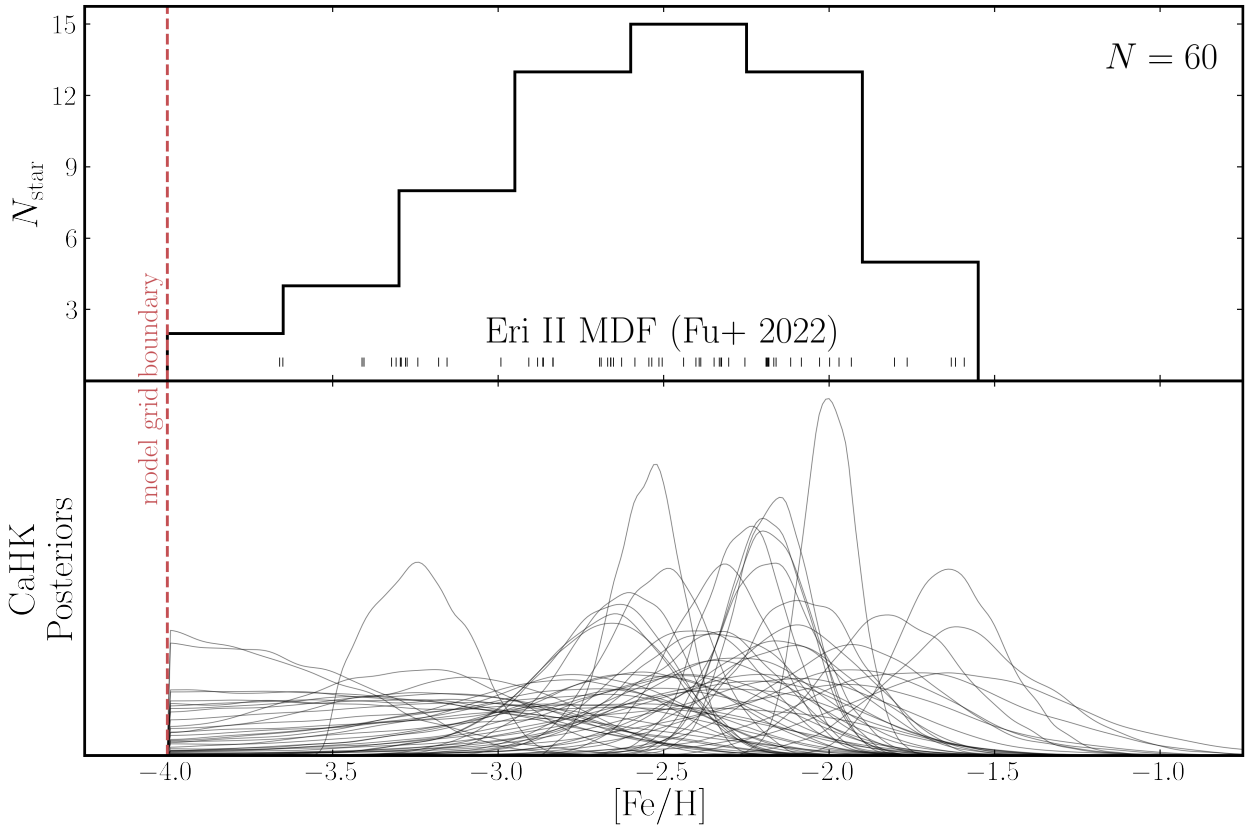


Figure 4.1: Top: Distribution of narrowband CaHK metallicity measurements for 60 RGB stars in the Eri II UFD reported by F22. The median measurement uncertainty (0.35 dex) is reflected in the choice of bin size. Each individual measurement is represented by a tick in the rug plot. Bottom: Posterior [Fe/H] distributions for each star approximated by applying bounded Gaussian kernel density estimation to the MCMC samples of F22. The posteriors of several stars exhibit long tails towards low metallicity and/or truncation at the limit of the MIST model grid ([Fe/H] = −4.0).

4.4 Methods

4.4.1 Chemical Evolution Model

We adopt the galactic chemical evolution framework presented and discussed extensively in WAF17. In brief, this analytic model tracks the time evolution of abundances in a fully mixed (one-zone) system experiencing gas accretion, star formation, supernova enrichment, and outflows. Relative to previous analytic models, the key innovation of the WAF17 model is its ability to separately track both rapid enrichment from core collapse supernovae (CC SNe) and delayed enrichment from Type Ia supernovae (SNe Ia). In the limit of prompt enrichment and no gas accretion, the model approaches "closed box" (no outflow) or "leaky box" scenarios, but the behavior in this limit is quite different from that of models with ongoing accretion. A complete description of the model and its input parameters can be found in WAF17 (see especially their Table 1). We summarize parameter choices for our fiducial Eri II model below and in Table 4.1. We also consider several alternative models with variations on the fiducial choices, which we describe in Section 4.4.3. A discussion of key model assumptions and their potential impact on the interpretation of the chemical evolution of Eri II is presented in Section 4.6.3.

4.4.1.1 Star Formation

Motivated by the star formation history (SFH) measured for Eri II by Simon, Brown, et al. (2021) and Gallart et al. (2021), we adopt a truncated exponentially declining star formation rate

$$\dot{M}_* \propto \begin{cases} \exp\left(-t/\tau_{\rm SFH}\right), & \text{if } t \leq t_{\rm trunc} \\ 0, & \text{if } t > t_{\rm trunc} \end{cases} \tag{4.1}$$

where $\tau_{\rm SFH}$ is the SFH timescale and $t_{\rm trunc}$ is the time at which all star formation ceases. The sharp truncation of the SFH is adopted to simulate the abrupt quenching of low-mass galaxies (e.g., from ram pressure stripping or reionization). We leave both $\tau_{\rm SFH}$ and $t_{\rm trunc}$ as free parameters.

The conversion of gas into stars is governed by a linear star formation law characterized by the star formation efficiency (SFE) timescale (or inverse SFE) according to

$$\tau_{\rm SFE} \equiv {\rm SFE}^{-1} \equiv M_g/\dot{M}_*, \tag{4.2}$$

where M_g and $\dot{M}_*$ are the gas mass and star formation rate (SFR) respectively. We leave $\tau_{\rm SFE}$ as a free parameter. (In WAF17, the SFE timescale is denoted τ_*.)

4.4.1.2 Gas Flows

The mass recycling fraction, r, sets the fraction of mass formed into stars that is immediately returned to the ISM without further chemical enrichment by CCSNe and asymptotic giant branch (AGB) stars. Because this recycling is not a source of new metals, its main effect

is to slow the rate at which metals in the ISM are depleted by star formation. We adopt a recycling fraction $r = 0.37$, which is appropriate for a Kroupa (2001) IMF after 1 Gyr. As shown by WAF17, treating this recycling as instantaneous is an accurate approximation, because much of the recycled material originates from stars with short lifetimes (see their Fig. 7). Moreover, the effect of this approximation is small when the metallicity is low or when galactic winds are important as is the case for Eri II.

Gas ejected from the ISM by stellar feedback (i.e., CCSNe and AGB winds) scales linearly with the SFR according to

$$\eta = \dot{M}_{\rm outflow}/\dot{M}_*, \tag{4.3}$$

where η is the mass-loading factor. We leave η as a free parameter.

Gas inflow is specified implicitly in the model through the provided SFH, SFE, mass recycling fraction, and mass-loading factor such that the depletion of gas by star formation and outflows is sufficiently balanced to maintain the SFR given in Equation 4.1. WAF17 demonstrates that the gas infall rate can be obtained analytically in terms of other model parameters as

$$\dot{M}_{\rm inf} = (1 + \eta - r)\dot{M}_* + \tau_{\rm SFE}\ddot{M}_* \tag{4.4}$$

(see their Equation 9). For our exponential SFH, $\ddot{M}_* = -\dot{M}_*/\tau_{\rm SFH}$. We assume accreted gas is pristine and free of previous enrichment.

4.4.1.3 Chemical Enrichment

Enrichment from CC SNe is assumed to occur instantaneously following star formation. Enrichment from SNe Ia, on the other hand, is assumed to follow a delay time distribution (DTD) that accurately approximates the $t^{-1.1}$ power-law found empirically by Maoz, Mannucci, et al. (2012)[2]. We adopt a minimum time delay, t_D, of 0.05 Gyr corresponding to the lifetime of the most massive white dwarf progenitors.

The WAF17 model parameterizes chemical enrichment using dimensionless IMF-weighted yield parameters, which are presumed to be independent of metallicity. These yield parameters represent the mass of elements produced per unit mass of star formation. We adopt lower yield values than WAF17, motivated by the recent study of Rodríguez et al. (2022), who infer a population-averaged mean Fe yield of $0.058\,{\rm M}_\odot$ per CC SN. A Kroupa (2001) IMF predicts approximately one $M > 8\,{\rm M}_\odot$ star per $100\,{\rm M}_\odot$ of star formation, so this estimate suggests a dimensionless CC SN Fe yield $y^{\rm cc}_{\rm Fe} \approx 6 \times 10^{-4}$, which we adopt for our models. Although our data do not include Mg abundances, we present predictions of [Mg/Fe] vs. [Fe/H] that could be tested with future data. We choose $y^{\rm cc}_{\rm Mg} = 0.001$, which puts the low-metallicity α "plateau" at [Mg/Fe] ≈ 0.5, roughly consistent with measurements in the MW disk from the H3 Survey (Conroy, Weinberg, et al. 2022). We assume that Mg has no SN Ia contribution, i.e., $y^{\rm Ia}_{\rm Mg} = 0$. Finally, we choose $y^{\rm Ia}_{\rm Fe} = 0.0012$ so that models evolved with "Milky Way disk" parameters similar to WAF17 reach [Mg/Fe] ≈ 0 at late times. For an Fe yield of $0.7\,{\rm M}_\odot$ per

[2]As discussed in WAF17, we approximate the power-law distribution using a sum of two exponentials to allow for an analytic solution.

SN Ia, this $y_{\rm Fe}^{\rm Ia}$ corresponds to 1.7×10^{-3} SNe Ia per $M_\odot$ of star formation, approximately consistent with the rate found by Maoz and Graur (2017). Moderate changes to the yields would change our best fit parameter values, especially for η, but they would not change our qualitative conclusions.

The products of CC SNe and SNe Ia that are deposited into the ISM are assumed to mix completely and instantaneously such that they are available for star formation immediately. This simplification, known as the instantaneous mixing approximation, has been shown to be a reasonable assumptions for CC SNe and SNe Ia products in low-mass, ancient galaxies like Eri II (e.g., Escala, Wetzel, et al. 2018).

The WAF17 model assumes that outflows are comprised of gas at the ISM metallicity, so that the associated metal loss rate is $\eta \dot{M}_* Z_{\rm ISM}$. We also consider an alternative formulation in which a fraction of supernova-produced metals are directly ejected from the galaxy and only a fraction $f_{\rm ret}$ are retained within the star-forming ISM. In this case, all yields are multiplied by the factor $f_{\rm ret}$, which we assume to be the same for CC SNe and SNe Ia because without [α/Fe] measurements we have little leverage to separate the two retention factors. The outflows described by η are still assumed to be at the ISM metallicity, but the total metal loss rate is larger because of the direct ejection, which implicitly occurs at a rate $y(1 - f_{\rm ret})\dot{M}_*$ for each channel. In our Fiducial model, we fix $f_{\rm ret} = 1$, reproducing the scenario in which all supernova-produced metals are deposited initially into the star-forming ISM.

4.4.1.4 Initial and Final Conditions

Initial conditions of the model are largely set by the aforementioned model parameters. An exponential SFH as assumed in our Fiducial model requires that Eri II begin with a non-zero gas mass at $t = 0\,$Gyr such that

$$M_g(t = 0) = \tau_{\rm SFE} \dot{M}_*(t = 0). \tag{4.5}$$

This initial gas mass is assumed to be primordial in composition (e.g., $Z = 0$). The stellar mass of Eri II at $t = 0\,$Gyr is assumed to be zero.

The evolution of the model effectively ends when the SFR is abruptly truncated at $t = t_{\rm trunc}$. Within the framework of this model, such a truncation could be achieved by removing all gas from the ISM and shutting off gas accretion (e.g., setting $M_g = \dot{M}_{\rm inf} = 0$), by heating gas in the ISM such that it cannot form stars (e.g., setting $\tau_{\rm SFE} = \infty$), or some combination of these effects. Both ram pressure stripping and reionization provide plausible physical explanations for the truncation of star formation, though the latter seems more likely given the relative isolation of Eri II. We do not attempt to include a more detailed prescription for star formation truncation as our dataset is of insufficient size and quality to yield meaningful insight.

Finally, all model parameters listed in Table 4.1 are assumed to be constant as a function of time, though we do not expect this to be strictly true in reality. For example, SN yields might vary with stellar metallicity, and the mass-loading factor could decrease over cosmic

time as the mass of the dark matter halo grows. We leave more detailed analysis using time- and metallicity-dependent parameters for future study, noting that the parameters used in this work can be thought of as time-averaged quantities characteristic of the evolution of Eri II.

4.4.1.5 Constructing the MDF

The number of stars born as a function of metallicity predicted by the model can be defined using the chain rule in terms of the SFR and the rate of change in [Fe/H] with time:

$$\frac{dN}{d[\mathrm{Fe/H}]} = \frac{dN/dt}{d[\mathrm{Fe/H}]/dt} \propto \frac{\dot{M}_*}{d[\mathrm{Fe/H}]/dt}. \tag{4.6}$$

We caution that Equation 4.6 only holds for a monotonically increasing [Fe/H], which is universally true for the WAF17 model. Time steps of $dt = 10^{-5}$ Gyr and metallicity sampling of $d[\mathrm{Fe/H}] = 0.01$ dex sufficiently minimize numerical artifacts. To convert between mass fractions of Fe and Mg predicted by the model and solar-scaled abundances [Fe/H] and [Mg/H], we adopt the photospheric abundance scale from Asplund et al. (2009), corresponding to solar mass fractions of 0.0013 and 0.0056 respectively. This choice of solar abundance scale purposefully matches the solar abundance scale used by the MIST isochrones that underpin the F22 CaHK measurements.

Param.	Description	Value/Priors	Units	References
	Fixed Parameters			
Δt	Time step	10^{-5}	Gyr	...
$Z_{\mathrm{Fe},\odot}$	Solar iron abundance by mass	0.0013	...	[1]
$Z_{\mathrm{Mg},\odot}$	Solar magnesium abundance by mass	0.0007	...	[1]
$y^{\mathrm{cc}}_{\mathrm{Mg}}$	IMF-integrated CCSN magnesium yield	0.0026	...	[2]
$y^{\mathrm{cc}}_{\mathrm{Fe}}$	IMF-integrated CCSN iron yield	0.0012	...	[2]
$y^{\mathrm{Ia}}_{\mathrm{Mg}}$	IMF-integrated SN Ia magnesium yield	0.0	...	[2]
$y^{\mathrm{Ia}}_{\mathrm{Fe}}$	IMF-integrated SN Ia iron yield	0.003	...	[2]
r	Mass recycling fraction	0.4	...	[3]
α_{Ia}	Slope of SN Ia power-law delay time distribution	-1.1	...	[4]
t_D	Minimum delay time for SNe Ia	0.05	Gyr	[3]
f_{ret}	Fraction of newly produced metals retained by the ISM	1.0	⋯	⋯
	Free Parameters			
τ_{SFH}	star formation history timescale, $\dot{M}_* \propto e^{-t/\tau_{\mathrm{SFH}}}$	$\mathcal{TN}(0.7, 0.3, 0.0, \infty)$	Gyr	[5]
τ_{SFE}	$= M_g/\dot{M}_*$, star formation efficiency timescale	$\mathcal{U}(0, 10^4)$	Gyr	⋯
t_{trunc}	Time of SFH Truncation	$\mathcal{TN}(1.0, 0.5, 0.0, \infty)$	Gyr	[5]
η	$= \dot{M}_{\mathrm{outflow}}/\dot{M}_*$, mass-loading factor	$\mathcal{U}(0, 10^3)$	...	..

Table 4.1: Fiducial model parameters adopted in this work. Priors for the free parameters are introduced in Section 4.4.2. The implementation of these parameters is described in detail in WAF17. The SN Ia DTD is a sum of two exponentials that accurately approximates a $t^{-1.1}$ power-law. **References.** [1] Asplund et al. (2009), [2] Conroy, Weinberg, et al. (2022), [3] Weinberg et al. (2017), [4] Maoz, Mannucci, et al. (2012), [5] Gallart et al. (2021)

4.4.2 Likelihood and Priors

We employ Bayesian hierarchical modelling to fit our chemical evolution model to the MDF of Eri II. As in previous analyses of dwarf galaxy MDFs (e.g., Kirby, Lanfranchi, et al. 2011), we normalize the metallicity distribution, $dN/d[\mathrm{Fe/H}]$, predicted by our chemical evolution model (Equation 4.6) such that $\int_{-\infty}^{\infty} dN/d[\mathrm{Fe/H}]\ d[\mathrm{Fe/H}] = 1$ and adopt it as a probability distribution function (PDF) for the observed stellar abundances.

To account for the lower limit on observed [Fe/H] imposed by the MIST isochrone grid, we truncate this PDF below $[\mathrm{Fe/H}] < -4.0$ and redistribute the truncated mass at the boundary following a half-normal distribution with width $\sigma = 0.35$ dex in accordance with the median measurement uncertainty from F22. We present an example of a (non-)truncated PDF predicted by the model in Figure 4.2.

Unlike in previous studies, we do not directly incorporate the observed [Fe/H] abundances into our likelihood function. Instead, we adopt the posterior distributions from F22 (described in Section 4.3) as priors on the "latent" [Fe/H] of each star, which we denote with a prime:

$$\mathrm{P}_{\mathrm{prior}}([\mathrm{Fe/H}]_i') = \mathrm{P}_{\mathrm{Fu22}}([\mathrm{Fe/H}] = [\mathrm{Fe/H}]_i' | \mathrm{CaHK}_i) \tag{4.7}$$

These latent abundances, $[\mathrm{Fe/H}]_i'$, are fit simultaneously along with the free model parameters (τ_{SFE}, τ_{SFH}, t_{trunc}, η, and where relevant f_{ret}). The total log-likelihood is then

$$\ln \mathcal{L} = \sum_{i=1}^{N_*} \ln \left. \frac{d\mathrm{N}}{d[\mathrm{Fe/H}]} \right|_{[\mathrm{Fe/H}]_i'}, \tag{4.8}$$

where the sum is over all N_* observed stars and the probability distribution function, $dN/d[\mathrm{Fe/H}]$, is evaluated at the latent abundance $[\mathrm{Fe/H}]_i'$ of each star. Equation 4.8 ensures that we do not infer an $[\mathrm{Fe/H}]_i'$ for any star beyond the maximum [Fe/H] predicted by the model, while Equation 4.7 penalizes the model for requiring $[\mathrm{Fe/H}]_i'$ values in tension with the CaHK_i measurements.

We adopt a truncated Gaussian prior on the SFH timescale, τ_{SFH}, centered at 0.7 Gyr with width 0.3 Gyr and bounded to be positive definite:

$$\tau_{\mathrm{SFH}} \sim \mathcal{TN}(0.7, 0.3, 0.0, \infty).$$

This choice of prior is informed by Gallart et al. (2021) who derived a SFH from deep *HST*/ACS photometry of Eri II that is peaked at the oldest possible age with half-width at half-maximum (HWHM) of ∼0.5 Gyr corresponding to a ∼0.7 Gyr e-folding timescale for an exponentially declining SFH. While a negatively skewed prior may be appropriate—Gallart et al. (2021) suggest that the true duration of the star formation burst is likely unresolved by their study and could be as short as 100 Myr—we adopt a Gaussian prior for simplicity.

Similarly, we adopt a positive definite truncated Gaussian prior on the SFH truncation time, t_{trunc}, centered at 1 Gyr with width 0.5 Gyr.

$$t_{\mathrm{trunc}} \sim \mathcal{TN}(1.0, 0.5, 0.0, \infty).$$

This too is motivated by the lack of evidence found by Gallart et al. (2021) for star formation in Eri II within the last $\sim 13\,\mathrm{Gyr}$ years.

We utilize broad uniform priors for the remaining model parameters as follows:

$$\log_{10} \tau_{\mathrm{SFE}} \sim \mathcal{U}(0.0, 4.0),$$
$$\eta \sim \mathcal{U}(0.0, 10^3)$$

Note that we have reparameterized to fit for τ_{SFE} in log-space given the large dynamic range we wish to explore.

Together, the sum of the 64 log-priors (4 for the model parameters and 1 for each stellar [Fe/H]′) and the log-likelihood presented in Equation 4.8 yield the log-posterior distribution that we wish to sample.

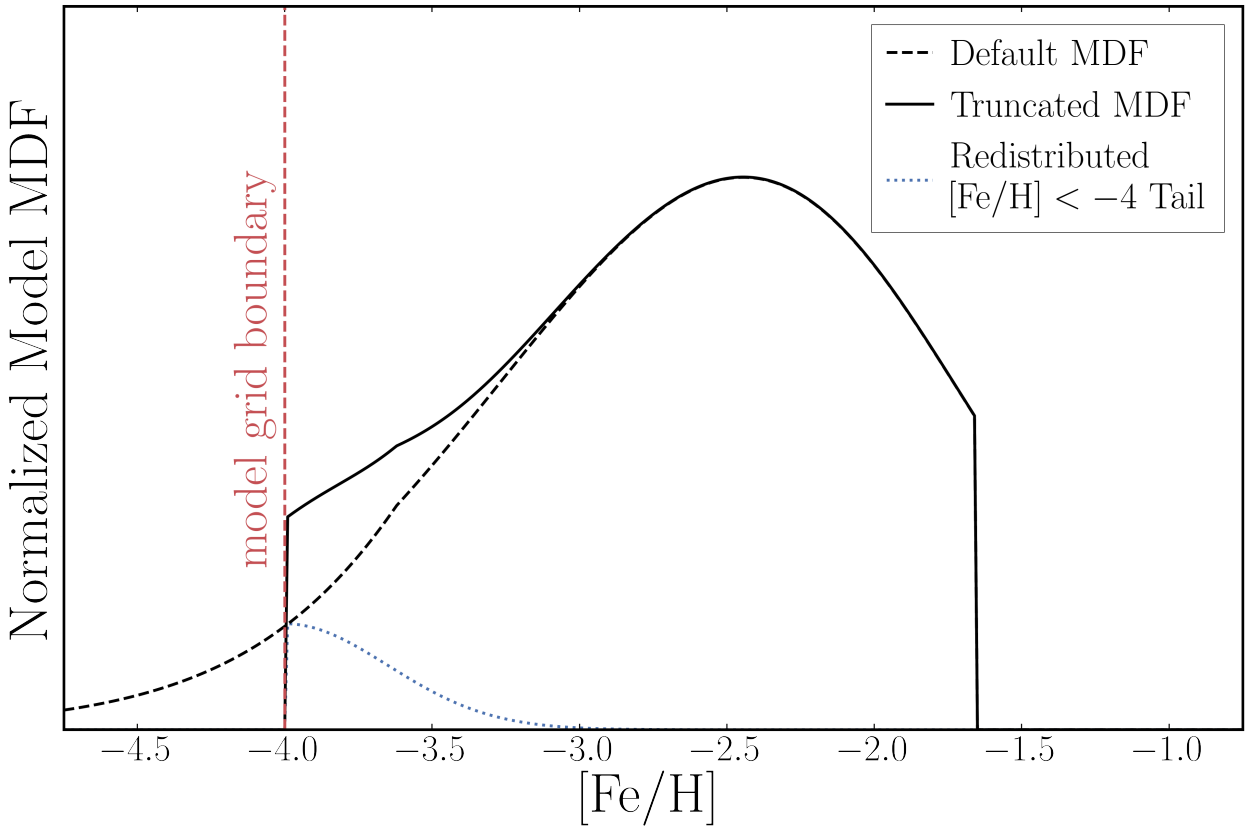


Figure 4.2: Example model MDF (see Section 4.4.1) before and after truncation below [Fe/H] < -4 (dashed and solid black lines respectively). The dotted blue line illustrates the redistribution of the truncated probability following a half-normal distribution with $\sigma = 0.35\,\mathrm{dex}$. This model was generated with the following parameters: $\tau_{\mathrm{SFE}} = 100\,\mathrm{Gyr}$, $\tau_{\mathrm{SFH}} = 0.5\,\mathrm{Gyr}$, $t_{\mathrm{trunc}} = 1.0\,\mathrm{Gyr}$, and $\eta = 50$.

4.4.3 Alternative Models

In addition to the Fiducial model described above, we consider several alternative models to build physical intuition and test specific scenarios. We briefly describe the motivation and adjustments for each below. Any parameters not explicitly referenced are identical to those in the Fiducial model (Table 4.1).

Linear-Exponential SFR Model In the Fiducial model, we assume an exponentially declining SFR, which requires a non-zero gas mass at the onset of star formation. With this model, we test an alternative "linear-exponential" functional form for the SFR functional form, given by

$$\dot{M}_* \propto \begin{cases} t \exp\left(-t/\tau_{\rm SFH}\right), & \text{if } t \leq t_{\rm trunc} \\ 0, & \text{if } t > t_{\rm trunc} \end{cases}. \tag{4.9}$$

In this model, the need for an initial gas reservoir is avoided as the galaxy begins with no star formation at $t = 0\,$Gyr. Rather, the SFR increases rapidly from zero to its peak at $t = \tau_{\rm SFH}$ before declining more gradually. As in the fiducial case, we adopt a prior for $\tau_{\rm SFH}$ motivated by the $\sim$0.5 Gyr SFH HWHM measured by Gallart et al. (2021). For the above linear-exponential SFH, this corresponds to $\tau_{\rm SFH} \sim 0.2\,$Gyr, and so we use the following truncated Gaussian prior:

$$\tau_{\rm SFH} \sim \mathcal{TN}(0.2, 0.1, 0.0, \infty).$$

Constant SFR Model With this model, we test whether a constant SFR could reproduce the Eri II MDF. No $\tau_{\rm SFH}$ is fit for this model as a constant SFR is equivalent to letting $\tau_{\rm SFH} \to \infty$.

Metal-Loading Model In this model, we enable supernova-produced metals to be directly ejected from the galaxy by allowing a non-zero retention factor. Specifically, we adopt a uniform prior on $f_{\rm ret}$:

$$f_{\rm ret} \sim \mathcal{U}(0.0, 1.0).$$

Due to the large degeneracy between $f_{\rm ret}$ and η, we find it desirable to impose a tight prior on η, which we force to be roughly 4 times smaller than preferred in the fiducial case (see Section 4.5):

$$\eta \sim \mathcal{N}(50, 10).$$

While we do achieve converged Monte Carlo chains with $f_{\rm ret}$ and η both free (see Section 4.4.4), the degeneracy between the parameters makes the results hard to interpret.

High SFE Model With this model, we investigate whether the MDF of Eri II can be modelled assuming a short SFE timescale of $\log_{10}\tau_{\rm SFE} \sim 0.4$, which is roughly an order of magnitude smaller than preferred by the Fiducial model (see Section 4.5). Such a high SFE might be expected if a large fraction ($\sim$75%) of the gas in Eri II was in the molecular phase. We force this enhanced SFE by implementing a tight prior on $\log_{10}\tau_{\rm SFE}$ of

$$\log_{10}\tau_{\rm SFE} \sim \mathcal{N}(0.4, 0.1).$$

Longer SN Ia Delay Model In the Fiducial model, we implement a minimum time delay for SN Ia of $t_{\rm D} = 0.05\,$Gyr corresponding to the lifetimes of the most massive SN Ia progenitors. However, previous chemical evolution studies (e.g., Schönrich et al. 2009; Andrews et al. 2017) have adopted a slightly longer time delay of 0.15 Gyr before the first SN Ia. In this model, we set $t_{\rm D} = 0.15\,$Gyr to test the impact of assuming a more delayed onset of SN Ia.

No SN Ia Model To evaluate the importance of SN Ia enrichment on the shape of the Eri II MDF, we consider a scenario in which SN Ia do not contribute at all to the enrichment of the galaxy. In this model, we set $y^{\rm Ia}_{\rm Fe} = 0$ but the same could be accomplished by setting $y^{\rm Ia}_{\rm Fe} = 0$.

Enhanced SN Ia Model In this model we assume the specific SN Ia rate scales with metallicity proportional to $Z^{-0.5}$ as found in the recent analysis of Johnson, Kochanek, et al. (2022). For Eri II, this scaling would imply an enhancement of the SN Ia rates by roughly an order of magnitude, which we implement by simply increasing the fiducial SN Ia Fe yield $y^{\rm Ia}_{\rm Fe}$ by a factor of 10 to $y^{\rm Ia}_{\rm Fe} = 0.012$.

4.4.4 Sampling

To sample our posterior distributions, we employ the *Preconditioned Monte Carlo* (PMC) method for Bayesian inference implemented in the publicly available Python package `pocoMC`[3] (Karamanis, Beutler, et al. 2022; Karamanis, Nabergoj, et al. 2022). PMC uses a combination of a normalizing flow with a sequential Monte Carlo sampling scheme to decorrelate and efficiently sample high-dimensional distributions with non-trivial geometry.

We initialize 5000 walkers from the prior distributions described in Section 4.4.2, imposing an arbitrary log-posterior threshold to ensure walkers are not too distant from the bulk of the posterior mass. We adopt default hyperparameters for `pocoMC`, run until the sampler has converged (i.e., when the "inverse temperature" $\beta = 1$), and then draw an additional 5000 samples for a total of 10,000 samples from the posterior distribution.

[3] `https://github.com/minaskar/pocomc`

4.5 Results

4.5.1 Fiducial Fit to Eri II MDF

We begin by briefly summarizing the recovered posterior distribution for the model parameters $\log_{10}\tau_{\rm SFE}$, $\tau_{\rm SFH}$, $t_{\rm trunc}$, and η, which we display in Figure 4.3. For each parameter, we report the median of each marginalized posterior distribution (blue lines) and the 16th and 84th percentiles (dashed black lines) in Table 4.2; for brevity, we refer to the posterior medians as "best-fit" values hereafter, with the percentile ranges treated as $\pm1\sigma$ uncertainties. These are discussed individually in Section 4.5.1.1. In short, we find that the MDF of Eri II is sufficient to place constraints on $\log_{10}\tau_{\rm SFE}$ ($1.44\pm^{0.28}_{0.27}$), $\tau_{\rm SFH}$ ($0.39\pm^{0.18}_{0.13}$ Gyr), and η ($194.53\pm^{33.37}_{42.67}$) but not $t_{\rm trunc}$ ($1.37\pm^{0.37}_{0.37}$ Gyr), which remains prior-dominated. We explore the aspects of the MDF shape that contribute to these constraints (or lack thereof) in Appendix 4.8.1.

Importantly, we find that our Fiducial model produces realistic predictions for the Eri II MDF, which we illustrate in Figure 4.4 using both continuous (top panel) and binned (middle panel) representations. The blue dashed lines in these panels represent the latent MDFs for the best-fit model parameters, which is the sum of the latent posterior distributions of the individual stars (in the top panel) integrated over the bins (in the middle panel). To visualize the uncertainties on the latent MDF, we make a bootstrap selection from our set of 60 stars (allowing replacement) and draw from the latent [Fe/H] posterior distribution of each, capturing both the uncertainties from finite sample size and the measurement uncertainties for each star. The resulting 95% confidence interval is depicted by the blue shaded region. The best-fit model MDF (thick red line) is in good agreement with the latent MDF, predicting a negatively skewed distribution with a small low-metallicity tail and little to no truncation below the model grid boundary. We additionally perform a posterior predictive check of our model by generating model MDFs for 1000 random draws from the parameter posteriors (thin red lines), which illustrates the range of MDFs consistent with the uncertainties on our best-fit model parameters. We include the observed CaHK [Fe/H] MDF from F22 (solid gray line) for reference, but reiterate that reproducing the latent MDF, not the input CaHK MDF, maximizes the likelihood.

In the bottom panel of Figure 4.4, we present the posterior distribution of the underlying [Fe/H] for each of the 60 stars. Compared to the input CaHK posteriors (Figure 4.1; bottom panel), these updated posteriors exhibit less pronounced low-metallicity tails as well as less frequent and less severe truncation at the model boundary of [Fe/H] $= -4$. The mean metallicity, $\langle$[Fe/H]$\rangle = -2.52\pm^{0.04}_{0.04}$, and metallicity dispersion, $\sigma_{\rm [Fe/H]} = 0.45\pm^{0.04}_{0.04}$, are still in good agreement with the values found by F22.

Model	$\log_{10}\tau_{\rm SFE}$	$\tau_{\rm SFH}$ [Gyr]	$t_{\rm trunc}$ [Gyr]	η	$f_{\rm ret}$	t_D [Gyr]	$y^{\rm Ia}_{\rm Fe}$	BF
Fiducial	$1.44\pm^{0.28}_{0.27}$	$0.39\pm^{0.18}_{0.13}$	$1.37\pm^{0.37}_{0.37}$	$194\pm^{43}_{33}$	1	0.05	0.0012	1.000
Linear-Exponential SFR	$1.41\pm^{0.22}_{0.23}$	$0.21\pm^{0.06}_{0.06}$	$1.26\pm^{0.36}_{0.32}$	$186\pm^{35}_{40}$	1	0.05	0.0012	3.099
Constant SFR	$1.85\pm^{0.18}_{0.23}$	∞	$1.03\pm^{0.37}_{0.34}$	$17\pm^{21}_{12}$	1	0.05	0.0012	0.232
Metal-Loading	$1.00\pm^{0.48}_{0.34}$	$0.47\pm^{0.19}_{0.16}$	$1.26\pm^{0.39}_{0.38}$	$52\pm^{8}_{8}$	$0.32\pm^{0.22}_{0.07}$	0.05	0.0012	0.711
High SFE	$0.47\pm^{0.08}_{0.08}$	$0.08\pm^{0.04}_{0.08}$	$1.06\pm^{0.39}_{0.37}$	$145\pm^{37}_{30}$	1	0.05	0.0012	0.002

Longer SN Ia Delay	$1.44\pm^{0.23}_{0.21}$	$0.49\pm^{0.16}_{0.14}$	$1.33\pm^{0.35}_{0.34}$	$155\pm^{31}_{35}$	1	0.15	0.0012	1.018
No SN Ia	$1.57\pm^{0.23}_{0.18}$	$0.45\pm^{0.21}_{0.18}$	$1.26\pm^{0.36}_{0.33}$	$71\pm^{21}_{39}$	1	0.05	0.0000	2.239
Enhanced SN Ia	$1.44\pm^{0.48}_{0.31}$	$0.48\pm^{0.18}_{0.12}$	$1.06\pm^{0.32}_{0.27}$	$879\pm^{61}_{91}$	1	0.05	0.0120	0.024

Table 4.2: Inferred Eri II parameters. Median values and 1σ uncertainties inferred for the model parameters from the fiducial and alternative model fits. Values without uncertainties were held fixed. The estimated Bayes factor relative to the Fiducial model is presented in the right-most column.

4.5.1.1 Inferred Parameters of Eri II

Star Formation Efficiency We infer the log-SFE timescale of Eri II to be $\log_{10}\tau_{\rm SFE} = 1.44\pm^{0.28}_{0.27}$ ($\tau_{\rm SFE} = 27.56\pm^{25.14}_{12.92}\,{\rm Gyr}$) or in terms of the SFE ($\tau^{-1}_{\rm SFE}$): SFE $= 0.036\pm^{0.032}_{0.017}\,{\rm Gyr}^{-1}$. This timescale is quite large compared to the SFE timescale of molecular gas ($\tau_{\rm SFE} = 2\,{\rm Gyr}$; Leroy et al. 2008) but in line with the current paradigm that low-mass galaxies are the least efficient at converting their gas into stars (e.g., Behroozi et al. 2013, and references therein).

In Figure 4.5, we compare the inferred SFE of Eri II with the best-fit SFE reported by previous chemical evolution studies of 13 LG dwarf galaxies spanning a wide range in stellar masses(Lanfranchi and Matteucci 2004; Lanfranchi, Matteucci, and Cescutti 2006; Lanfranchi and Matteucci 2007, 2010; Vincenzo et al. 2014; Romano et al. 2015; Lacchin et al. 2020; Alexander et al. 2023). Here, we adopt a stellar mass for Eri II of $2\times10^5\,{\rm M}_\odot$ from Gallart et al. (2021). Despite the range in chemical evolution models and assumptions adopted in these studies, a clear trend between a the SFE of a galaxy and its stellar mass is visible. As expected, galaxies more massive than Eri II are found to be more efficient at converting gas to stars (SFE $\sim$ 0.5–1.0 ${\rm Gyr}^{-1}$), while galaxies less massive than Eri II are found to be less efficient (SFE $\sim$ 0.003–0.03 ${\rm Gyr}^{-1}$). Given this apparent relationship, the SFE we infer for Eri II is in good agreement with expectations given its stellar mass.

A low SFE like that found for Eri II may be indicative that the majority of its gas is in the atomic phase. Following the reasoning of Johnson and Weinberg (2020) that $\tau_{\rm SFE} = (2\,{\rm Gyr})(1+M_{\rm HI}/M_{\rm H_2})$, we infer that the molecular gas fraction in Eri II was only $7.26\pm^{6.41}_{3.46}\%$. This, of course, is only a rough approximation given the assumptions made in our model. It is possible that the molecular gas fraction of Eri II changed over its star-forming lifetime, resulting in a time-varying SFE. WAF17 find that smooth evolution of $\tau_{\rm SFE}$ has little impact on chemical evolution tracks if the SFH remains fixed, but we have not yet investigated the low-metallicity regime relevant here. While numerical solutions with time-dependent SFE are feasible, it is not clear what behavior would be appropriate to assume for Eri II, and so we stick with the simplest assumption of constant SFE.

Star Formation History We recover the SFH timescale of Eri II to be $\tau_{\rm SFH} = 0.39\pm^{0.18}_{0.13}\,{\rm Gyr}$. This is slightly shorter than the SFH reported by Gallart et al. (2021), which informed our choice of prior ($\tau_{\rm SFH,prior} = 0.7\pm0.3\,{\rm Gyr}$). The SFH timescale we infer corresponds to a star formation HWHM of $270\pm^{130}_{90}\,{\rm Myr}$—about half the duration found by

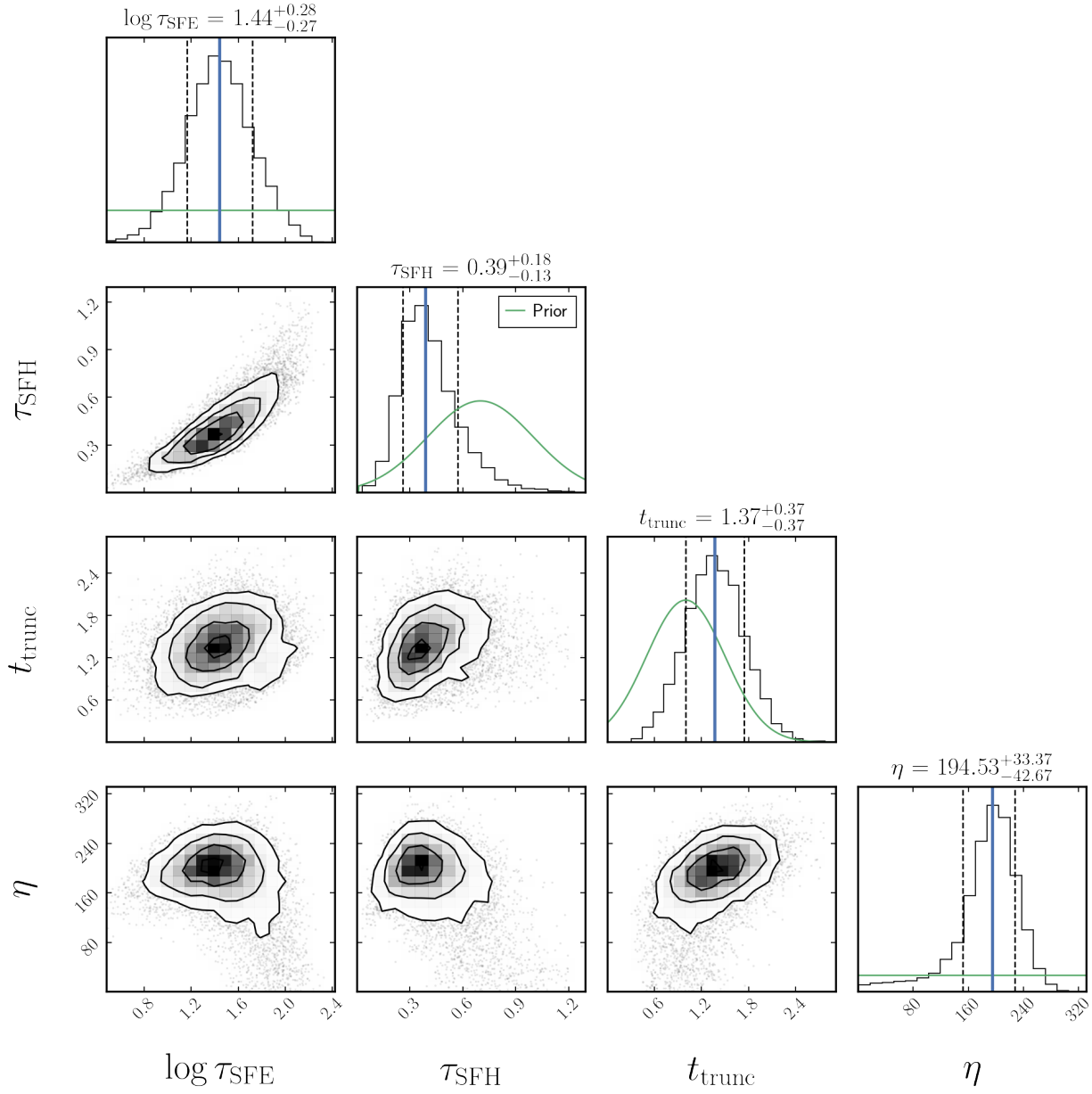


Figure 4.3: Posterior distribution corner plot of the model parameters $\log_{10} \tau_{\rm SFE}$, $\tau_{\rm SFH}$, $t_{\rm trunc}$, and η. Median values and 1σ uncertainties from the 16th and 84th percentiles are reported for each label and denoted by solid blue and dashed black lines respectively. The adopted prior distributions are included for reference as solid green lines. The Eri II MDF provides informative constraints on $\log_{10} \tau_{\rm SFE}$, $\tau_{\rm SFH}$, and η, while $t_{\rm trunc}$ remains prior-dominated.

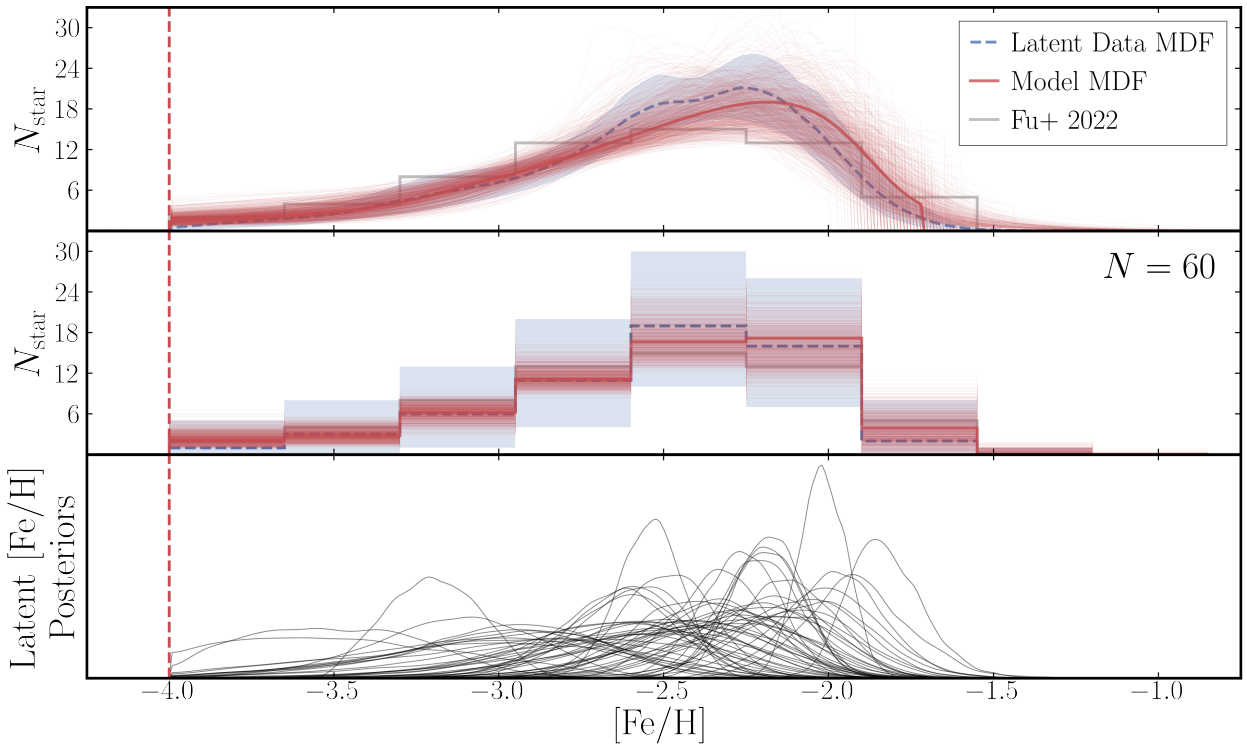


Figure 4.4: Top: Best-fit model MDF (solid red line; computed from the parameter values listed in Table 4.2) compared to the latent [Fe/H] MDF (dashed blue line). Uncertainties on the latent MDF (blue shaded regions) are estimated via bootstrapping (with replacement) from our sample of stars and their latent [Fe/H] posterior distribution. Model MDFs generated from 1000 random draws of the posterior distribution are displayed in thin red lines. We include the observed CaHK [Fe/H] MDF from F22 (solid gray line; same as in Figure 4.1) for reference but note that the model is not directly fit to this MDF. Middle: Same as the top panel but binned for comparison to the F22 CaHK MDF. Bottom: Updated posteriors distributions of [Fe/H] for each star in the sample. Compared to the CaHK posteriors presented in the bottom panel of Figure 4.1, long low-metallicity tails and the degree of truncation at [Fe/H] $= -4$ are substantially reduced because the model predicts that only a small fraction of stars have such low metallicities. Our fit indicates that it is unlikely that any of the stars in our sample are truly ultra metal-poor stars with [Fe/H] $\lesssim -4$.

Gallart et al. (2021). This supports the hypothesis of Gallart et al. (2021) that the true duration of the star formation burst is shorter than they could resolve with their CMD fitting techniques. Assuming star formation commenced immediately, the inferred $\tau_{\rm SFH}$ implies that Eri II had formed ∼65% of its stellar mass by $z \sim 11.5$ and ∼95% of its stellar mass by $z \sim 5.7$, which would independently confirm that Eri II is a relic of pre-reionization

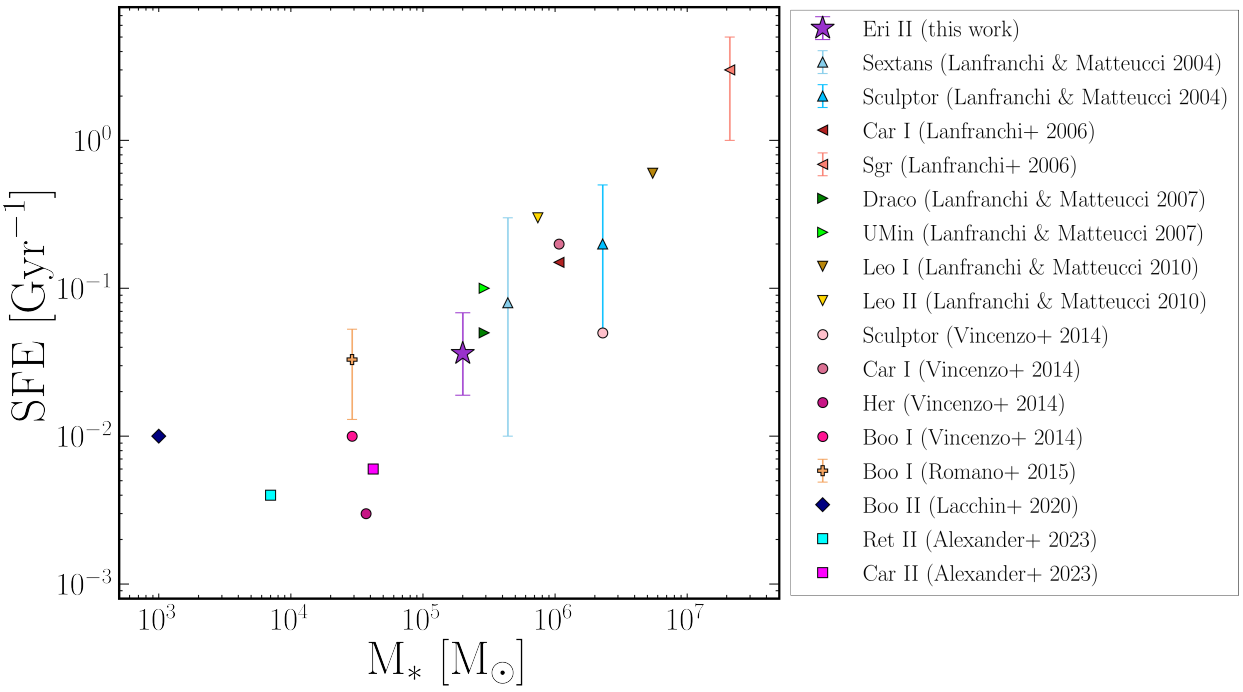


Figure 4.5: The inferred SFE of Eri II (purple star) compared to the SFEs reported by previous chemical evolution studies of LG dwarf galaxies. Though these studies employed a range of chemical evolution models, a clear relationship between galaxy stellar mass and SFE is apparent. The SFE inferred for Eri II is in good agreement with this relationship given its stellar mass.

era galaxy formation.

While we can place tight constraints on the SFH timescale, the inference of $t_{\rm trunc}$ ($t_{\rm trunc} = 1.37\pm^{0.37}_{0.37}$ Gyr; $z_{\rm trunc} = 4.40\pm^{1.25}_{0.81}$) remains dominated by the imposed prior. Tests allowing $t_{\rm trunc}$ to be unconstrained find no evidence in the Eri II MDF that the SFR truncated abruptly within the first 5 Gyr. That said, we know from the Eri II CMD that there has been effectively no star formation for the last $\sim$13 Gyr. Our inability to provide independent constraints on $t_{\rm trunc}$ is not indicative of tension between the MDF and CMD but rather a result of how subtle the impact of truncation is given the short star formation timescale of Eri II. In our best-fit model, truncation occurs after $\sim 3.5\tau_{\rm SFH}$, when the SFR is already quite low—the stellar mass of Eri II would only increase by $\sim$3% in the absence of truncation. In other words, the inferred exponential suppression of the SFH is already strong enough that a final, super-exponential truncation is difficult to detect. Constraining a sharp truncation in the SFH of Eri II from its MDF would require a larger sample of stars with abundance measurements precise enough to map the high-metallicity tail of the MDF.

Mass-Loading Factor We recover a broad but clearly peaked posterior for the mass-loading factor of Eri II, $\eta = 195^{+33}_{-43}$. This means that for every $1\,M_\odot$ of star formation, nearly $200\,M_\odot$ of ISM gas is ejected from the galaxy by SNe feedback. While extreme in comparison to the mass-loading factors of MW-like galaxies ($\eta \sim 1$), mass-loading factors of this magnitude are frequently invoked for low-mass galaxies in order to match simulations to empirical scaling relations (e.g., Benson et al. 2003; Somerville et al. 2015; Mitchell et al. 2020). State-of-the-art hydrodynamic simulations have also found $\eta \sim 100$ for the lowest-mass dwarf galaxies (e.g., Muratov et al. 2015; Emerick et al. 2019; Pandya et al. 2021).

In Figure 4.6, we compare our inferred mass-loading factor for Eri II (purple star) to the mass-loading factors inferred from chemical evolution studies of the disrupted dwarf galaxies *Gaia*-Sausage Enceladus (GSE) and Wukong/LMS-1 by Johnson, Conroy, et al. (2022) using the `VICE` one-zone chemical evolution model (red and blue circles respectively) and of the UFDs Carina II and Reticulum II by Alexander et al. (2023) using the `i-getool` inhomogenous chemical evolution model (magenta and cyan squares respectively). In addition, we include measurements of the mass-loading factor of galaxies from Chisholm et al. (2017) and McQuinn et al. (2019) (black diamonds and triangles respectively). Lastly, we include the dwarf starburst galaxy Pox 186, which was observed by Eggen et al. (2022) to currently have a suppressed mass-loading factor (open black pentagon) due to the efficient removal of gas by earlier SN-driven outflows. By estimating the total amount of gas lost to outflows, Eggen et al. (2022) concluded that the mass-loading factor of Pox 186 was substantially larger during its previous outflow episode (filled black pentagon).

Direct comparison between mass-loading factors measured through direct observational indicators (e.g., Chisholm et al. 2017; McQuinn et al. 2019) and chemical evolution models (Johnson, Conroy, et al. 2022; Alexander et al. 2023, and this work) is challenging for a number of reasons. For one, the manner in which outflows are parameterized in models frequently do not map directly to the observable quantities being measured. Additionally, the strength of outflows experienced presently at $z = 0$ by a galaxy of a given mass may not be representative of the outflows experienced at high redshift by galaxies of a similar mass. Nevertheless, our result for Eri II is in good qualitative agreement with the observed trend that less massive galaxies have stronger outflows.

Two scaling relationships have historically been invoked for relating the mass-loading factor of a galaxy to its stellar velocity dispersion, σ (as a proxy for its mass). In the physical scenario of momentum-driven winds governed by radiation pressure, the mass-loading factor scales as $\eta \propto V_c^{-1}$ (Murray et al. 2005). This scaling has been argued for by Finlator et al. (2008) and Peeples et al. (2011) based on the observed mass-metallicity relationship. Based on results from the FIRE-1 cosmological zoom-in simulation (Hopkins, Kereš, et al. 2014), Muratov et al. (2015) found a scaling relationship between the mass-loading factor and galactic stellar mass of $\eta = 3.6(M_*/10^{10}\ M_\odot)^{-0.35}$, which is in good agreement with the expectations given a momentum-driven wind scaling. Alternatively, in the physical scenario of energy-driven winds from SNe, the mass-loading factor scales as $\eta \propto V_c^{-2}$ (Chevalier et al. 1985). This scaling has been argued to be more important in low-mass galaxies with $\sigma < 75\,\mathrm{km\,s^{-1}}$ (e.g., Murray et al. 2005, 2010; Hopkins, Quataert, et al. 2012; Davé et al.

2013). Furthermore, Pandya et al. (2021) found a steeper scaling relationship in the updated FIRE-2 simulations (Hopkins, Wetzel, et al. 2018) of $\eta = 0.6(M_*/10^{10}\ \mathrm{M}_\odot)^{-0.45}$, which is more characteristic of the energy-driven wind scaling.

For reference, we have included both the momentum-driven scaling ($\eta \propto M_*^{-0.35}$) of Muratov et al. (2015) and the energy-driven scaling ($\eta \propto M_*^{-0.45}$) of Pandya et al. (2021) in Figure 4.6. The normalization of the chemistry-based mass-loading factors, including Eri II, is more in-line with the findings of Muratov et al. (2015). However, their scaling with stellar mass is marginally closer to that of Pandya et al. (2021), just offset by a small factor to larger values. A larger sample of galaxies, especially at the lowest masses, is required before these measurements can discriminate between these two physical outflow scenarios.

While smaller mass-loading factors ($\eta \lesssim 50$) are not prohibited by the model, they are 2σ disfavored and would require longer SFE and SFH timescales. Allowing for direct ejection of SN products (e.g., letting $f_{\mathrm{ret}} < 1$) has the potential to temper large mass-loading factors, but preliminary tests suggest that 1) mass-loading values of $\eta \sim 100$ are still preferred and 2) lower mass-loading factors require both low retention fractions ($f_{\mathrm{ret}} \sim 0.3$) and higher SFEs ($\tau_{\mathrm{SFE}} \sim 1.0$)—see the Metal-Loading model in Section 4.5.2. As with the τ_{SFE}, η could in principle vary with time as the dark matter halo of Eri II grows and its potential well deepened, though large changes in η over the duration of the star-forming lifetime are disfavored by its small τ_{SFH}. We leave investigation of a time-dependent η for future study.

4.5.2 Alternative Model Fits

Here we present the results of fitting the Eri II MDF with the alternative models described in Section 4.4.3. Median and 16th- and 84th-percentiles of the marginalized posteriors are presented alongside the fiducial best-fit values in Table 4.2. In Figure 4.7, we compare the MDF predicted for these median posterior values of each alternative model (colored lines) to that of the Fiducial model (black line) and its latent [Fe/H] distribution (gray dashed line and shaded region). We refer to these predicted MDFs as the "best-fit" for each model, though strictly speaking the model with the highest posterior probability does not have exactly the median posterior values of each parameter.

It is not entirely fair to judge the quality of the alternative model fits to the fiducial latent MDF, as each model may predict a different underlying distribution. In practice, however, we find that latent MDF of most models is quite similar to the fiducial case. The two exceptions to this are the latent distributions of the Constant SFR and High SFE models, which are more negatively skewed and centrally peaked respectively. While not statistically prohibited given the wide posteriors of the CaHK measurements, such underlying MDFs would be unusual in comparison to the MDFs observed in other dwarf galaxies using more precise spectroscopic abundances.

One quantitative metric for judging the goodness-of-fit of each model relative to the Fiducial model is the Bayes factor (BF). The BF is defined as the ratio of the Bayesian evidence of each model, which expresses the posterior probability of one model relative to the Fiducial model under the a priori assumption that both models are equally probable. A

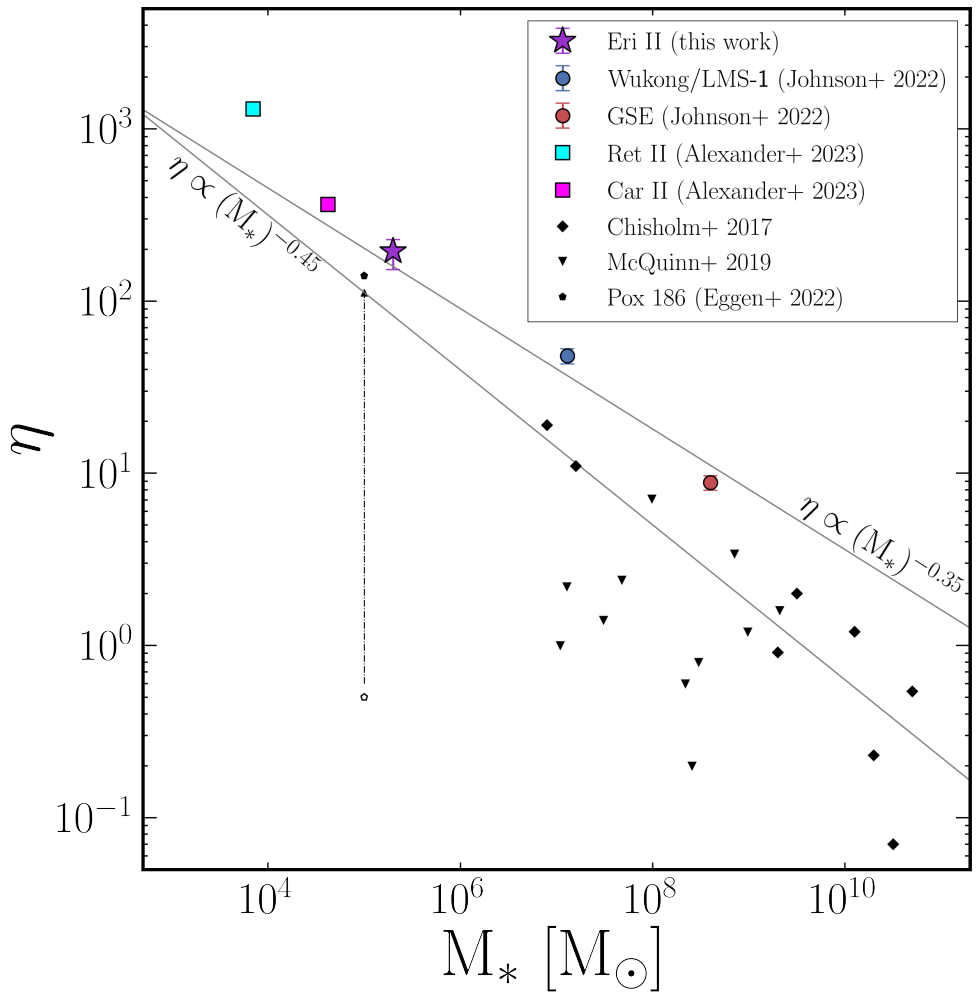


Figure 4.6: The inferred Mass-loading factor of Eri II (purple star) compared to the mass-loading factors inferred by the chemical evolution studies of Johnson, Conroy, et al. (2022) for Wukong/LMS-1 and GSE (blue and red circles respectively) and Alexander et al. (2023) for Car II and Ret II (magenta and cyan squares respectively) as a function of stellar mass. Mass-loading factors for galaxies observed by Chisholm et al. (2017) and McQuinn et al. (2019) are included as black diamonds and triangles respectively. The current observed mass-loading factor of Pox 186 and its previous estimated mass-loading factor from Eggen et al. (2022) are represented by the open and filled black pentagons respectively. The scaling found by Muratov et al. (2015) in FIRE-1 simulations indicative of momentum-driven winds ($\eta \propto M_*^{-0.35}$) and the scaling found by Pandya et al. (2021) in FIRE-2 simulations indicative of energy-driven winds ($\eta \propto M_*^{-0.45}$) are included for reference as solid black lines.

BF < 1 indicates that the Fiducial model is more probable while a BF > 1 indicates the alternative model is more probable. Using the Bayesian evidence estimated by `pocomc`, we calculate the BF for each model; we present these in the right-most column of Table 4.2.

In general, we find BFs < 1, indicating that the Fiducial model is preferred. However, we caution that the BF is inherently sensitive to choices in priors, so BFs of $\mathcal{O}(1)$ should not be over-interpreted. Nevertheless, because the BFs of the High SFE and Enhanced SN Ia are $\ll$1, we can be fairly confident that those models are disfavored. With a BF of $\sim$0.2, the Constant SFR model is also disfavored but to a more modest degree. One could argue that many of the alternative models require specific values for model parameters, making them a priori less likely. In this case, the BF we calculate would overestimate the probability of the alternative models relative to the Fiducial model. Definitively ruling out alternative models with BF ~ 1 would require additional observational constraints, including a larger sample of stars, more precise [Fe/H] measurements, and/or measurements of [α/Fe] (see Section 4.6.2).

We discuss each alternative model and their MDF predictions below.

Linear-Exponential SFR Model Despite the distinctly different functional form of the Linear-Exponential SFR model, we find that it provide a fit to the data which is just as good (if not slightly better) than that of the Fiducial model, while inferring very similar values for $\tau_{\rm SFE}$, $t_{\rm trunc}$, and η. Though the MDF predicted by this model is more symmetric, peaking at a slightly lower metallicity ([Fe/H] ~ -2.5) and extending to a higher metallicity ([Fe/H] ~ -1.5), the dispersion and mean metallicity of the MDF are in good agreement with the predictions of the Fiducial model. The Bayes factor slightly favors this fit over the Fiducial, but only marginally so. Discriminating between these two models would require a larger sample of high quality stellar metallicities than is currently available.

Our primary conclusion from this model is that our constraints on $\tau_{\rm SFE}$ and η from the fiducial model are *not* sensitive to the assumed SFH at early times. In particular, our findings are insensitive to the assumption of the Fiducial model that the gas reservoir was already in place at the onset of star formation, instead of growing rapidly through gas accretion as it does in the Linear-Exponential SFR model (see Section 4.4.1.4).

Constant SFR Model The Constant SFR model predicts an MDF that is rapidly increasing until it sharply truncates at the high-metallicity end. Such an MDF is atypical for dwarf galaxies in the Local Group, which universally exhibit MDFs that turn over at the high-metallicity end (e.g., Kirby, Lanfranchi, et al. 2011). The outflow mass-loading required to achieve the observed metallicity range is an order-of-magnitude below that of the Fiducial model because a constant SFR requires rapid continuing gas accretion that dilutes the metal production from stars. Relaxing the prior on $t_{\rm trunc}$ results in an earlier truncation ($\sim$0.6 Gyr) and a slightly smaller $\log_{10}\tau_{\rm SFE}$ ($\sim$1.7), but the shape of the predicted MDF and the inferred mass-loading factor remain largely the same. Although the predicted MDF shape is radically different from the F22 histogram and from the latent MDF of the Fiducial model, the Bayes Factor only mildly disfavors the Constant SFR mode. The uncertainties

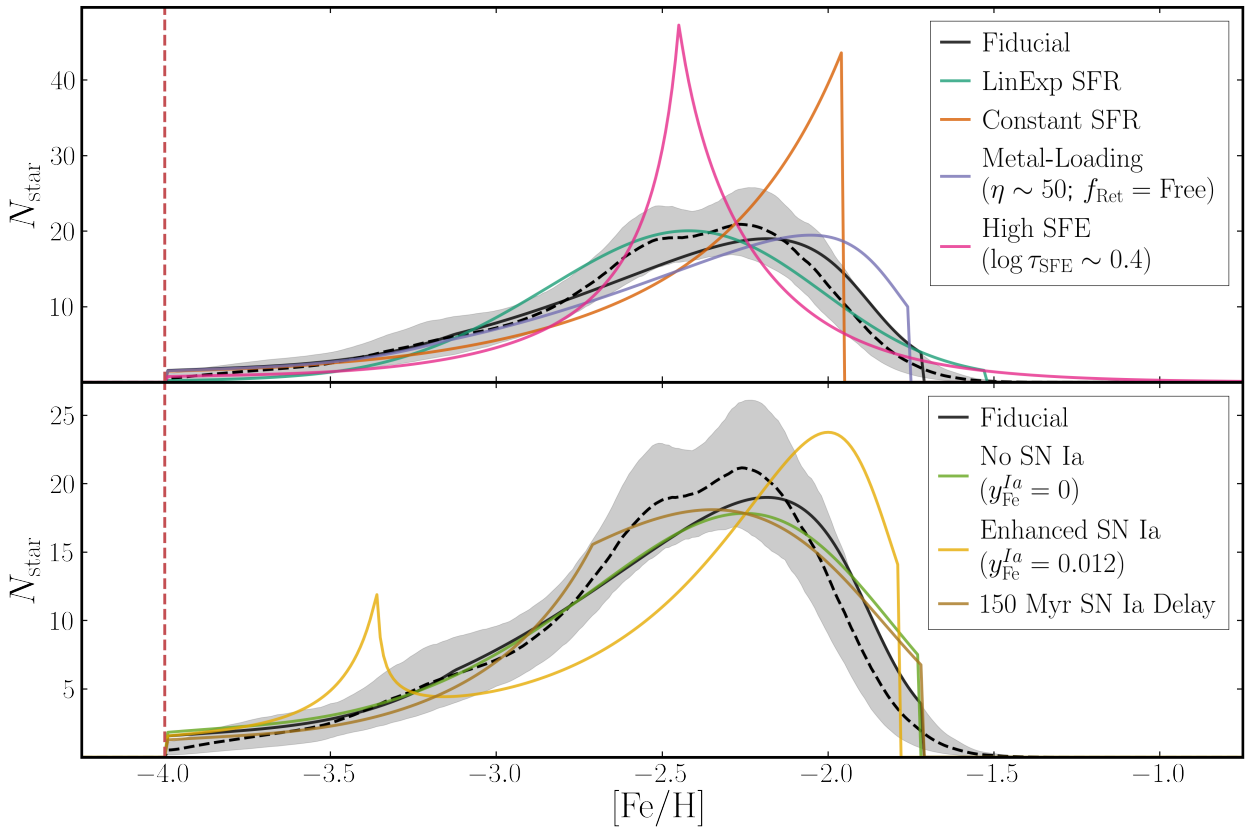


Figure 4.7: Best-Fit MDFs generated with the alternative models described in Section 4.4.3 (colored lines) compared to the best-fit MDF of the Fiducial model (solid black line) and latent [Fe/H] distribution (dashed black line and grey shaded region).

in the CaHK [Fe/H] values are large enough that a sharply truncated MDF is not in large tension with the measurements. More precise [Fe/H] values for the highest metallicity stars in Eri II would be needed to decisively distinguish the Fiducial and Constant SFR models.

Metal-Loading Model We find that forcing a lower mass-loading factor ($\eta \sim 50$) is able to produce realistic fits to the Eri II MDF if the retention fraction of SN products is low ($f_{\rm ret} \sim 0.3$)—that is, 70% of the metals produced by SNe are directly ejected from the galaxy. The covariance we find between η and $f_{\rm ret}$ in models where both are free implies that $(f_{\rm ret},\ \eta) = (0.3,\ 50)$ is roughly consistent with the mass-loading factor of $\eta \sim 200$ inferred in the Fiducial model when $f_{\rm ret} = 1$. Conversely, this covariance implies that to produce a similar MDF with an even smaller mass-loading factor of $\eta \sim 10$ would require a retention factor of only 10% (and also a substantially lower SFE timescale of $\log_{10} \tau_{\rm SFE} \sim 0.5$). While the best-fit MDFs of the Fiducial and Metal-Loading models are both generally consistent

with the data (the former more-so than the latter), they do produce qualitatively different MDFs. The Metal-Loading model predicts a slightly more skewed MDF with a higher metallicity peak and a truncated metal-rich tail. These differences result from the fact that the Metal-Loading model is always losing a significant fraction of the metals produced, while the Fiducial model only experiences significant metal losses once the ISM metallicity approaches its final value. Discriminating between these two models (i.e., breaking the degeneracy between η and $f_{\rm ret}$) can therefore be accomplished by acquiring more precise [Fe/H] measurements of stars in the high-metallicity end of the MDF.

High SFE Model When we force a short SFE timescale ($\log_{10}\tau_{\rm SFE} \sim 0.4$), we find that the best-fit MDF is sharply peaked at [Fe/H] ~ -2.4 with roughly exponential tails on either side. The peak of this distribution corresponds to the CC equilibrium Fe-abundance that the model would evolve to in the absence of SN Ia (see WAF17 for derivation):

$$[\mathrm{Fe/H}]^{\rm cc}_{\rm eq} = \log_{10}\left(\frac{y^{\rm cc}_{\rm Fe}}{Z_{\rm Fe,\odot}(1+\eta-r-\tau_{\rm SFE}/\tau_{\rm SFH})}\right). \tag{4.10}$$

For a galaxy with an exponentially declining SFH, this equilibrium value is approached on the "harmonic difference timescale" set by $\tau_{\rm SFH}$ and the gas depletion timescale $\tau_{\rm dep}$,

$$\tau^{\rm cc}_{\rm Fe,eq} \equiv \left(\tau^{-1}_{\rm dep} - \tau^{-1}_{\rm SFH}\right)^{-1}, \tag{4.11}$$

where the depletion timescale is defined to be

$$\tau_{\rm dep} = \frac{\tau_{\rm SFE}}{1+\eta-r}. \tag{4.12}$$

Equation 4.11 can be equivalently expressed as

$$\tau^{\rm cc}_{\rm Fe,eq} = \frac{\tau_{\rm SFE}}{1+\eta-r-\tau_{\rm SFE}/\tau_{\rm SFH}}. \tag{4.13}$$

While both this model and the Fiducial model have $[\mathrm{Fe/H}]^{\rm cc}_{\rm eq} \sim -2.4$, the short $\tau_{\rm SFE}$ imposed here results in a substantially smaller $\tau^{\rm cc}_{\rm Fe,eq}$ ($\sim$30 Myr vs. $\sim$200 Myr), which is shorter than the 50 Myr minimum time delay for SN Ia. As a result, the model quickly evolves to the equilibrium metallicity where it forms stars until the onset of SN Ia at which point the model evolves to higher metallicity. The sharp decline in the MDF above $[\mathrm{Fe/H}]^{\rm cc}_{\rm eq}$ is due to the short SFH timescale (0.08 Gyr) inferred for this model—by the time SN Ia begin increasing the metallicity, the rate of star formation is rapidly declining. In contrast, the longer $\tau^{\rm cc}_{\rm Fe,eq}$ of the Fiducial model is sufficiently long that SNe Ia begin contributing to Fe production before $[\mathrm{Fe/H}]^{\rm cc}_{\rm eq}$ is reached, resulting in a smoother MDF with no sharp peaks. The High SFE model is strongly disfavored by the Bayes factor ($\sim$0.002), and its extremely short $\tau_{\rm SFH}$ appears physically implausible.

Longer SN Ia Delay Within the uncertainties of the latent MDF, we find that a model assuming a minimum SN Ia time delay of 0.15 Gyr provides a fit that is roughly as good as that of the Fiducial model. With this longer time delay, we infer a SFH timescale that is slightly larger and a mass-loading factor that is slightly smaller than the Fiducial model. The kink in the evolutionary track at [Fe/H] ~ -2.75 corresponds to the onset of SN Ia enrichment (a milder version of the sharp transition found in the High SFE model). In principle one could distinguish the $t_D = 0.05$ and 0.15 Gyr scenarios from the different shapes of the predicted MDFs, but this would require a larger sample of high quality stellar metallicities than is currently available.

No SN Ia We find that a model with no SN Ia enrichment can reproduced the Eri II MDF reasonably well with only slight differences from the MDF of the Fiducial model. Like the Metal-Loading model, this model is most distinguishable in the high-metallicity tail. While this model infers values for the SFE and SFH timescales and $t_{\rm trunc}$ that are consistent with the Fiducial model, it requires a mass-loading factor that is $\sim$2.5 times smaller because the total Fe yield is lower. Unlike the other best-fit models, this model has a SFH timescale that is *shorter* than the depletion timescale $\tau_{\rm dep}$ (Equation 4.12), 0.45 Gyr vs. 0.52 Gyr, which is necessary in the absence of SN Ia enrichment to avoid forming an MDF peaked at $[{\rm Fe/H}]_{\rm eq}^{\rm cc}$. Physically achieving $\tau_{\rm SFH} < \tau_{\rm dep}$ would require the removal of gas from the galaxy by a process not associated with star formation. If this is indeed the case for Eri II, then reionization-driven photo-evaporation might be responsible for the removal of gas, though this would require additional investigation.

While the Bayes factor of this model is marginally larger than the Fiducial model, we caution that this alone does not indicate that the No SN Ia model is better. The BF is only informative insofar as the two models are equally likely a priori. The scenario considered here, in which no SN Ia contributed in any part to the Fe enrichment of the stars in our sample, is highly improbable given the expectation that $\sim$200 SN Ia should have occurred in a galaxy of Eri II's mass. The No SN Ia model could be easily distinguished from the Fiducial model with measurements of [α/Fe] ratios, which should remain elevated in the absence of SN Ia enrichment (see Section 4.6.2.1).

Enhanced SN Ia Increasing the SN Ia yield by a factor of 10, as a SN Ia rate $\propto Z^{-0.5}$ would imply, results in an MDF with a higher and narrower peak close to the eventual sharp truncation, as well as a secondary low-metallicity peak. The high $y_{\rm Fe}^{\rm Ia}$ forces a high η, which in turn leads to a short depletion time $\tau_{\rm dep} \sim 30$ Myr. As with the High SFE model, this low-metallicity peak is the result of the model evolving to its equilibrium CC Fe abundance (in this case $[{\rm Fe/H}]_{\rm eq}^{\rm CC} \sim -3.25$) before the commencement of SN Ia at $t_D = 50$ Myr. The Enhanced SN Ia model is disfavored by the Bayes factor ($\sim$0.02).

The discrepancy on the high-metallicity end of the MDF is alleviated if we allow for $f_{\rm ret} < 1$, which compensates by decreasing the effective SN yield. Indeed, if we set $f_{\rm ret} = 0.1$ for SN Ia but $f_{\rm ret} = 1$ for CCSN, then this model is equivalent to the Fiducial model, with

direct SN Ia metal loss exactly cancelling the higher yield. However, there is no obvious reason to have a high retention fraction for CCSN but a low retention fraction for SN Ia. We do find that adopting a single $f_{\rm ret} = 0.1$ produces an MDF in better agreement with the Fiducial model. While the best-fit mass-loading factor in this case is similar to that of the Fiducial model ($\eta \sim 200$), the inferred SFE timescale is substantially lower ($\log_{10} \tau_{\rm SFE} \sim 0.7$). If we allow $f_{\rm ret}$ to be free, the model prefers larger retention fractions and mass-loading factors: $f_{\rm ret} \sim 0.6$ and $\eta \sim 700$ (though these parameters, as always, are very degenerate). In all of these permutations, the peak around $[{\rm Fe/H}]_{\rm eq}^{\rm CC}$ remains.

4.6 Discussion

4.6.1 Physical Interpretation of the Model

Our Fiducial model achieves a good match to the observed MDF with physically plausible values of its four evolutionary parameters, $\tau_{\rm SFE} = 27.5\,{\rm Gyr}$, $\tau_{\rm SFH} = 0.39\,{\rm Gyr}$, $t_{\rm trunc} = 1.37\,{\rm Gyr}$, and $\eta = 194$. As previously discussed, a low SFE (large $\tau_{\rm SFE}$) is characteristic of low-M_* dwarfs, and the high η value is consistent with scaling relationships from numerical and analytic models extrapolated to the low mass of Eri II. The e-folding timescale for star formation is consistent with direct estimates of the SFH (Gallart et al. 2021), but the value of $t_{\rm trunc}$ is not independently well constrained by the MDF data.

Figure 4.8 elucidates the evolution of the best-fit Fiducial model. Because of the short $\tau_{\rm SFH}$, the model has already formed $\sim$65% of its stellar mass by $t = 0.4\,{\rm Gyr}$ ($z = 11.3$). The gas mass, $M_g = \tau_{\rm SFE}\dot{M}_*$, follows the same exponential decline as the SFH, given the assumption of a constant $\tau_{\rm SFE}$. Through most of the evolution the stellar mass fraction $M_*/(M_* + M_g)$ is $\ll$1, though by the end it has risen to 0.5. However, because the value of η is so high, the mass of gas ejected from the galaxy exceeds the mass remaining in the ISM at all times $t \gtrsim 0.1\,{\rm Gyr}$. The model has vigorous ongoing gas accretion that fuels continuing star formation despite the strong outflow, with an infall rate $\dot{M}_{\rm inf} \approx (\eta - \tau_{\rm SFE}/\tau_{\rm SFH})\dot{M}_* \approx 120\dot{M}_*$ (see Equation 4.4). An exponential SFH requires a non-zero gas mass at $t = 0$. In the Fiducial model, the mass of gas accreted exceeds this initial mass for $t \gtrsim 0.35\,{\rm Gyr}$.

The bottom panel of Figure 4.8 tracks the Fe mass budget. In the Fiducial model, the total Fe produced by CCSN and SN Ia is nearly equal over the life of the galaxy. However, CCSN enrichment dominates the early evolution, and by late times the enrichment rate from SN Ia greatly exceeds that from CCSN, as one can see by comparing the slopes of the blue and red curves. These conclusions rely on our adopted values of $y_{\rm Fe}^{\rm cc}$ and $y_{\rm Fe}^{\rm Ia}$; as shown in Figure 4.7, the MDF can be reasonably well reproduced even in a model with no SN Ia enrichment. For $t > 0.4\,{\rm Gyr}$, the mass of Fe ejected from the galaxy exceeds the mass remaining in the ISM by a substantial factor. The Fe mass in stars is small compared to that in the ISM because the star-to-gas mass ratio is low and because the mean metallicity of stars is always lower than the ISM metallicity.

Although the SFE timescale is long, the gas depletion time (Equation 4.12) is short

because of the high η, $\tau_{\rm dep} \sim 0.14\,{\rm Gyr}$. As discussed by WAF17, obtaining an MDF that turns over rather than peaking sharply requires a rapidly declining SFH so that newly produced metals are deposited in a dwindling gas supply, resulting in high ISM metallicity at late times when only a small number of stars are produced. In the case of the Fiducial model, $\tau_{\rm SFE} = 0.39\,{\rm Gyr}$ is still significantly longer than $\tau_{\rm dep}$, but it is shorter than the characteristic enrichment time for SN Ia (roughly 1.5 Gyr), so the SN Ia enrichment drives the MDF turnover. In the $y_{\rm Fe}^{\rm Ia} = 0$ model, by contrast, the MDF turnover arises because the depletion time is longer, and with $\tau_{\rm SFH} \sim \tau_{\rm dep}$ the model approaches the leaky box scenario, which was shown to provide a reasonable fit to the Eri II MDF by F22. Our Fiducial model is still rather far from this limit.

We can summarize the physical properties of the Fiducial model as follows. It begins with an initial gas mass $M_g \sim 1.6 \times 10^7\,{\rm M}_\odot$ and accretes gas vigorously but at an exponentially declining rate. Only a small fraction of the accreted gas forms into stars because the SFE is low and because feedback from star formation drives ISM gas out of the shallow potential well with a high mass-loading factor $\eta \sim 200$. Fe enrichment is dominated by CCSN at early times and by SN Ia at later times, with the two channels producing similar total amounts of Fe over the life of the galaxy. However, more than 90% of the Fe produced by the stars is ejected from the galaxy. This low metal retention is the main reason for the low final metallicity, not the truncation of star formation. The turnover in the MDF arises because Fe from SN Ia is deposited in a dwindling gas supply, enabling a small fraction of stars to form at relatively high [Fe/H]. The exponentially declining SFH arises because gas accretion does not keep up with gas losses from the feedback-induced galactic wind. Star formation ceases abruptly at $t = t_{\rm trunc} \sim 1.4\,{\rm Gyr}$ ($z \sim 4.3$), presumably because reionization evaporates the remaining gas supply.

4.6.2 Additional Model Predictions

4.6.2.1 Predictions of [Mg/Fe]

While we do not consider [Mg/Fe] in our fit (there being presently no stars in Eri II with [α/Fe] measurements of any kind), we can use our model to make predictions of the [Mg/Fe] evolution in Eri II that next-generation spectroscopic facilities will soon be able to test (e.g., Sandford, Weisz, et al. 2020). In principle, measurements of [Mg/Fe] for even a few stars in our sample should provide tighter constraints on our posteriors.

In Figure 4.9 (top), we display the distribution of stars in [α/Fe]-[Fe/H] space that is compatible with our fiducial Eri II model (gray-scale histogram). This distribution is generated by sampling 60 stars from each of the 1000 randomly drawn posterior predictions included in Figure 4.4. Evolutionary tracks of these models are also included here as thin red lines. Red circles depict 10 Myr snapshots of the best-fit model, where the size of the marker is proportional to the relative SFR at that time; time-steps corresponding to 10 Myr, 100 Myr, and 1 Gyr are outlined in black.

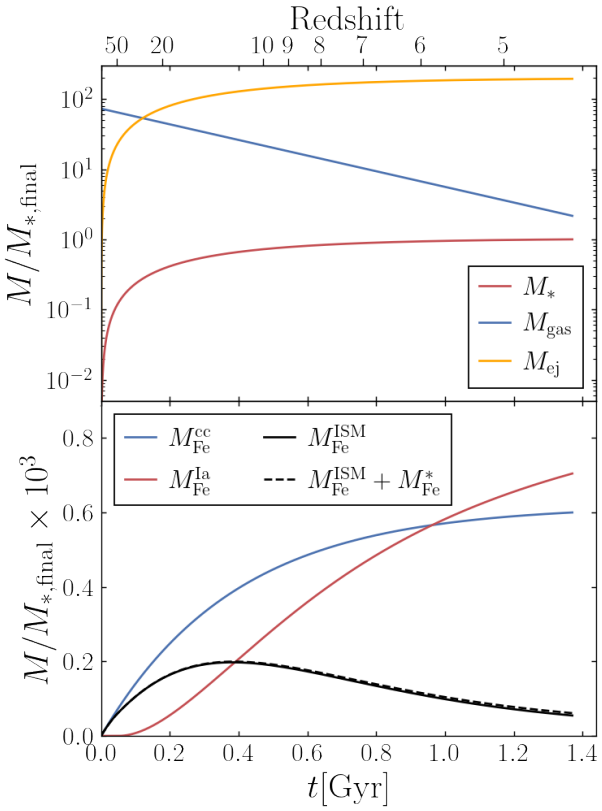


Figure 4.8: Top: The stellar mass (red), gas mass (blue), and cumulative ejected gas mass (orange) of the best-fit model as a function of time. The star-to-gas ratio is very low at early times but evolves close to unity by the time star formation ceases. At the end of the simulation, the model has lost roughly 100× its stellar mass in gas outflows. Bottom: The mass of Fe in the ISM (solid black) and the cumulative mass of Fe produced by CCSNe and SN Ia (blue and red respectively) for the best-fit model as a function of time. At early times, CCSNe dominate the production of Fe, while at late times the total contribution of CCSNe and SN Ia are roughly equal. While the amount of Fe in the ISM is slowly decreasing for $t > 0.3\,\mathrm{Gyr}$, the rapidly decreasing gas supply (top panel) results in a monotonically increasing gas-phase metallicity. The mass of Fe in stars is negligible as can be seen from the dashed black line, which shows the combined mass of Fe in both the ISM and long-lived stars.

As described in the Section 4.4.1, the low-metallicity plateau of [Mg/Fe] ~ 0.5 is produced by design given our adopted CC Mg and Fe yields. The turn-over or "knee" in the [Mg/Fe]-[Fe/H] distribution occurs at [Fe/H] ~ -3.0, which is the metallicity of the model at $t = 50\,\mathrm{Myr}$ when SN Ia begin to contribute to Fe production. The SN Ia Fe yield was set to

$y_{\rm Fe}^{\rm Ia} = 0.0012$ such that a MW Disk-like model evolves to [Fe/H] $\approx$ [Mg/Fe] ≈ 0 at late times. Unsurprisingly, the evolution of a UFD-like model presented here only evolves to [Fe/H] ~ -1.75, but it reaches sub-solar Mg abundances of [Mg/Fe] ~ -0.2 because of the short $\tau_{\rm SFH}$. [Mg/Fe] measurements in UFDs are sparse and uncertain, especially for [Fe/H] $\lesssim -3$, but our predictions are generally consistent with observations (e.g., Simon 2019, and references therein).

Though the locus of possible [Mg/Fe]-[Fe/H] falls relatively tightly around the best-fit Fiducial model, the spread in the posterior predictions is substantial both in the location of the knee and the final [α/Fe]. This suggests, similar to the comparison of alternative models in Figure 4.7, that the more metal-rich stars in Eri II hold increased constraining power. Fortunately, these are also the stars for which spectroscopic measurements should be (comparatively) easier. That being said, measuring [Mg/Fe] in stars with [Fe/H] < -3.0 will provide valuable constraints on the CC SN yields that determine the high-[Mg/Fe] plateau.

The constraining power of [Mg/Fe] measurements is further exemplified in the bottom panel of Figure 4.9, where the [Mg/Fe]-[Fe/H] evolution of the alternative models are compared to that of the Fiducial model. The No SN Ia model (green) is easily distinguishable from the Fiducial model (black) because without SN Ia enrichment [Mg/Fe] remains elevated. Meanwhile the Enhanced SN Ia model (yellow) is distinguishable for the opposite reason because the extra SN Ia enrichment drives [Mg/Fe] lower faster. The low SFE of the Constant SFR model (purple) leads to SN Ia decreasing [Mg/Fe] at lower [Mg/Fe] than other models, while its low mass-loading factor means more of the CCSN products produced early are retained, keeping [Mg/Fe] from decreasing as steeply. In the High SFE model (pink), [Fe/H] evolves much more rapidly so the knee occurs at higher metallicity, but given the short $\tau_{\rm SFH}$ inferred for this model, few stars are formed at lower [Mg/Fe]. The differences predicted by the Longer SN Ia Delay model (brown) and the Metal-Loading model (orange) compared to the Fiducial model are smaller. The longer time delay before SN Ia start shifts the [Mg/Fe]-[Fe/H] track to higher metallicities, while the direct loss of metals makes SN Ia slightly less effective at decreasing [Mg/Fe] at late times. Once again, precise abundance measurements of stars at the high-metallicity end will provide the best opportunity to discriminate between these models.

4.6.2.2 Ultra Metal-Poor Stars

Our hierarchical Bayesian framework enables us to recover the posterior distribution of the latent [Fe/H] for each star in our sample. We caution, however, that these inferred values are influenced by the model MDF and represent the "true" [Fe/H] of each star only insofar as the model represents the true MDF of Eri II. Further, because a truncation of the MDF and [Fe/H] priors below [Fe/H] < -4 is imposed by the limitations of the stellar grid used in the CaHK measurements, we cannot recover the metallicity of a star to be more metal-poor than [Fe/H] < -4 even if such a star was in our sample. That being said, our framework does allow us to compare the relative probability that each star in our sample was drawn from the un-truncated MDF and the redistributed metal-poor tail of the MDF (see Figure

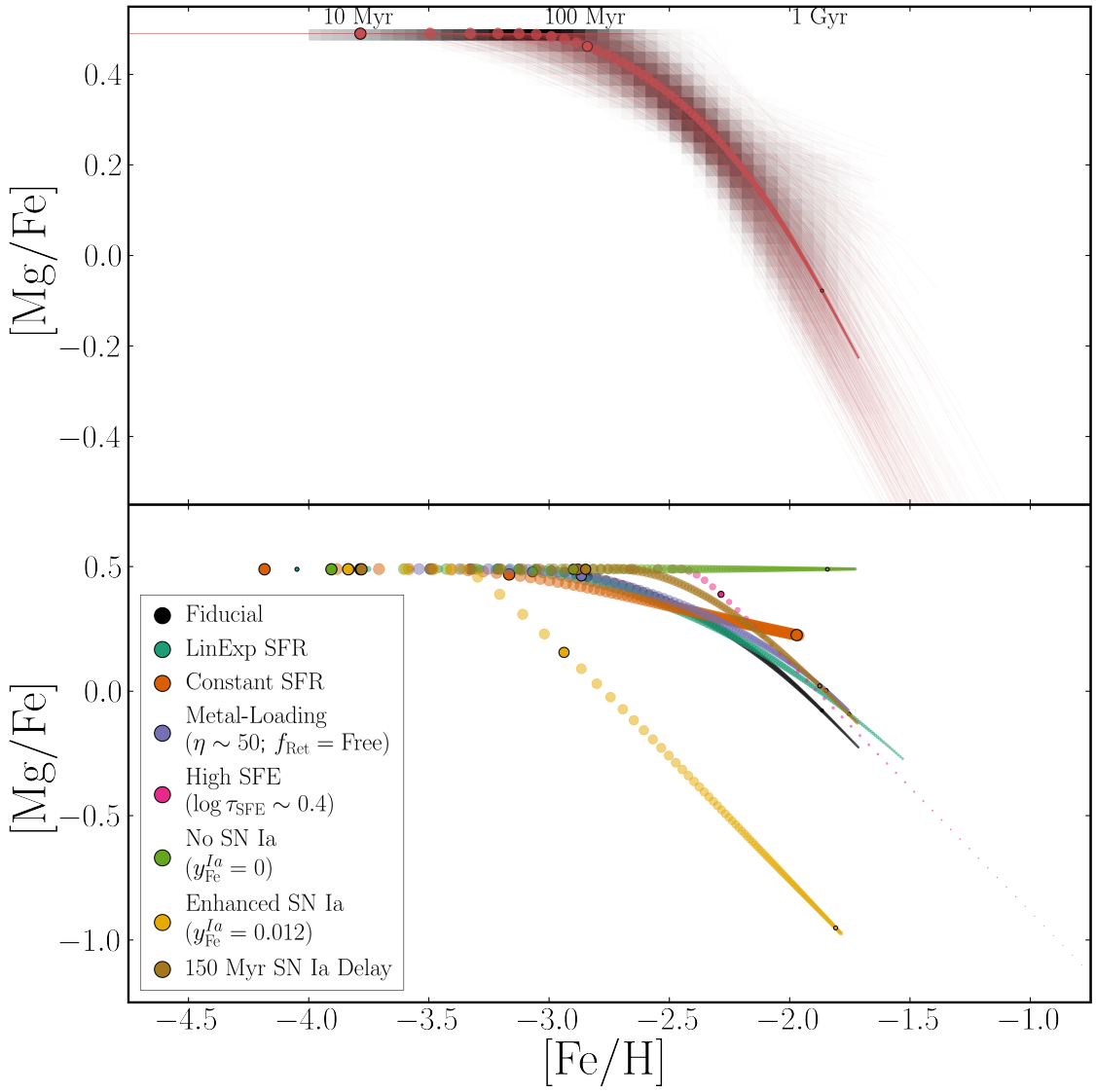


Figure 4.9: Top: [Mg/Fe]-[Fe/H] distribution of stars (gray-scale histogram) predicted by models sampled from our posterior distribution (thin red lines). The evolution of the best-fit Fiducial model in steps of 10 Myr is depicted in red circles; the size of the marker is proportional to the SFR at each step. Bottom: The evolution of the best-fit alternative models (colored circles) compared to the Fiducial model evolution (black circles) following the same plotting convention as the top panel. While many of the alternative models produce MDFs similar to the Fiducial model, they predict quite distinct [Mg/Fe]-[Fe/H] distributions.

4.2). By doing this for the entire posterior sample, we can infer the probability that each star is truly an ultra-metal poor star with [Fe/H] < -4.

Figure 4.10 shows the probability of being an ultra-metal poor star for the 10 highest probability stars. We find that the posterior distribution on $P([\mathrm{Fe/H}] < -4)$ is consistent with zero for every star, strongly disfavoring the presence of any UMP star in our sample. This result suggests that no pre-enrichment of the gas supply or metallicity floor is necessary to explain the dearth of UMP stars, though a larger sample of stars is necessary to conclusively rule out these scenarios. We place 95% upper limits on the ultra-metal poor probability of each star and find 10 stars with upper limits greater than 15%. Stars 11 and 21 have upper limits greater than 40% and were previously identified as extremely metal-poor ([Fe/H] < -3) candidates by F22. Searches for UMP stars in Eri II should prioritize this subset of our sample for spectroscopic follow-up in order to confirm their true metallicity.

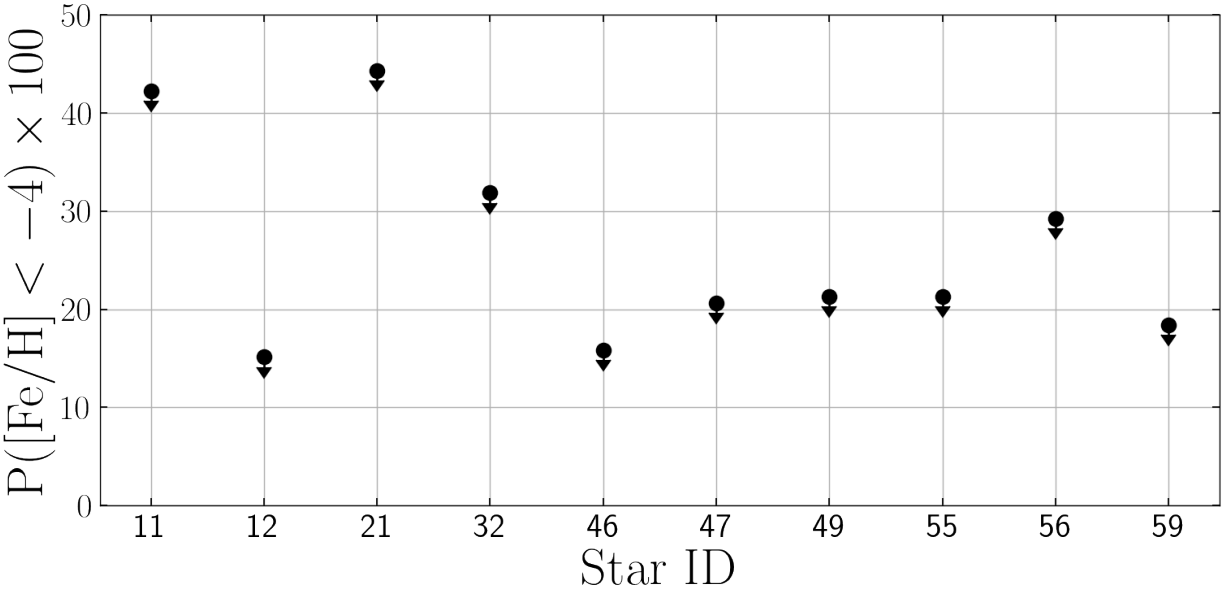


Figure 4.10: 95% upper limits on the inferred probability of metallicities below [Fe/H] < -4 for the highest probability stars in our sample. The remaining 50 stars have probabilities <15%. Stars 11 and 21 were previously identified as extremely metal-poor ([Fe/H] < -3) candidates by F22.

4.6.3 Potential Limitations of a One-Zone Model

The analytic solutions employed in our analysis require several idealizations, such as constant values of τ_{SFE} and η. The most important idealization is the one-zone framework itself, i.e.,

the assumption that the star-forming ISM can be treated as a single, fully-mixed gas reservoir with abundances that evolve in time but do not vary with position. Our key finding is that the Fiducial model reproduces the observed MDF of Eri II with parameter values that appear plausible on empirical and theoretical grounds. More complicated models for the chemical evolution of Eri II are certainly possible, but they are not required by the observed MDF.

The sharpest conclusion from our modeling is that the observed MDF implies a rapidly declining SFH, with $\tau_{\rm SFH} \sim 0.4\,{\rm Gyr}$, in addition to an eventual truncation. This conclusion is driven by the turnover and slow decline of the MDF, as opposed to the high peak and rapid cutoff predicted for models with roughly constant SFR (WAF17). As discussed in Section 4.5.2, clearly ruling out this model would require more precise [Fe/H] values for the highest metallicity stars in Eri II, but the predicted MDF shape for a constant SFR clearly differs from that inferred in most studies of low-luminosity dwarfs. Returning to Figure 4.4, we note that the portion of the MDF beyond the maximum at [Fe/H] ~ -2.25 is populated by by stars whose posterior p([Fe/H]) peaks at higher [Fe/H], not by the long tails of stars whose most probably [Fe/H] is lower. Thus, there is no indication in the data that the smooth turnover of the MDF (as opposed to a sharp cutoff) is caused by observational scatter.

There are two ways in which departures from a one-zone model could explain a turnover in the MDF without a rapidly declining SFH or otherwise bias our results: (1) spatial metallicity gradients (2) stochastic enrichment events. We discuss these scenarios below, but note that either of these would require additional degrees of freedom to model. If it remains consistent with future data, the parsimony of the 4-parameter one-zone model is an argument in its favor.

4.6.3.1 Spatial Gradients

Many, though not all, dwarf galaxies are known to host mild radial stellar metallicity gradients of $\nabla_{\rm [Fe/H]} \sim -0.1\,{\rm dex}/r_h$ (see Taibi et al. 2022, and references therein). The presence of a metallicity gradient in Eri II could impact our results in one of two ways. Our sample may be biased to higher metallicity because our CaHK measurements only include stars within $\sim$1 r_h thereby missing the most metal-poor stars at large radii. Alternatively, the shape of the Eri II MDF may be altered by the inclusion of stars at a range of radii that do not share identical chemical enrichment histories, thus violating the assumption of a one-zone chemical evolution model.

In most cases, the metallicity gradients of dwarf galaxies are thought to be primarily generated by feedback-driven outflows, which heat stellar orbits and preferentially drive outward migration of old stars (El-Badry, Wetzel, et al. 2016). Because the oldest stars are also likely to be the most metal-poor, this migration can create a negative stellar metallicity gradient with more metal-rich stars at small radii and more metal-poor stars at large radii. While El-Badry, Wetzel, et al. (2016) find feedback-driven stellar migration to be more pronounced for low-mass galaxies in a slightly larger mass regime ($M_* \sim 10^{7-9.6}$)[4], we cannot

[4]Specifically, El-Badry, Wetzel, et al. (2016) find that the higher dark matter fractions and lower SFE

entirely rule out the possibility that our sample of stars in Eri II has been impacted by radial migration.

Alternatively, radial metallicity gradients may be indicative of radial gradients in galactic physics. For example, if the outflow mass-loading increased with radius or the SFE decreased with radius, then the central regions of the galaxy could evolve to higher [Fe/H] than the outer regions. In principle, this could produce a small fraction of stars with [Fe/H] beyond the peak of the MDF, creating the turnover we see in the Eri II MDF without needing to invoke an exponentially declining SFH. We have not experimented with such models, but with freedom to choose the density profile and $\eta(r)$ or $\tau_{\rm SFE}(r)$ we expect one could produce a range of MDF shapes. In chemical evolution studies of more massive galaxies like the Milky Way, radial variations in model parameters is indeed important and has motivated the replacement of the single one-zone model with a series of concentric one-zone models, each representing a radial annulus of the galaxy (e.g., Matteucci and Francois 1989; Schönrich et al. 2009; Minchev et al. 2013; Johnson, Weinberg, et al. 2021; Sharma et al. 2021). However, such effects are likely to be much less important for UFDs like Eri II, which formed the bulk of their stars on very short timescales and very small spatial scales.

So far, only two studies, Martínez-Vázquez et al. (2021) and F22, have attempted to measure a stellar radial metallicity gradient in Eri II. Using a sample of 67 RR Lyrae stars, Martínez-Vázquez et al. (2021) measured a strong negative metallicity gradient of $-0.46\,{\rm dex}/r_h$ in the inner half-light radius of Eri II. However, a gradient of this magnitude is highly unusual for an isolated dwarf galaxy with the mass of Eri II and is more characteristic of dwarf galaxies known to have experienced a past merger event (e.g., Sextans, Andromeda II, Phoenix, NGC 6822; Taibi et al. 2022). Moreover, F22, which provides the observational basis for our analysis, found no evidence for a spatial trend in stellar [Fe/H] within one half-light radius where the gradient was reported to be strongest by Martínez-Vázquez et al. (2021). The origin of this discrepancy remains uncertain and merits future investigation. Nevertheless, because no radial metallicity gradient exists within our sample, we do not believe the shape of our MDF to be substantially altered by the presence of spatial gradients in Eri II.

It is still possible that we are biased by our centrally concentrated sample, which may not include old, metal-poor stars that formed or migrated beyond the inner half-light radius. A more spatially extended survey of stellar metallicities in Eri II is necessary in order to rule this possibility out. If it turns out that our current sample is missing a sizeable population of metal-poor stars, then the SFE or SFH timescale we infer may be biased high. A more spatially extended survey of stellar metallicities in Eri II is necessary in order to quantify the magnitude of the bias or rule this possibility out.

---

of less massive galaxies leads to smaller fluctuations in the galactic potential and therefore weaker coupling between feedback-driven outflows and stellar kinematics.

4.6.3.2 Stochastic Supernova Enrichment

A second possible departure from our model assumptions is Poisson sampling of the supernova population. For $M_* \sim 2 \times 10^5\,\mathrm{M}_\odot$ the number of CCSN is ∼2000 and the number of SN Ia is ∼200. If the reservoir is fully mixed, as appears to be a reasonable assumption for ancient dwarf galaxies (see Escala, Wetzel, et al. 2018), then Poisson fluctuations would produce only minor variations in the enrichment history, at least at the high metallicity end of the MDF. However, if the galaxy is divided into smaller zones that do not efficiently share metals with each other then the number of supernovae that contribute to the composition of any given star is smaller. In this scenario, the high metallicity tail of the MDF could be populated by stars that happened to be enriched by unusually large numbers of supernovae—most likely SN Ia because of their smaller numbers and larger Fe yield per supernova, though stochastic sampling of the IMF at late times when the SFR is low may also contribute to fluctuations in the number of CCSNe. In principle this scenario could be tested by measuring stochastic fluctuations in element ratios, following the arguments presented by Griffith et al. (2023). While stochastic sampling may be a small effect in this study of Eri II, it is likely to be more important in lower mass UFDs where the total number of CCSN and SN Ia may be smaller and in analyses that involve additional element ratios (e.g., Alexander et al. 2023).

4.6.4 Comparison to Johnson, Conroy, et al. (2022)

Although the formulation is quite different, our method has features in common with the recently proposed method of Johnson, Conroy, et al. (2022, hereafter J22), which also fits dwarf galaxy abundance data with one-zone chemical evolution models. The J22 method considers the probability that each star can be associated with each point on a model evolutionary track. Weighting these probabilities by the model SFR enforces a good match to the MDF of the data set. Our method works directly from the MDF, though the treatment of measured $P([\mathrm{Fe/H}]|\mathrm{CaHK})$ as a prior on the latent $P([\mathrm{Fe/H}]')$ of each star makes the calculation resemble the likelihood calculation of J22. J22 consider data with both [α/Fe] and [Fe/H] measurements, and the [α/Fe] turnover provides leverage on the model timescales given the DTD of SN Ia enrichment. For Eri II we have been able to derive surprisingly strong constraints from $P([\mathrm{Fe/H}])$ alone, though we are aided by the turnover form of the MDF and by the known early truncation of star formation.

We suspect, but have not yet tested, that the two methods would give similar results from equivalent input data. We use the WAF17 analytic solutions while J22 use numerical computations from `VICE` (Johnson and Weinberg 2020), but in principle either method could be implemented using analytic or numerical chemical evolution calculations. Our method could be generalized to model a joint $P([\alpha/\mathrm{Fe}], [\mathrm{Fe/H}])$ distribution, but the J22 method may be simpler to implement when multiple observables per star are included. Conversely, our approach may be better adapted to complex non-Gaussian [Fe/H] uncertainties like those derived from CaHK photometry. Further work is merited to understand the consistency of these approaches and their relative strengths for different classes of observational data.

4.7 Summary

In this work, we use an analytic one-zone galactic chemical evolution model to fit the CaHK MDF of Eri II in a hierarchical Bayesian framework that appropriately accounts for non-Gaussian measurement uncertainties. Our Fiducial model achieves a good match to the observed MDF from which we infer reasonable constraints on the SFH of Eri II ($\tau_{\rm SFH} = 0.39\pm^{0.18}_{0.13}\,{\rm Gyr}$), SFE ($\tau_{\rm SFE} = 27.56\pm^{25.14}_{12.92}\,{\rm Gyr}$), and mass-loading factor ($\eta = 194.53\pm^{33.37}_{42.67}$). These results are consistent with expectations of both low SFE and high η in low-mass galaxies and with direct estimates of the Eri II SFH from deep photometric data.

Our best-fit Fiducial model paints the following picture of the evolution of Eri II. When star formation began, Eri II had an initial gas mass of $\sim 10^7\,{\rm M}_\odot$ and continued to accrete gas vigorously but at an exponentially declining rate. Because of its low SFE and the presence of strong stellar feedback which drives ISM gas out of its shallow potential well, only a small fraction of the accreted gas is converted into stars. The production of Fe is dominated at early times by CCSNe and at late times by SN Ia, though feedback-induced galactic winds remove >90% of all Fe from the galaxy, resulting in the low final metallicity of Eri II. Gas loss from these large outflows outpaces gas accretion, resulting in an exponentially declining SFH that truncates at $\sim$1.4 Gyr—likely as a result of reionization evaporating its remaining gas supply.

In addition to our Fiducial model, we consider several alternative models to build physical intuition and test specific formation scenarios (e.g., a constant SFR). These models, by-and-large, yield less natural fits to the Eri II MDF compared to the Fiducial model, though in some cases they remain statistically acceptable because of the uncertainties of the stellar [Fe/H] measurements. Additional investigation is required to evaluate the role that spatial variation and stochastic SN enrichment may play in the Eri II MDF. Similarly, given its low mass and early star formation, a more physically-motivated treatment of reionization is warranted. It is encouraging that the stellar MDF alone gives informative constraints on the evolution of Eri II within the framework of the Fiducial model.

Regarding future observations of Eri II and other UFDs, we stress the importance of acquiring precise spectroscopic abundances of not just the lowest metallicity stars, but also—and especially—stars at the high-metallicity end of the MDF. The metallicities and element abundance ratios of these stars will provide some of the strongest constraints on the inferred evolution of their host galaxies.

4.8 Appendices

4.8.1 Appendix: Isolating Parameter Influences on the Model MDF

The ability of the Eri II MDF to constrain the model parameters can be understood by investigating how each parameter changes the predicted model MDF. In Figure 4.11, we

show how the MDF changes as we increase and decrease each parameter from the best-fit Fiducial model. With all other parameters held fixed, increasing $\log_{10}\tau_{\rm SFE}$ results in a higher CC equilibrium abundance $[{\rm Fe/H}]^{\rm cc}_{\rm eq}$ and a longer timescale CC equilibrium timescale $\tau^{\rm cc}_{\rm Fe,eq}$ (Equations 4.10–4.12). This leads to a broader MDF with a more extended low-metallicity tail (top-left panel). In the SFE regime of Eri II, changing $\log_{10}\tau_{\rm SFE}$ does not strongly change the location of the MDFs peak. Increasing $\tau_{\rm SFH}$ has roughly the opposite effect of $\log_{10}\tau_{\rm SFE}$, decreasing both $[{\rm Fe/H}]^{\rm cc}_{\rm eq}$ and $\tau^{\rm cc}_{\rm Fe,eq}$ and resulting in a narrower MDF with a smaller low-metallicity tail (top-right panel). While the similarities in impact between $\tau_{\rm SFH}$ and $\tau_{\rm SFE}$ lead to the covariance in their posteriors seen in Figure 4.3, they are not fully degenerate. Unlike for $\tau_{\rm SFE}$, decreasing $\tau_{\rm SFH}$ shifts the peak of the MDF to lower metallicity. Furthermore, a more extended SFH leads to an MDF that is more sharply truncated at the high-metallicity end by the abrupt end to star formation at $t_{\rm trunc}$. In comparison, the effect of η on the shape of the MDF is more distinct (bottom-right panel). Increasing the mass-loading factor removes more metals from the galaxy, slowing the rate of enrichment and decreasing the final metallicity that the system evolves to. Higher outflows result in a narrower MDF with a lower metallicity peak and maximum [Fe/H]. Changing η, however, has little impact on the low-metallicity tail of the MDF ($[{\rm Fe/H}] < -3.5$). The direct impact of $t_{\rm trunc}$ itself is more subtle than the aforementioned parameters, because it is only responsible for the truncation of the MDF at higher metallicities (and the induced re-normalization; bottom-left panel). In our models of Eri II, $\tau_{\rm SFH}$ is sufficiently short that the SFR is quite low at the time of truncation, and thus the portion of the MDF that is truncated is small. It is therefore understandable that $t_{\rm trunc}$ is prior dominated in our fits; there is no clear signature in the Eri II MDF indicating the abrupt cessation of star formation.

In Figure 4.12, we illustrate how varying a handful of model parameters that were held fixed in the Fiducial model, including $f_{\rm ret}$, t_D, and $y^{\rm Ia}_{\rm Fe}$, influences the predicted MDF in Figure 4.12. Reducing $f_{\rm ret}$ effectively reduces the yield of all SNe, which shifts the entire MDF to lower metallicities (left panel). A factor of two reduction as shown here results in a 0.3 dex shift to lower metallicity. When t_D is increased, SN Ia contribute less to the enrichment of the system overall and especially at early times, resulting in an MDF with a lower-metallicity peak (middle panel). The shape of the MDF below $[{\rm Fe/H}] < -2.5$ when $t_D = 0.15$ Gyr is due to the model approaching the CC equilibrium metallicity, which it does on a timescale roughly equivalent to the minimum SN Ia time delay ($\tau^{\rm cc}_{\rm Fe,eq} \sim t_D$). Increasing $y^{\rm Ia}_{\rm Fe}$ has much the same effect as decreasing η. Large SN Ia yields drives the system to higher metallicity for $t > t_D$, resulting in broader, more metal-rich MDFs. This degeneracy with η explains why the Enhanced SN Ia model required such large mass-loading factors to reproduce the Eri II MDF.

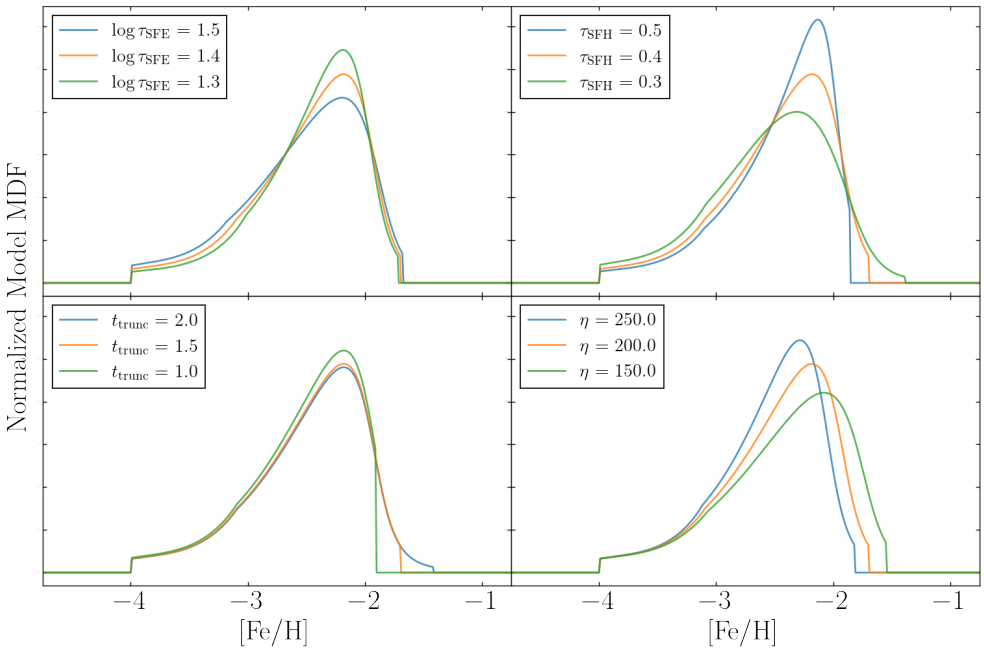


Figure 4.11: Predicted MDF of the Fiducial model as each of the free model parameters is individually increased (blue) and decreased (green) from the approximate best-fit value (orange).

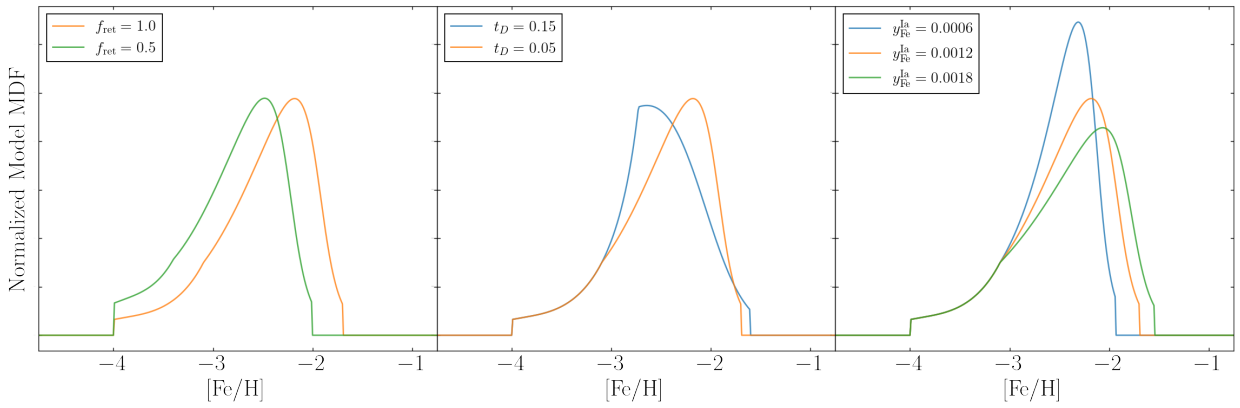


Figure 4.12: Predicted MDF of the best-fit Fiducial model (orange) compared to the predicted MDF when different values of $f_{\rm ret}$, t_D, and $y_{\rm Fe}^{\rm Ia}$ are adopted.

Chapter 5

Conclusions & Future Work

This dissertation has focused on enabling robust chemical abundance measurements from low-resolution spectra of resolved stars and using those measurements to investigate the evolution of the Local Group's low-mass galaxies. In Chapters 2 and 3, I worked at the technical frontier of stellar spectroscopy, using an array of statistical and machine learning techniques along with stellar atmosphere and radiative transfer models to identify and push the limits of what chemical information can be confidently extracted from the next-generation of low-resolution resolved star spectroscopy. In Chapter 4, I utilized chemical evolution models within a probabilistic framework to constrain the reionization-era evolution of the ancient ultra-faint dwarf galaxy, Eridanus II, from its stellar chemistry alone. Below, I briefly summarize the primary results and outstanding science questions presented in each chapter. I close with a short summary of the dissertation's overarching conclusions and the work that remains to be undertaken.

In Chapter 2, I expanded on the application of Cramér-Rao Lower Bounds (CRLBs) to forecast the precision of spectroscopic chemical abundance measurements, incorporating realistic observing conditions, instrument response functions, and Bayesian Priors. This methodology was then employed to quantify the expected chemical abundance precision for >40 existing, future, and proposed spectrograph configurations on 14 telescopes, providing a comprehensive preview of the next decade of Local Group stellar spectroscopy. Highlights of this work include the following forecasts: 1) despite the presence of heavily blended absorption features, low-resolution ($R \lesssim 5,000$) spectra with information-rich blue-optical wavelength coverage ($\lambda < 5000$ Å), can enable precision multi-element abundance measurements, 2) high-resolution ($R \gtrsim 20,000$) contain substantial chemical information even at low S/N ($\sim$10 pixel^{-1}), accessible via full-spectrum fitting techniques, and 3) next-generation low-resolution spectrographs on JWST and 30-m class telescopes will enable, in modest integration time, measurements of bulk metallicity and α-enhancements for individual stars in galaxies throughout the Local Group and out to several Mpc. Preliminary investigation of Keck/DEIMOS and LAMOST spectroscopy indicate general agreement between the forecasted and achieved precision, though further validation of the CRLBs is necessary to

identify potential shortcomings.

As part of the work presented in Chapter 2, I released an open-source python package, `Chem-I-Calc: the Chemical Information Calculator`, the facilitates the calculation of these Fisher forecasts for the astronomical community. Already, `Chem-I-Calc` has been used to plan numerous spectroscopic observations, including a JWST Cycle 2 program to observe isolated LG dwarf galaxies (JWST-GO-3788; PI D. Weisz). `Chem-I-Calc` has also been used to inform the science requirements and science cases of several next-generation spectrographs, including Keck/FOBOS (Bundy et al. 2019), VLT/HRMOS, and MUST[1].

In Chapter 3, I introduced a novel method to validate the forecasted chemical abundance precisions presented in Chapter 2 and assess the impact of the synthetic gap at the low resolutions characteristic of extragalactic stellar spectroscopy. Systematic biases and uncertainties in chemical abundance recovery as a function of decreasing spectroscopic resolution were quantified by analyzing initially exquisite Keck/HIRES spectra of stars in the M15 globular cluster that were incrementally degraded in resolution from $R \sim 80,000$ down to $R \sim 2500$. In order to perform this analysis, which required self-consistent measurement of $\sim$40 elemental abundances and $\sim$100–200 nuisance parameters at spectral resolving powers spanning an order of magnitude, I designed and implemented a probabilistic full-spectrum fitting framework for echelle spectra. This study found that 20 (9) elements can be robustly recovered at resolutions of $R \sim 10,000$ (2500)—a promising result for upcoming extragalactic spectroscopic surveys. As expected, the predominant source of systematic bias and uncertainty found in this study is the blending of poorly modeled absorption features, which most greatly impacts elements with few and/or weak lines (e.g., Cu, Er). This analysis provides further validation that CRLBs are robust estimators of expected measurement precision for most elemental abundances. It also confirms that most chemical abundance measurements are robust at low S/N ($\lesssim$10 pixel^{-1}), though the explanation for the rare exceptions to this requires deeper investigation.

Additional application of the methodology presented in Chapter 3, however, must be undertaken to extend these conclusions more broadly to a more diverse set of stellar types and metallicities and to a wider range of spectroscopic wavelength coverage. This chapter concludes optimistically that even with imperfect models, consistent results can be achieved with both low- and high-resolutions for a representative sample of elements.

In Chapter 4, I presented novel constraints on the underlying galaxy formation physics (e.g., star formation, gas outflows) at $z \gtrsim 6$ for the UFD Eri II. Using a hierarchical Bayesian framework, I fit Eri II's CaHK-based stellar metallicity distribution function with an analytic chemical evolution model and found that Eri II formed the majority of its stars over a short ($\sim$400 Myr) and inefficient (SFE $\sim$ 0.4 Gyr^{-1}) period of star formation during which the galaxy experienced incredibly large supernova-driven gas outflows ($\eta \sim 200$). It is these combination of factors, but especially the efficient outflows, which remove >90% of all Fe

[1] `https://must.astro.tsinghua.edu.cn/en`

from the galaxy, that result in Eri II's low final metallicity. The star formation efficiency and outflows inferred in Eri II are consistent with the current understanding of the field that low-mass galaxies struggle to both convert their gas content into stars and retain their gas reservoirs.

In addition, this chapter considers alternative formation scenarios, which largely yield less natural fits to the chemistry of Eri II's stellar population. Several of these alternative models, however, remain statistically acceptable due to the large uncertainties of the stellar [Fe/H] measurements. A larger sample of stars, more precise [Fe/H] measurements, and/or additional elemental abundances (especially at the more metal-rich end of Eri II's stellar distribution) are necessary to confirm or rule out these models. With JWST/NIRSpec operational and 30-m class telescopes around the corner, these measurements are now, or soon will be, within reach.

As a whole, this dissertation has laid necessary groundwork for the next decade of low-resolution spectroscopic surveys, which will acquire millions of low-resolution spectra of resolved stars though which we can investigate the evolution of the smallest, faintest, and farthest galaxies in the Local Group. It has built upon a long and storied legacy of advancement in the fields of stellar spectroscopy and galactic chemical evolution, combining our present understanding of stellar structure, star formation, galaxy-scale baryonic physics, and nucleosynthesis with powerful statistical techniques to advance our understanding of the Universe.

Much work, of course, remains to be done. While Chapters 2 and 3 demonstrate the power and promise of low-resolution spectrographs for extragalactic stellar spectroscopy, they also highlight the critical work necessary to fully deliver on this potential. Continued investments in updating stellar models with more accurate physical treatments (e.g., 3D-NLTE) and atomic and molecular line data will be crucial to enable robust interpretation of stellar spectra at all resolutions. This is especially important for low-metallicity red giant branch stars, which are the primary targets of extragalactic observations but have seen less attention than the solar-metallicity main sequence stars that dominate Milky Way surveys. In parallel, the refinement of novel spectroscopic techniques like the domain adaptation used by `CYCLE-STARNET` will provide further mitigation of the synthetic gap. Even in advance of the highly multiplexed extragalactic stellar surveys of the next decade, development and application of these methods to archival extragalactic resolved star spectroscopy are likely to yield substantial improvements and increased homogeneity of results over previous analyses.

Chapter 4 provides the first detailed investigation of Eri II's chemical evolution around the time of reionization, which merely scratches the surface of what we can hope to learn about the early evolution of Eri II and other ancient low-mass galaxies in the Local Group. As demonstrated in Chapter 4, substantial degeneracies remain in the specifics of Eri II's evolution (e.g., the initial gas reservoir mass, the metal-loading factor, the minimum time delay of Type Ia supernovae), which can only be broken with additional stellar chemical abundance measurements (e.g., [α/Fe]). Follow-up observations of Eri II should be pursued

with JWST/NIRSpec, which is the only currently operational spectroscopic facility with the angular resolution and sensitivity necessary to make these measurements at the distance of Eri II.

Eri II, of course, is just a singular galaxy, and the specifics of its evolution must be be placed in context with the larger population of low-mass galaxies in the LG. Application of the framework developed in Chapter 4 to the remaining sample of UFDs with CaHK abundances measured by Fu, Weisz, Starkenburg, Martin, Savino, et al. (2023) is already underway. Once complete, this analysis will provide constraints on the star formation efficiencies, star formation timescales, and outflow strengths for the largest sample of LG UFDs to date and enable a study of how these properties vary with stellar mass. In parallel, improvements to the chemical evolution framework should be pursued. Specifically, the current chemical evolution model is lacking in a physically motivated treatment of cosmic reionization and stochastic enrichment events, both of which will grow in importance for the lowest mass galaxies.

I conclude, at last, with one final reiteration of a sentiment that has been repeated countless times throughout this dissertation. In the next 5–10 years, we will see an enormous proliferation of low-resolution resolved star spectroscopy from existing and future facilities, which will extend the reach of stellar archaeology out to the edges of the Local Group and beyond. Though large technical challenges await the field, so too do numerous new avenues for understanding the astrophysics of low-mass galaxy evolution.

Bibliography

Adshead, Peter and Richard Easther (Oct. 2008). "Constraining Inflation". In: *Journal of Cosmology and Astroparticle Physics* 10, p. 047. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2008/10/047. (Visited on 12/30/2019).

Aghanim, N. et al. (Sept. 2020). "Planck 2018 Results. VI. Cosmological Parameters". In: *Astronomy and Astrophysics* 641, A6. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201833910. (Visited on 10/24/2022).

Albrecht, Andreas et al. (Sept. 2006). "Report of the Dark Energy Task Force". In: *arXiv:astro-ph/0609591*. arXiv: astro-ph/0609591. (Visited on 09/25/2019).

Aldenius, M. et al. (Jan. 2007). "Experimental Mg I Oscillator Strengths and Radiative Lifetimes for Astrophysical Applications on Metal-Poor Stars. New Data for the Mg I b Triplet". In: *Astronomy and Astrophysics* 461, pp. 767–773. ISSN: 0004-6361. DOI: 10.1051/0004-6361:20066266. (Visited on 01/03/2020).

Alexander, R. K. et al. (July 2023). "Inhomogeneous Galactic Chemical Evolution: Modelling Ultra-Faint Dwarf Galaxies of the Large Magellanic Cloud". In: *Monthly Notices of the Royal Astronomical Society* 522, pp. 5415–5433. ISSN: 0035-8711. DOI: 10.1093/mnras/stad1312. (Visited on 06/06/2023).

Alfaro-Cuello, Mayte et al. (Sept. 2019). "A Deep View into the Nucleus of the Sagittarius Dwarf Spheroidal Galaxy with MUSE. I. Data and Stellar Population Characterization". In: *arXiv e-prints*, arXiv:1909.10529. (Visited on 10/08/2019).

Allende Prieto, Carlos (July 2016). "Solar and Stellar Photospheric Abundances". In: *Living Reviews in Solar Physics* 13, p. 1. DOI: 10.1007/s41116-016-0001-6. (Visited on 12/17/2019).

Aller, Lawrence H. (Nov. 1942). "Curves of Growth for the a Dwarfs γ Geminorum and Sirius." In: *The Astrophysical Journal* 96, p. 321. ISSN: 0004-637X. DOI: 10.1086/144468. (Visited on 01/21/2020).

— (Nov. 1946). "The Atmosphere of 10 Lacertae." In: *The Astrophysical Journal* 104, p. 347. ISSN: 0004-637X. DOI: 10.1086/144864. (Visited on 01/21/2020).

Alves-Brito, A. et al. (Dec. 2006). "VLT-UVES Abundance Analysis of Four Giants in NGC 6553". In: *Astronomy and Astrophysics* 460, pp. 269–276. ISSN: 0004-6361. DOI: 10.1051/0004-6361:20065488. (Visited on 05/04/2020).

Amarsi, A. M., M. Asplund, et al. (Nov. 2015). "The Galactic Chemical Evolution of Oxygen Inferred from 3D Non-LTE Spectral-Line-Formation Calculations." In: *Monthly Notices*

of the Royal Astronomical Society 454, pp. L11–L15. ISSN: 0035-8711. DOI: 10.1093/mnrasl/slv122. (Visited on 07/17/2023).

Amarsi, A. M., K. Lind, et al. (Dec. 2016). "Non-LTE Line Formation of Fe in Late-Type Stars - III. 3D Non-LTE Analysis of Metal-Poor Stars". In: *Monthly Notices of the Royal Astronomical Society* 463, pp. 1518–1533. ISSN: 0035-8711. DOI: 10.1093/mnras/stw2077. (Visited on 01/03/2020).

Andreasen, D. T. et al. (Jan. 2016). "Near-Infrared Spectroscopy of the Sun and HD 20010. Compiling a New Line List in the near-Infrared". In: *Astronomy and Astrophysics* 585, A143. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201527308. (Visited on 01/03/2020).

Andrews, Brett H. et al. (Feb. 2017). "Inflow, Outflow, Yields, and Stellar Population Mixing in Chemical Evolution Models". In: *The Astrophysical Journal* 835, p. 224. ISSN: 0004-637X. DOI: 10.3847/1538-4357/835/2/224. (Visited on 03/09/2023).

Aoki, W. et al. (Aug. 2009). "Chemical Composition of Extremely Metal-Poor Stars in the Sextans Dwarf Spheroidal Galaxy". In: *Astronomy and Astrophysics* 502, pp. 569–578. ISSN: 0004-6361. DOI: 10.1051/0004-6361/200911959. (Visited on 01/21/2020).

Arcones, Almudena and Friedrich-Karl Thielemann (Dec. 2023). "Origin of the Elements". In: *Astronomy and Astrophysics Review* 31, p. 1. ISSN: 0935-4956. DOI: 10.1007/s00159-022-00146-x. (Visited on 07/25/2023).

Arentsen, Anke et al. (Sept. 2022). "On the Inconsistency of [C/Fe] Abundances and the Fractions of Carbon-Enhanced Metal-Poor Stars among Various Stellar Surveys". In: *Monthly Notices of the Royal Astronomical Society* 515, pp. 4082–4098. ISSN: 0035-8711. DOI: 10.1093/mnras/stac2062. (Visited on 09/14/2022).

Arp, H. C. (Jan. 1955). "Color-Magnitude Diagrams for Seven Globular Clusters." In: *The Astronomical Journal* 60, p. 317. ISSN: 0004-6256. DOI: 10.1086/107232. (Visited on 06/29/2021).

Asplund, Martin (2005). "New Light on Stellar Abundance Analyses: Departures from LTE and Homogeneity". In: *Annual Review of Astronomy and Astrophysics* 43.1, pp. 481–530. DOI: 10.1146/annurev.astro.42.053102.134001. (Visited on 04/19/2023).

Asplund, Martin et al. (Sept. 2009). "The Chemical Composition of the Sun". In: *Annual Review of Astronomy and Astrophysics* 47, p. 481. DOI: 10.1146/annurev.astro.46.060407.145222. (Visited on 10/08/2019).

Bacon, R. et al. (July 2010). "The MUSE Second-Generation VLT Instrument". In: *Ground-Based and Airborne Instrumentation for Astronomy III*. Vol. 7735. International Society for Optics and Photonics, p. 773508. DOI: 10.1117/12.856027. (Visited on 10/07/2019).

El-Badry, Kareem, Hans-Walter Rix, et al. (Feb. 2018). "Signatures of Unresolved Binaries in Stellar Spectra: Implications for Spectral Fitting". In: *Monthly Notices of the Royal Astronomical Society* 473, pp. 5043–5049. ISSN: 0035-8711. DOI: 10.1093/mnras/stx2758. (Visited on 07/23/2021).

El-Badry, Kareem, Andrew Wetzel, et al. (Apr. 2016). "Breathing FIRE: How Stellar Feedback Drives Radial Migration, Rapid Size Fluctuations, and Population Gradients in Low-mass Galaxies". In: *The Astrophysical Journal* 820, p. 131. ISSN: 0004-637X. DOI: 10.3847/0004-637X/820/2/131. (Visited on 06/07/2023).

Bagnasco, Giorgio et al. (Sept. 2007). "Overview of the Near-Infrared Spectrograph (NIRSpec) Instrument on-Board the James Webb Space Telescope (JWST)". In: *Cryogenic Optical Systems and Instruments XII* 6692, p. 66920M. ISSN: 0277-786X. DOI: 10.1117/12.735602. (Visited on 11/21/2019).

Bailer-Jones, C. a. L. (May 2000). "Stellar Parameters from Very Low Resolution Spectra and Medium Band Filters. T_eff, Log g and [M/H] Using Neural Networks". In: *Astronomy and Astrophysics* 357, p. 197. (Visited on 10/04/2019).

Baschek, B. (1959). "Aufbau Und Chemische Zusammensetzung Der Atmosphäre Des Subdwarfs HD 140283. Mit 10 Textabbildungen". In: *Zeitschrift fur Astrophysik* 48, p. 95. ISSN: 0372-8331. (Visited on 01/21/2020).

Bastian, Nate and Carmela Lardo (Sept. 2018). "Multiple Stellar Populations in Globular Clusters". In: *Annual Review of Astronomy and Astrophysics* 56, pp. 83–136. ISSN: 0066-4146. DOI: 10.1146/annurev-astro-081817-051839. (Visited on 12/31/2019).

Battaglia, G., A. Helmi, et al. (July 2008). "The Kinematic Status and Mass Content of the Sculptor Dwarf Spheroidal Galaxy". In: *The Astrophysical Journal Letters* 681, p. L13. ISSN: 0004-637X. DOI: 10.1086/590179. (Visited on 01/21/2020).

Battaglia, G., S. Taibi, et al. (Jan. 2022). "Gaia Early DR3 Systemic Motions of Local Group Dwarf Galaxies and Orbital Properties with a Massive Large Magellanic Cloud". In: *Astronomy and Astrophysics* 657, A54. ISSN: 0004-6361. DOI: 10.1051/0004-6361/202141528. (Visited on 10/17/2022).

Battaglia, G., E. Tolstoy, A. Helmi, M. Irwin, P. Parisi, et al. (Feb. 2011). "Study of the Sextans Dwarf Spheroidal Galaxy from the DART Ca II Triplet Survey". In: *Monthly Notices of the Royal Astronomical Society* 411, pp. 1013–1034. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2010.17745.x. (Visited on 01/24/2020).

Battaglia, G., E. Tolstoy, A. Helmi, M. J. Irwin, B. Letarte, et al. (Nov. 2006). "The DART Imaging and CaT Survey of the Fornax Dwarf Spheroidal Galaxy". In: *Astronomy and Astrophysics* 459, pp. 423–440. ISSN: 0004-6361. DOI: 10.1051/0004-6361:20065720. (Visited on 01/21/2020).

Baumgardt, H. and E. Vasiliev (Aug. 2021). "Accurate Distances to Galactic Globular Clusters through a Combination of Gaia EDR3, HST, and Literature Data". In: *Monthly Notices of the Royal Astronomical Society* 505, pp. 5957–5977. ISSN: 0035-8711. DOI: 10.1093/mnras/stab1474. (Visited on 08/30/2022).

Bechtol, K. et al. (July 2015). "Eight New Milky Way Companions Discovered in First-year Dark Energy Survey Data". In: *The Astrophysical Journal* 807, p. 50. ISSN: 0004-637X. DOI: 10.1088/0004-637X/807/1/50. (Visited on 10/17/2022).

Becker, Adam, Dragan Huterer, and Kenji Kadota (Dec. 2012). "Constraining Scale-Dependent Non-Gaussianity with Future Large-Scale Structure and the CMB". In: *Journal of Cosmology and Astroparticle Physics* 12, p. 034. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2012/12/034. (Visited on 12/30/2019).

Bedell, Megan et al. (Nov. 2014). "Stellar Chemical Abundances: In Pursuit of the Highest Achievable Precision". In: *The Astrophysical Journal* 795, p. 23. ISSN: 0004-637X. DOI: 10.1088/0004-637X/795/1/23. (Visited on 01/27/2020).

Behroozi, Peter S., Risa H. Wechsler, and Charlie Conroy (Jan. 2013). "On the Lack of Evolution in Galaxy Star Formation Efficiency". In: *The Astrophysical Journal* 762, p. L31. ISSN: 0004-637X. DOI: 10.1088/2041-8205/762/2/L31. (Visited on 10/24/2022).

Bell, R. A. (1970). "The Calibration of Narrow Band Photometry - I. Cambridge Observations of G and K Giants." In: *Monthly Notices of the Royal Astronomical Society* 148, pp. 25–52. ISSN: 0035-8711. DOI: 10.1093/mnras/148.1.25. (Visited on 01/21/2020).

Bell, R. A. and D. Branch (Apr. 1976). "Carbon Abundances in G Dwarfs". In: *Monthly Notices of the Royal Astronomical Society* 175, pp. 25–32. ISSN: 0035-8711. DOI: 10.1093/mnras/175.1.25. (Visited on 01/21/2020).

Bensby, T. et al. (Sept. 2017). "Chemical Evolution of the Galactic Bulge as Traced by Microlensed Dwarf and Subgiant Stars. VI. Age and Abundance Structure of the Stellar Populations in the Central Sub-Kpc of the Milky Way". In: *Astronomy and Astrophysics* 605, A89. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201730560. (Visited on 07/25/2023).

Benson, A. J. et al. (Dec. 2003). "What Shapes the Luminosity Function of Galaxies?" In: *The Astrophysical Journal* 599, pp. 38–49. ISSN: 0004-637X. DOI: 10.1086/379160. (Visited on 10/19/2022).

Bergemann, M. and G. Cescutti (Nov. 2010). "Chromium: NLTE Abundances in Metal-Poor Stars and Nucleosynthesis in the Galaxy". In: *A&A* 522, A9. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201014250. (Visited on 03/02/2023).

Bergemann, M. and T. Gehren (Dec. 2008). "NLTE Abundances of Mn in a Sample of Metal-Poor Stars". In: *Astronomy and Astrophysics, Volume 492, Issue 3, 2008, pp.823-831* 492.3, p. 823. ISSN: 0004-6361. DOI: 10.1051/0004-6361:200810098. (Visited on 07/21/2021).

Bergemann, Maria (May 2011). "Ionization Balance of Ti in the Photospheres of the Sun and Four Late-Type Stars". In: *Monthly Notices of the Royal Astronomical Society* 413, pp. 2184–2198. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2011.18295.x. (Visited on 07/21/2021).

Bergemann, Maria, Remo Collet, et al. (Sept. 2017). "Non-Local Thermodynamic Equilibrium Stellar Spectroscopy with 1D and <3D> Models. I. Methods and Application to Magnesium Abundances in Standard Stars". In: *The Astrophysical Journal* 847, p. 15. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aa88cb. (Visited on 07/21/2021).

Bergemann, Maria, Rolf-Peter Kudritzki, et al. (Feb. 2013). "Red Supergiant Stars as Cosmic Abundance Probes. II. NLTE Effects in J-band Silicon Lines". In: *The Astrophysical Journal* 764, p. 115. ISSN: 0004-637X. DOI: 10.1088/0004-637X/764/2/115. (Visited on 07/21/2021).

Bergemann, Maria, K. Lind, et al. (Nov. 2012). "Non-LTE Line Formation of Fe in Late-Type Stars - I. Standard Stars with 1D and <3D> Model Atmospheres". In: *Monthly Notices of the Royal Astronomical Society* 427, pp. 27–49. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2012.21687.x. (Visited on 01/03/2020).

Bergemann, Maria, Juliet C. Pickering, and Thomas Gehren (Jan. 2010). "NLTE Analysis of CoI/CoII Lines in Spectra of Cool Stars with New Laboratory Hyperfine Splitting

Constants". In: *Monthly Notices of the Royal Astronomical Society* 401, pp. 1334–1346. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2009.15736.x. (Visited on 07/21/2021).

Bernstein, Rebecca et al. (Mar. 2003). "MIKE: A Double-Echelle Spectrograph for the Magellan Telescopes at Las Campanas Observatory". In: *Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*. Vol. 4841. International Society for Optics and Photonics, pp. 1694–1704. DOI: 10.1117/12.461502. (Visited on 10/07/2019).

Betoule, M. et al. (Aug. 2014). "Improved Cosmological Constraints from a Joint Analysis of the SDSS-II and SNLS Supernova Samples". In: *Astronomy and Astrophysics* 568, A22. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201423413. (Visited on 12/30/2019).

Blanco-Cuaresma, Sergi (June 2019). "Modern Stellar Spectroscopy Caveats". In: *Monthly Notices of the Royal Astronomical Society* 486, pp. 2075–2101. ISSN: 0035-8711. DOI: 10.1093/mnras/stz549. (Visited on 12/15/2019).

Boesgaard, Ann Merchant, Jeremy R. King, et al. (Aug. 2005). "Chemical Composition in the Globular Cluster M71 from Keck HIRES Spectra of Turnoff Stars". In: *The Astrophysical Journal* 629, p. 832. ISSN: ['0004-637X']. DOI: 10.1086/431645. (Visited on 11/21/2019).

Boesgaard, Ann Merchant, Alex Stephens, et al. (June 2000). "Chemical Abundances in Globular Cluster Turn-off Stars from Keck/HIRES Observations". In: *Discoveries and Research Prospects from 8- to 10-Meter-Class Telescopes*. Vol. 4005. International Society for Optics and Photonics, pp. 274–284. DOI: 10.1117/12.390137. (Visited on 11/21/2019).

Boylan-Kolchin, Michael et al. (Oct. 2016). "The Local Group: The Ultimate Deep Field". In: *Monthly Notices of the Royal Astronomical Society* 462, pp. L51–L55. ISSN: 0035-8711. DOI: 10.1093/mnrasl/slw121. (Visited on 07/25/2023).

Braak, Ter and Cajo J. F (Sept. 2006). "A Markov Chain Monte Carlo Version of the Genetic Algorithm Differential Evolution: Easy Bayesian Computing for Real Parameter Spaces". In: *Statistics and Computing* 16.3, pp. 239–249. DOI: 10.1007/s11222-006-8769-1. (Visited on 08/30/2022).

Buder, Sven et al. (Sept. 2021). "The GALAH+ Survey: Third Data Release". In: *Monthly Notices of the Royal Astronomical Society* 506, pp. 150–201. ISSN: 0035-8711. DOI: 10.1093/mnras/stab1242. (Visited on 09/14/2022).

Bundy, Kevin et al. (Sept. 2019). "FOBOS: A Next-Generation Spectroscopic Facility". In: *Bulletin of the American Astronomical Society* 51, p. 198. (Visited on 11/21/2019).

Caffau, E., A. Koch, et al. (Mar. 2013). "Velocity and Abundance Precisions for Future High-Resolution Spectroscopic Surveys: A Study for 4MOST". In: *Astronomische Nachrichten* 334, p. 197. ISSN: 0004-6337. DOI: 10.1002/asna.201211814. (Visited on 01/10/2020).

Caffau, E., H.-G. Ludwig, et al. (Feb. 2011). "Solar Chemical Abundances Determined with a CO5BOLD 3D Model Atmosphere". In: *Solar Physics* 268, pp. 255–269. ISSN: 0038-0938. DOI: 10.1007/s11207-010-9541-4. (Visited on 01/03/2020).

Carbon, D. F. et al. (June 1982). "Carbon and Nitrogen Abundances in Giant Stars of the Metal-Poor Globular Cluster M92". In: *The Astrophysical Journal Supplement Series* 49, pp. 207–258. ISSN: 0067-0049. DOI: 10.1086/190796. (Visited on 01/21/2020).

Cargile, Phillip A. et al. (Sept. 2020). "MINESweeper: Spectrophotometric Modeling of Stars in the Gaia Era". In: *The Astrophysical Journal* 900, p. 28. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aba43b. (Visited on 01/27/2023).

Carlberg, Joleen K. et al. (May 2011). "The Frequency of Rapid Rotation Among K Giant Stars". In: *The Astrophysical Journal* 732, p. 39. ISSN: 0004-637X. DOI: 10.1088/0004-637X/732/1/39. (Visited on 08/30/2022).

Carlin, Jeffrey L. et al. (Sept. 2009). "Kinematics and Metallicities in the Boötes III Stellar Overdensity: A Disrupted Dwarf Galaxy?" In: *The Astrophysical Journal* 702.1, p. L9. DOI: 10.1088/0004-637X/702/1/L9. (Visited on 10/07/2019).

Carnall, A. C. (May 2017). "SpectRes: A Fast Spectral Resampling Tool in Python". In: *arXiv e-prints*, arXiv:1705.05165. (Visited on 07/20/2021).

Carrera, R. et al. (Sept. 2013). "The Near-Infrared Ca II Triplet as a Metallicity Indicator - II. Extension to Extremely Metal-Poor Metallicity Regimes". In: *Monthly Notices of the Royal Astronomical Society* 434, pp. 1681–1691. ISSN: 0035-8711. DOI: 10.1093/mnras/stt1126. (Visited on 01/24/2020).

Carretta, E., A. Bragaglia, R. Gratton, and S. Lucatello (Oct. 2009). "Na-O Anticorrelation and HB. VIII. Proton-capture Elements and Metallicities in 17 Globular Clusters from UVES Spectra". In: *Astronomy and Astrophysics* 505.1, pp. 139–155. ISSN: 0004-6361. DOI: 10.1051/0004-6361/200912097. (Visited on 06/29/2021).

Carretta, E., A. Bragaglia, R. G. Gratton, S. Lucatello, et al. (Oct. 2009). "Na-O Anticorrelation and HB. VII. The Chemical Composition of First and Second-Generation Stars in 15 Globular Clusters from GIRAFFE Spectra". In: *Astronomy and Astrophysics* 505.1, pp. 117–138. ISSN: 0004-6361. DOI: 10.1051/0004-6361/200912096. (Visited on 08/30/2022).

Casagrande, L. et al. (June 2011). "New Constraints on the Chemical Evolution of the Solar Neighbourhood and Galactic Disc(s). Improved Astrophysical Parameters for the Geneva-Copenhagen Survey". In: *Astronomy and Astrophysics* 530, A138. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201016276. (Visited on 11/21/2019).

Cayrel, R. (1969). "Comparison of Synthetic Spectra with Real Spectra". In: ed. by Owen Gingerich. Proceedings of the 3rd Harvard-Smithsonian Conference on Stellar Atmospheres. Massachusetts Institute of Technology. Cambridge, MA, p. 237. (Visited on 01/21/2020).

Cayrel, R. et al. (Mar. 2004). "First Stars V - Abundance Patterns from C to Zn and Supernova Yields in the Early Galaxy". In: *A&A* 416.3, pp. 1117–1138. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361:20034074. (Visited on 08/30/2022).

Cenarro, A. J., N. Cardiel, et al. (Sept. 2001). "Empirical Calibration of the Near-Infrared Ca Ii Triplet - I. The Stellar Library and Index Definition". In: *Monthly Notices of the Royal Astronomical Society* 326, pp. 959–980. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2001.04688.x. (Visited on 01/14/2020).

Cenarro, A. J., J. Gorgas, N. Cardiel, S. Pedraz, et al. (Sept. 2001). "Empirical Calibration of the Near-Infrared Ca Ii Triplet - II. The Stellar Atmospheric Parameters". In: *Monthly*

Notices of the Royal Astronomical Society 326, pp. 981–994. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2001.04689.x. (Visited on 01/14/2020).

Cenarro, A. J., J. Gorgas, N. Cardiel, A. Vazdekis, et al. (Feb. 2002). "Empirical Calibration of the Near-Infrared Ca II Triplet - III. Fitting Functions". In: *Monthly Notices of the Royal Astronomical Society* 329, pp. 863–876. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2002.05029.x. (Visited on 01/14/2020).

Chapman, Scott C. et al. (Oct. 2005). "A Keck DEIMOS Kinematic Study of Andromeda IX: Dark Matter on the Smallest Galactic Scales". In: *The Astrophysical Journal* 632, p. L87. DOI: 10.1086/497686. (Visited on 10/07/2019).

Chevalier, R. A. and A. W. Clegg (Sept. 1985). "Wind from a Starburst Galaxy Nucleus". In: *Nature* 317, pp. 44–45. ISSN: 0028-0836. DOI: 10.1038/317044a0. (Visited on 06/07/2023).

Chisholm, John et al. (Aug. 2017). "The Mass and Momentum Outflow Rates of Photoionized Galactic Outflows". In: *Monthly Notices of the Royal Astronomical Society* 469, pp. 4831–4849. ISSN: 0035-8711. DOI: 10.1093/mnras/stx1164. (Visited on 06/07/2023).

Choi, Jieun et al. (June 2016). "Mesa Isochrones and Stellar Tracks (MIST). I. Solar-scaled Models". In: *The Astrophysical Journal* 823, 102, p. 102. DOI: 10.3847/0004-637X/823/2/102. arXiv: 1604.08592 [astro-ph.SR]. (Visited on 10/04/2019).

Cohen, Judith G. and Wenjin Huang (Aug. 2009). "The Chemical Evolution of the Draco Dwarf Spheroidal Galaxy". In: *The Astrophysical Journal* 701, pp. 1053–1075. ISSN: 0004-637X. DOI: 10.1088/0004-637X/701/2/1053. (Visited on 01/21/2020).

Collaboration, Gaia et al. (Aug. 2018). "Gaia Data Release 2 - Observational Hertzsprung-Russell Diagrams". In: *A&A* 616, A10. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201832843. (Visited on 03/02/2023).

Collins, Michelle L. M. et al. (May 2013). "A Kinematic Study of the Andromeda Dwarf Spheroidal System". In: *The Astrophysical Journal* 768.2, p. 172. ISSN: 0004-637X. DOI: 10.1088/0004-637X/768/2/172. (Visited on 11/21/2019).

Conroy, Charlie, Ana Bonaca, et al. (Sept. 2019). "Mapping the Stellar Halo with the H3 Spectroscopic Survey". In: *The Astrophysical Journal* 883, p. 107. ISSN: ['0004-637X']. DOI: 10.3847/1538-4357/ab38b8. (Visited on 11/21/2019).

Conroy, Charlie, Rohan P. Naidu, et al. (Sept. 2019). "Resolving the Metallicity Distribution of the Stellar Halo with the H3 Survey". In: *arXiv e-prints*, arXiv:1909.02007. (Visited on 11/21/2019).

Conroy, Charlie, David H. Weinberg, et al. (Apr. 2022). *Birth of the Galactic Disk Revealed by the H3 Survey*. DOI: 10.48550/arXiv.2204.02989. (Visited on 03/09/2023).

Cooper, Andrew P. et al. (Apr. 2023). "Overview of the DESI Milky Way Survey". In: *The Astrophysical Journal* 947, p. 37. ISSN: 0004-637X. DOI: 10.3847/1538-4357/acb3c0. (Visited on 05/26/2023).

Cowan, John J. et al. (June 2002). "The Chemical Composition and Age of the Metal-poor Halo Star BD +17°3248". In: *The Astrophysical Journal* 572, pp. 861–879. ISSN: 0004-637X. DOI: 10.1086/340347. (Visited on 01/10/2020).

Cramer, Harald (1946). *Mathematical Methods of Statistics*. Princeton Mathematical Series 9. Princeton: Princeton University Press.

Crnojević, D. et al. (June 2016). "Deep Imaging of Eridanus II and Its Lone Star Cluster". In: *The Astrophysical Journal* 824, p. L14. ISSN: 0004-637X. DOI: 10.3847/2041-8205/824/1/L14. (Visited on 01/25/2023).

Cui, Xiang-Qun et al. (Aug. 2012). "The Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST)". In: *Research in Astronomy and Astrophysics* 12.9, pp. 1197–1242. ISSN: 1674-4527. DOI: 10.1088/1674-4527/12/9/003. (Visited on 05/18/2020).

Czekala, Ian et al. (Oct. 2015). "Constructing a Flexible Likelihood Function for Spectroscopic Inference". In: *The Astrophysical Journal* 812.2, p. 128. DOI: 10.1088/0004-637X/812/2/128. (Visited on 09/25/2019).

Dalton, Gavin et al. (Aug. 2016). "Final Design and Progress of WEAVE: The next Generation Wide-Field Spectroscopy Facility for the William Herschel Telescope". In: *Ground-Based and Airborne Instrumentation for Astronomy VI*. Vol. 9908. SPIE, pp. 394–403. DOI: 10.1117/12.2231078. (Visited on 11/12/2022).

Darmois, G. (1945). "Sur Les Limites de La Dispersion de Certaines Estimations". In: *Revue de l'Institut International de Statistique / Review of the International Statistical Institute* 13.1/4, pp. 9–15. ISSN: 0373-1138. DOI: 10.2307/1400974. JSTOR: 1400974. (Visited on 01/10/2020).

Davé, Romeel et al. (Sept. 2013). "The Neutral Hydrogen Content of Galaxies in Cosmological Hydrodynamic Simulations". In: *Monthly Notices of the Royal Astronomical Society* 434, pp. 2645–2663. ISSN: 0035-8711. DOI: 10.1093/mnras/stt1274. (Visited on 06/07/2023).

de Jong, R. S. et al. (Mar. 2019). "4MOST: Project Overview and Information for the First Call for Proposals". In: *The Messenger* 175, pp. 3–11. ISSN: 0722-6691. DOI: 10.18727/0722-6691/5117. (Visited on 09/13/2022).

De Silva, G. M. et al. (May 2015). "The GALAH Survey: Scientific Motivation". In: *Monthly Notices of the Royal Astronomical Society* 449.3, p. 2604. ISSN: 0035-8711. DOI: 10.1093/mnras/stv327. (Visited on 11/21/2019).

Dekker, Hans et al. (Aug. 2000). "Design, Construction, and Performance of UVES, the Echelle Spectrograph for the UT2 Kueyen Telescope at the ESO Paranal Observatory". In: 4008, pp. 534–545. DOI: 10.1117/12.395512. (Visited on 05/04/2020).

Den Hartog, E. A. et al. (Aug. 2019). "Atomic Transition Probabilities for UV and Blue Lines of Fe II and Abundance Determinations in the Photospheres of the Sun and Metal-poor Star HD 84937". In: *The Astrophysical Journal Supplement Series* 243, p. 33. ISSN: 0067-0049. DOI: 10.3847/1538-4365/ab322e. (Visited on 09/13/2022).

DePoy, D. L. et al. (Sept. 2012). "GMACS: A Wide Field, Multi-Object, Moderate-Resolution, Optical Spectrograph for the Giant Magellan Telescope". In: *Ground-Based and Airborne Instrumentation for Astronomy IV*. Vol. 8446. International Society for Optics and Photonics, 84461N. DOI: 10.1117/12.926186. (Visited on 11/21/2019).

DESI Collaboration, Aghamousa, et al. (Oct. 2016). "The DESI Experiment Part II: Instrument Design". In: *arXiv e-prints*, arXiv:1611.00037. (Visited on 10/07/2019).

DESI Collaboration, Jessica Aguilar, et al. (Oct. 2016). "The DESI Experiment Part I: Science,Targeting, and Survey Design". In: *arXiv e-prints*, arXiv:1611.00036. (Visited on 10/07/2019).

Dotter, Aaron (Jan. 2016). "MESA Isochrones and Stellar Tracks (MIST) 0: Methods for the Construction of Stellar Isochrones". In: *The Astrophysical Journal Supplement Series* 222, p. 8. DOI: 10.3847/0067-0049/222/1/8. (Visited on 10/04/2019).

Duane, Simon et al. (Sept. 1987). "Hybrid Monte Carlo". In: *Physics Letters B* 195.2, pp. 216–222. ISSN: 0370-2693. DOI: 10.1016/0370-2693(87)91197-X. (Visited on 12/30/2019).

Duggan, Gina E. et al. (Dec. 2018). "Neutron Star Mergers Are the Dominant Source of the R-Process in the Early Evolution of Dwarf Galaxies". In: *The Astrophysical Journal* 869.1, p. 50. DOI: 10.3847/1538-4357/aaeb8e. (Visited on 10/07/2019).

Echeverria, Alex et al. (Oct. 2016). "Analysis of the Bayesian Cramér-Rao Lower Bound in Astrometry. Studying the Impact of Prior Information in the Location of an Object". In: *Astronomy and Astrophysics* 594, A111. DOI: 10.1051/0004-6361/201628220. (Visited on 09/25/2019).

Eggen, N. R. et al. (July 2022). *Blow-Away in the Extreme Low-Mass Starburst Galaxy Pox˜186*. DOI: 10.48550/arXiv.2207.02245. arXiv: 2207.02245 [astro-ph]. (Visited on 06/07/2023).

Eitner, P., M. Bergemann, and S. Larsen (July 2019). "NLTE Modelling of Integrated Light Spectra - Abundances of Barium, Magnesium, and Manganese in a Metal-Poor Globular Cluster". In: *A&A* 627, A40. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201935416. (Visited on 03/02/2023).

Emerick, Andrew, Greg L. Bryan, and Mordecai-Mark Mac Low (Jan. 2019). "Simulating an Isolated Dwarf Galaxy with Multichannel Feedback and Chemical Yields from Individual Stars". In: *Monthly Notices of the Royal Astronomical Society* 482, pp. 1304–1329. ISSN: 0035-8711. DOI: 10.1093/mnras/sty2689. (Visited on 06/07/2023).

Eriksen, Martin and Enrique Gaztañaga (Sept. 2015). "Combining Spectroscopic and Photometric Surveys Using Angular Cross-Correlations - II. Parameter Constraints from Different Physical Effects". In: *Monthly Notices of the Royal Astronomical Society* 452, pp. 2168–2184. ISSN: 0035-8711. DOI: 10.1093/mnras/stv1075. (Visited on 12/30/2019).

Escala, Ivanna, Karoline M. Gilbert, et al. (Feb. 2020). "Elemental Abundances in M31: A Comparative Analysis of Alpha and Iron Element Abundances in the the Outer Disk, Giant Stellar Stream, and Inner Halo of M31". In: *The Astrophysical Journal* 889, p. 177. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab6659. (Visited on 02/13/2023).

Escala, Ivanna, Evan N. Kirby, et al. (June 2019). "Elemental Abundances in M31: Alpha and Iron Element Abundances from Low-resolution Resolved Stellar Spectroscopy in the Stellar Halo". In: *The Astrophysical Journal* 878.1, p. 42. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab1eac. (Visited on 11/22/2019).

Escala, Ivanna, Andrew Wetzel, et al. (Feb. 2018). "Modelling Chemical Abundance Distributions for Dwarf Galaxies in the Local Group: The Impact of Turbulent Metal Diffu-

sion". In: *Monthly Notices of the Royal Astronomical Society* 474, pp. 2194–2211. ISSN: 0035-8711. DOI: 10.1093/mnras/stx2858. (Visited on 01/26/2023).

Evans, C. J. et al. (Feb. 2019). "First Stellar Spectroscopy in Leo P". In: *Astronomy and Astrophysics* 622, A129. DOI: 10.1051/0004-6361/201834145. (Visited on 10/08/2019).

Faber, Sandra M. et al. (Mar. 2003). "The DEIMOS Spectrograph for the Keck II Telescope: Integration and Testing". In: *Instrument Design and Performance for Optical/Infrared Ground-based Telescopes* 4841, p. 1657. DOI: 10.1117/12.460346. (Visited on 10/04/2019).

Fabricant, Daniel, Robert Fata, Harland Epps, et al. (July 2019). "Binospec: A Wide-field Imaging Spectrograph for the MMT". In: *Publications of the Astronomical Society of the Pacific* 131.1001, p. 075004. DOI: 10.1088/1538-3873/ab1d78. (Visited on 10/07/2019).

Fabricant, Daniel, Robert Fata, John Roll, et al. (Dec. 2005). "Hectospec, the MMT's 300 Optical Fiber-Fed Spectrograph". In: *Publications of the Astronomical Society of the Pacific* 117.838, p. 1411. DOI: 10.1086/497385. (Visited on 10/07/2019).

Fanson, James et al. (Dec. 2020). "Overview and Status of the Giant Magellan Telescope Project". In: *Ground-Based and Airborne Telescopes VIII*. Vol. 11445. SPIE, pp. 295–314. DOI: 10.1117/12.2561852. (Visited on 11/12/2022).

Feeney, Stephen M., Benjamin D. Wandelt, and Melissa K. Ness (Dec. 2019). "SSSpaNG! Stellar Spectra as Sparse, Data-Driven, Non-Gaussian Processes". In: *arXiv e-prints* 1912, arXiv:1912.09498. (Visited on 01/09/2020).

Finlator, Kristian and Romeel Davé (Apr. 2008). "The Origin of the Galaxy Mass-Metallicity Relation and Implications for Galactic Outflows". In: *Monthly Notices of the Royal Astronomical Society* 385, pp. 2181–2204. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2008.12991.x. (Visited on 10/19/2022).

Fischel, David (July 1964). "On the Influence of the Chemical Composition on the Structure and Metal Index of Stellar Atmospheres." In: *The Astrophysical Journal* 140, p. 221. ISSN: 0004-637X. DOI: 10.1086/147909. (Visited on 01/21/2020).

Font-Ribera, Andreu et al. (May 2014). "DESI and Other Dark Energy Experiments in the Era of Neutrino Mass Measurements". In: *Journal of Cosmology and Astroparticle Physics* 05, p. 023. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2014/05/023. (Visited on 12/30/2019).

Fraunhofer, Joseph (1817). "Bestimmung Des Brechungs- Und Des Farbenzerstreungs-Vermögens Verschiedener Glasarten, in Bezug Auf Die Vervollkommnung Achromatischer Fernröhre". In: *Annalen der Physik* 56, pp. 264–313. ISSN: 0003-3804. DOI: 10.1002/andp.18170560706. (Visited on 01/21/2020).

Frebel, Anna and John E. Norris (Aug. 2015). "Near-Field Cosmology with Extremely Metal-Poor Stars". In: *Annual Review of Astronomy and Astrophysics* 53, pp. 631–688. ISSN: 0066-4146. DOI: 10.1146/annurev-astro-082214-122423. (Visited on 07/25/2023).

Frebel, Anna, John E. Norris, et al. (Aug. 2016). "The Chemical Evolution of the Bootes I Ultra-faint Dwarf Galaxy". In: *The Astrophysical Journal* 826.2, p. 110. DOI: 10.3847/0004-637X/826/2/110. (Visited on 10/07/2019).

Frebel, Anna, Joshua D. Simon, Marla Geha, et al. (Jan. 2010). "High-Resolution Spectroscopy of Extremely Metal-Poor Stars in the Least Evolved Galaxies: Ursa Major II and Coma Berenices". In: *The Astrophysical Journal* 708, pp. 560–583. ISSN: 0004-637X. DOI: 10.1088/0004-637X/708/1/560. (Visited on 01/21/2020).

Frebel, Anna, Joshua D. Simon, and Evan N. Kirby (May 2014). "Segue 1: An Unevolved Fossil Galaxy from the Early Universe". In: *The Astrophysical Journal* 786.1, p. 74. DOI: 10.1088/0004-637X/786/1/74. (Visited on 10/07/2019).

Fréchet, Maurice (1943). "Sur l'extension de Certaines Evaluations Statistiques Au Cas de Petits Echantillons". In: *Revue de l'Institut International de Statistique / Review of the International Statistical Institute* 11.3/4, pp. 182–205. ISSN: 0373-1138. DOI: 10.2307/1401114. JSTOR: 1401114. (Visited on 01/10/2020).

Fu, Sal Wanying, Daniel R. Weisz, Else Starkenburg, Nicolas Martin, Alexander P. Ji, et al. (Jan. 2022). "Metallicity Distribution Function of the Eridanus II Ultra-faint Dwarf Galaxy from Hubble Space Telescope Narrowband Imaging". In: *The Astrophysical Journal* 925, p. 6. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ac3665. (Visited on 08/15/2022).

Fu, Sal Wanying, Daniel R. Weisz, Else Starkenburg, Nicolas Martin, Alessandro Savino, et al. (June 2023). *Metallicity Distribution Functions of 13 Ultra-Faint Dwarf Galaxy Candidates from Hubble Space Telescope Narrowband Imaging*. DOI: 10.48550/arXiv.2306.06260. (Visited on 07/18/2023).

Fulbright, Jon. P., R. Michael Rich, and Sandra Castro (Sept. 2004). "Draco 119: A Remarkable Heavy-Element-deficient Giant". In: *The Astrophysical Journal* 612, pp. 447–453. ISSN: 0004-637X. DOI: 10.1086/421712. (Visited on 01/21/2020).

Gaia Collaboration et al. (July 2022). *Gaia Data Release 3: Summary of the Content and Survey Properties*. (Visited on 08/30/2022).

Gallart, C. et al. (Mar. 2021). "The Star Formation History of Eridanus II: On the Role of Supernova Feedback in the Quenching of Ultrafaint Dwarf Galaxies". In: *The Astrophysical Journal* 909, p. 192. ISSN: 0004-637X. DOI: 10.3847/1538-4357/abddbe. (Visited on 08/24/2022).

García Pérez, Ana E. et al. (June 2016). "ASPCAP: The APOGEE Stellar Parameter and Chemical Abundances Pipeline". In: *The Astronomical Journal* 151.6, p. 144. ISSN: 0004-6256. DOI: 10.3847/0004-6256/151/6/144. (Visited on 11/21/2019).

Gardner, Jonathan P. et al. (Apr. 2006). "The James Webb Space Telescope". In: *Space Science Reviews* 123, pp. 485–606. ISSN: 0038-6308. DOI: 10.1007/s11214-006-8315-7.

Gilbert, Karoline M. et al. (Oct. 2019). "Elemental Abundances in M31: First Alpha and Iron Abundance Measurements in M31's Giant Stellar Stream". In: *The Astrophysical Journal* 883, p. 128. ISSN: ['0004-637X']. DOI: 10.3847/1538-4357/ab3807. (Visited on 11/21/2019).

Gilmore, G. et al. (Mar. 2012). "The Gaia-ESO Public Spectroscopic Survey". In: *The Messenger* 147, pp. 25–31. ISSN: 0722-6691.

Gilmozzi, R. and J. Spyromilio (Mar. 2007). "The European Extremely Large Telescope (E-ELT)". In: *The Messenger* 127, p. 11. ISSN: 0722-6691.

Goodman, Jonathan and Jonathan Weare (Jan. 2010). "Ensemble Samplers with Affine Invariance". In: *Communications in Applied Mathematics and Computational Science* 5, pp. 65–80. DOI: 10.2140/camcos.2010.5.65.

Gray, David F. (2021). *The Observation and Analysis of Stellar Photospheres*. 4th ed. Cambridge: Cambridge University Press. DOI: 10.1017/9781009082136.

Greenstein, Jesse L. (Mar. 1948). "Spectrophotometry of the F Stars and of τ Ursae Majoris. I." In: *The Astrophysical Journal* 107, p. 151. ISSN: 0004-637X. DOI: 10.1086/145002. (Visited on 01/21/2020).

Griffith, Emily J. et al. (Feb. 2023). "Untangling the Sources of Abundance Dispersion in Low-metallicity Stars". In: *The Astrophysical Journal* 944, p. 47. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aca659. (Visited on 03/09/2023).

Guiglion, G. et al. (Dec. 2020). "The RAdial Velocity Experiment (RAVE): Parameterisation of RAVE Spectra Based on Convolutional Neural Networks". In: *Astronomy and Astrophysics* 644, A168. ISSN: 0004-6361. DOI: 10.1051/0004-6361/202038271. (Visited on 07/17/2023).

Hansen, C. J. et al. (Sept. 2015). "Stellar Science from a Blue Wavelength Range. A Possible Design for the Blue Arm of 4MOST". In: *Astronomische Nachrichten* 336, p. 665. ISSN: 0004-6337. DOI: 10.1002/asna.201512206. (Visited on 01/10/2020).

Hayden, Michael R. et al. (Aug. 2015). "Chemical Cartography with APOGEE: Metallicity Distribution Functions and the Chemical Structure of the Milky Way Disk". In: *The Astrophysical Journal* 808, p. 132.

Hearnshaw, John (July 2010). "Auguste Comte's Blunder: An Account of the First Century of Stellar Spectroscopy and How It Took One Hundred Years to Prove That Comte Was Wrong!" In: *Journal of Astronomical History and Heritage* 13, pp. 90–104. ISSN: 1440-2807. (Visited on 12/15/2019).

Heiter, U., P. Jofré, et al. (Oct. 2015). "Gaia FGK Benchmark Stars: Effective Temperatures and Surface Gravities". In: *Astronomy and Astrophysics* 582, A49. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201526319. (Visited on 11/21/2019).

Heiter, U., C. Soubiran, et al. (Jan. 2014). "On the Metallicity of Open Clusters. II. Spectroscopy". In: *Astronomy and Astrophysics* 561, A93. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201322559. (Visited on 07/01/2021).

Hendricks, Benjamin et al. (Dec. 2014). "Insights from the Outskirts: Chemical and Dynamical Properties in the Outer Parts of the Fornax Dwarf Spheroidal Galaxy". In: *Astronomy and Astrophysics* 572, A82. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201424645. (Visited on 01/24/2020).

Hill, V. et al. (June 2019). "VLT/FLAMES High-Resolution Chemical Abundances in Sculptor: A Textbook Dwarf Spheroidal Galaxy". In: *Astronomy and Astrophysics* 626, A15. DOI: 10.1051/0004-6361/201833950. (Visited on 10/08/2019).

Hinkel, Natalie R. et al. (Sept. 2016). "A Comparison of Stellar Elemental Abundance Techniques and Measurements". In: *The Astrophysical Journal Supplement Series* 226, p. 4. ISSN: 0067-0049. DOI: 10.3847/0067-0049/226/1/4. (Visited on 09/13/2022).

Ho, Nhung et al. (Jan. 2015). "Metallicity Evolution of the Six Most Luminous M31 Dwarf Satellites". In: *The Astrophysical Journal* 798, p. 77. ISSN: 0004-637X. DOI: 10.1088/0004-637X/798/2/77. (Visited on 01/24/2020).

Holtzman, Jon A. et al. (Nov. 2015). "Abundances, Stellar Parameters, and Spectra from the SDSS-III/APOGEE Survey". In: *The Astronomical Journal* 150.5, p. 148. DOI: 10.1088/0004-6256/150/5/148. (Visited on 10/04/2019).

Hopkins, Philip F., Dušan Kereš, et al. (Nov. 2014). "Galaxies on FIRE (Feedback In Realistic Environments): Stellar Feedback Explains Cosmologically Inefficient Star Formation". In: *Monthly Notices of the Royal Astronomical Society* 445, pp. 581–603. ISSN: 0035-8711. DOI: 10.1093/mnras/stu1738. (Visited on 10/19/2022).

Hopkins, Philip F., Eliot Quataert, and Norman Murray (Apr. 2012). "Stellar Feedback in Galaxies and the Origin of Galaxy-Scale Winds". In: *Monthly Notices of the Royal Astronomical Society* 421, pp. 3522–3537. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2012.20593.x. (Visited on 06/07/2023).

Hopkins, Philip F., Andrew Wetzel, et al. (Oct. 2018). "FIRE-2 Simulations: Physics versus Numerics in Galaxy Formation". In: *Monthly Notices of the Royal Astronomical Society* 480, pp. 800–863. ISSN: 0035-8711. DOI: 10.1093/mnras/sty1690. (Visited on 06/07/2023).

Huggins, William and W. A. Miller (1864). "On the Spectra of Some of the Fixed Stars". In: *Philosophical Transactions of the Royal Society of London Series I* 154, pp. 413–435. (Visited on 01/21/2020).

Husser, Tim-Oliver et al. (Apr. 2016). "MUSE Crowded Field 3D Spectroscopy of over 12 000 Stars in the Globular Cluster NGC 6397. I. The First Comprehensive HRD of a Globular Cluster". In: *Astronomy and Astrophysics* 588, A148. DOI: 10.1051/0004-6361/201526949. (Visited on 10/08/2019).

Ireland, J. (Feb. 2005). "Precision Limits to Emission-Line Profile Measuring Experiments". In: *The Astrophysical Journal* 620.2, p. 1132. DOI: 10.1086/427230. (Visited on 10/04/2019).

Ishigaki, M. N., W. Aoki, and M. Chiba (July 2013). "Chemical Abundances of the Milky Way Thick Disk and Stellar Halo. II. Sodium, Iron-peak, and Neutron-capture Elements". In: *The Astrophysical Journal* 771, p. 67. ISSN: 0004-637X. DOI: 10.1088/0004-637X/771/1/67. (Visited on 08/30/2022).

Ivezić, Željko et al. (Mar. 2019). "LSST: From Science Drivers to Reference Design and Anticipated Data Products". In: *The Astrophysical Journal* 873, p. 111. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab042c. (Visited on 07/17/2023).

Jagourel, P. et al. (July 2018). "MOSAIC: The ELT Multi-Object Spectrograph". In: *Ground-Based and Airborne Instrumentation for Astronomy VII*. Vol. 10702. International Society for Optics and Photonics, 10702A4. DOI: 10.1117/12.2314135. (Visited on 11/21/2019).

Ji, Alexander P., Anna Frebel, Rana Ezzeddine, et al. (Nov. 2016). "Chemical Diversity in the Ultra-faint Dwarf Galaxy Tucana II". In: *The Astrophysical Journal* 832.1, p. L3. DOI: 10.3847/2041-8205/832/1/L3. (Visited on 10/07/2019).

Ji, Alexander P., Anna Frebel, Joshua D. Simon, and Anirudh Chiti (Oct. 2016). "Complete Element Abundances of Nine Stars in the R-Process Galaxy Reticulum II". In: *The Astrophysical Journal* 830, p. 93. DOI: 10.3847/0004-637X/830/2/93. (Visited on 10/07/2019).

Ji, Alexander P., Anna Frebel, Joshua D. Simon, and Marla Geha (Jan. 2016). "High-Resolution Spectroscopy of Extremely Metal-poor Stars in the Least-evolved Galaxies: Bootes II". In: *The Astrophysical Journal* 817, p. 41. DOI: 10.3847/0004-637X/817/1/41. (Visited on 10/07/2019).

Ji, Alexander P., Joshua D. Simon, et al. (Jan. 2019). "Chemical Abundances in the Ultra-faint Dwarf Galaxies Grus I and Triangulum II: Neutron-capture Elements as a Defining Feature of the Faintest Dwarfs". In: *The Astrophysical Journal* 870, p. 83. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aaf3bb. (Visited on 10/19/2022).

Jofré, P., U. Heiter, C. C. Worley, et al. (May 2017). "Gaia FGK Benchmark Stars: Opening the Black Box of Stellar Element Abundance Determination". In: *A&A* 601, A38. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201629833. (Visited on 03/02/2023).

Jofré, Paula, Ulrike Heiter, and Caroline Soubiran (Aug. 2019). "Accuracy and Precision of Industrial Stellar Abundances". In: *Annual Review of Astronomy and Astrophysics* 57, p. 571. ISSN: 0066-4146. DOI: 10.1146/annurev-astro-091918-104509.

Johnson, James W., Charlie Conroy, et al. (Oct. 2022). *Dwarf Galaxy Archaeology from Chemical Abundances and Star Formation Histories*. DOI: 10.48550/arXiv.2210.01816. (Visited on 03/09/2023).

Johnson, James W., Christopher S. Kochanek, and K. Z. Stanek (Oct. 2022). *Binaries Drive High Type Ia Supernova Rates in Dwarf Galaxies*. DOI: 10.48550/arXiv.2210.01818. (Visited on 03/09/2023).

Johnson, James W. and David H. Weinberg (Oct. 2020). "The Impact of Starbursts on Element Abundance Ratios". In: *Monthly Notices of the Royal Astronomical Society* 498, pp. 1364–1381. ISSN: 0035-8711. DOI: 10.1093/mnras/staa2431. (Visited on 03/09/2023).

Johnson, James W., David H. Weinberg, et al. (Dec. 2021). "Stellar Migration and Chemical Enrichment in the Milky Way Disc: A Hybrid Model". In: *Monthly Notices of the Royal Astronomical Society* 508, pp. 4484–4511. ISSN: 0035-8711. DOI: 10.1093/mnras/stab2718. (Visited on 06/07/2023).

Jönsson, Henrik et al. (Sept. 2020). "APOGEE Data and Spectral Analysis from SDSS Data Release 16: Seven Years of Observations Including First Results from APOGEE-South". In: *The Astronomical Journal* 160, p. 120. ISSN: 0004-6256. DOI: 10.3847/1538-3881/aba592.

Jorgensen, U. G., M. Carlsson, and H. R. Johnson (Feb. 1992). "The Calcium Infrared Triplet Lines in Stellar Spectra". In: *Astronomy and Astrophysics* 254, p. 258. ISSN: 0004-6361. (Visited on 12/18/2019).

Kalirai, Jason S. et al. (Mar. 2010). "The SPLASH Survey: Internal Kinematics, Chemical Abundances, and Masses of the Andromeda I, II, III, VII, X, and XIV Dwarf Spheroidal

Galaxies". In: *The Astrophysical Journal* 711, pp. 671–692. ISSN: 0004-637X. DOI: 10.1088/0004-637X/711/2/671. (Visited on 01/15/2020).

Kamann, S., T.-O. Husser, S. Dreizler, et al. (Feb. 2018). "A Stellar Census in Globular Clusters with MUSE: The Contribution of Rotation to Cluster Dynamics Studied with 200 000 Stars". In: *Monthly Notices of the Royal Astronomical Society* 473, p. 5591. DOI: 10.1093/mnras/stx2719. (Visited on 10/08/2019).

Kamann, S., T.-O. Husser, M. Wendt, et al. (June 2016). "A Stellar Census in NGC 6397 with MUSE". In: *The Messenger* 164, p. 18. (Visited on 10/08/2019).

Käppeler, F. et al. (Jan. 2011). "The s Process: Nuclear Physics, Stellar Models, and Observations". In: *Reviews of Modern Physics* 83, pp. 157–194. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.83.157. (Visited on 07/25/2023).

Karamanis, Minas, Florian Beutler, et al. (Oct. 2022). "Accelerating Astronomical and Cosmological Inference with Preconditioned Monte Carlo". In: *Monthly Notices of the Royal Astronomical Society* 516, pp. 1644–1653. ISSN: 0035-8711. DOI: 10.1093/mnras/stac2272. (Visited on 03/09/2023).

Karamanis, Minas, David Nabergoj, et al. (Nov. 2022). "pocoMC: A Python Package for Accelerated Bayesian Inference in Astronomy and Cosmology". In: *The Journal of Open Source Software* 7, p. 4634. DOI: 10.21105/joss.04634. (Visited on 03/09/2023).

Kay, Steven M. (Apr. 1993). *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. 1 edition. Englewood Cliffs, N.J: Prentice Hall. ISBN: 978-0-13-345711-7.

King, Anthea L. et al. (July 2014). "High-Redshift Standard Candles: Predicted Cosmological Constraints". In: *Monthly Notices of the Royal Astronomical Society* 441, pp. 3454–3476. ISSN: 0035-8711. DOI: 10.1093/mnras/stu793. (Visited on 12/30/2019).

Kirby, Evan N., Judith G. Cohen, Joshua D. Simon, and Puragra Guhathakurta (Nov. 2015). "Triangulum II: Possibly a Very Dense Ultra-faint Dwarf Galaxy". In: *The Astrophysical Journal Letters* 814, p. L7. ISSN: 0004-637X. DOI: 10.1088/2041-8205/814/1/L7. (Visited on 01/15/2020).

Kirby, Evan N., Judith G. Cohen, Joshua D. Simon, Puragra Guhathakurta, et al. (Apr. 2017). "Triangulum II. Not Especially Dense After All". In: *The Astrophysical Journal* 838, p. 83. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aa6570. (Visited on 01/15/2020).

Kirby, Evan N., Judith. G. Cohen, Graeme H. Smith, et al. (Feb. 2011). "Multi-Element Abundance Measurements from Medium-resolution Spectra. IV. Alpha Element Distributions in Milky Way Satellite Galaxies". In: *The Astrophysical Journal* 727, p. 79. ISSN: 0004-637X. DOI: 10.1088/0004-637X/727/2/79.

Kirby, Evan N., Karoline M. Gilbert, et al. (Feb. 2020). "Elemental Abundances in M31: The Kinematics and Chemical Evolution of Dwarf Spheroidal Satellite Galaxies". In: *The Astronomical Journal* 159, p. 46. ISSN: 0004-6256. DOI: 10.3847/1538-3881/ab5f0f. (Visited on 01/27/2023).

Kirby, Evan N., Puragra Guhathakurta, Michael Bolte, et al. (Nov. 2009). "Multi-Element Abundance Measurements from Medium-resolution Spectra. I. The Sculptor Dwarf Spheroidal Galaxy". In: *The Astrophysical Journal* 705.1, p. 328. DOI: 10.1088/0004-637X/705/1/328. (Visited on 09/25/2019).

Kirby, Evan N., Puragra Guhathakurta, Joshua D. Simon, et al. (Dec. 2010). "Multi-Element Abundance Measurements from Medium-resolution Spectra. II. Catalog of Stars in Milky Way Dwarf Satellite Galaxies". In: *The Astrophysical Journal Supplement Series* 191.2, p. 352. DOI: 10.1088/0067-0049/191/2/352. arXiv: 1011.4516 [astro-ph.GA].

Kirby, Evan N., Puragra Guhathakurta, and Christopher Sneden (Aug. 2008). "Metallicity and Alpha-Element Abundance Measurement in Red Giant Stars from Medium-Resolution Spectra". In: *The Astrophysical Journal* 682.2, p. 1217. DOI: 10.1086/589627. (Visited on 09/25/2019).

Kirby, Evan N., Michelle Guo, et al. (Mar. 2015). "Carbon in Red Giants in Globular Clusters and Dwarf Spheroidal Galaxies". In: *The Astrophysical Journal* 801, p. 125. ISSN: 0004-637X. DOI: 10.1088/0004-637X/801/2/125. (Visited on 01/15/2020).

Kirby, Evan N., Gustavo A. Lanfranchi, et al. (Feb. 2011). "Multi-Element Abundance Measurements from Medium-resolution Spectra. III. Metallicity Distributions of Milky Way Dwarf Satellite Galaxies". In: *The Astrophysical Journal* 727.2, p. 78. DOI: 10.1088/0004-637X/727/2/78. (Visited on 10/04/2019).

Kirby, Evan N., Luca Rizzi, et al. (Jan. 2017). "Chemistry and Kinematics of the Late-forming Dwarf Irregular Galaxies Leo A, Aquarius, and Sagittarius DIG". In: *The Astrophysical Journal* 834, p. 9. ISSN: 0004-637X. DOI: 10.3847/1538-4357/834/1/9. (Visited on 01/15/2020).

Kirby, Evan N., Joshua D. Simon, and Judith G. Cohen (Sept. 2015). "Spectroscopic Confirmation of the Dwarf Galaxies Hydra II and Pisces II and the Globular Cluster Laevens 1". In: *The Astrophysical Journal* 810.1, p. 56. DOI: 10.1088/0004-637X/810/1/56. (Visited on 10/07/2019).

Kirby, Evan N., Justin L. Xie, et al. (July 2018). "Catalog of Chromium, Cobalt, and Nickel Abundances in Globular Clusters and Dwarf Galaxies". In: *The Astrophysical Journal Supplement Series* 237.1, p. 18. DOI: 10.3847/1538-4365/aac952.

Kirchhoff, G. (1860). "Ueber Das Verhältniss Zwischen Dem Emissionsvermögen Und Dem Absorptionsvermögen Der Körper Für Wärme Und Licht". In: *Annalen der Physik* 185, pp. 275–301. ISSN: 0003-3804. DOI: 10.1002/andp.18601850205. (Visited on 01/21/2020).

— (1863). "Zur Geschichte Der Spectral-Analyse Und Der Analyse Der Sonnenatmosphäre". In: *Annalen der Physik* 194, pp. 94–111. ISSN: 0003-3804. DOI: 10.1002/andp.18631940106. (Visited on 01/21/2020).

Kirchhoff, G. and R. Bunsen (1860). "Chemische Analyse Durch Spectralbeobachtungen". In: *Annalen der Physik* 186, pp. 161–189. ISSN: 0003-3804. DOI: 10.1002/andp.18601860602. (Visited on 01/21/2020).

Koch, A. and P. Côté (July 2010). "Chemical Abundances for the Outer Halo Cluster Pal 4 from Co-Added High-Resolution Spectroscopy". In: *Astronomy and Astrophysics* 517, A59. ISSN: ['0004-6361']. DOI: 10.1051/0004-6361/201014155. (Visited on 11/21/2019).

Koch, Andreas, Eva K. Grebel, Gerard F. Gilmore, et al. (Apr. 2008). "Complexity on Small Scales. III. Iron and α Element Abundances in the Carina Dwarf Spheroidal Galaxy". In:

The Astronomical Journal 135, p. 1580. DOI: 10.1088/0004-6256/135/4/1580. (Visited on 11/01/2019).

Koch, Andreas, Eva K. Grebel, Jan T. Kleyna, et al. (Jan. 2007). "Complexity on Small Scales. II. Metallicities and Ages in the Leo II Dwarf Spheroidal Galaxy". In: *The Astronomical Journal* 133, pp. 270–283. ISSN: 0004-6256. DOI: 10.1086/509889. (Visited on 01/21/2020).

Koch, Andreas, Andrew McWilliam, et al. (Nov. 2008). "The Highly Unusual Chemical Composition of the Hercules Dwarf Spheroidal Galaxy". In: *The Astrophysical Journal Letters* 688, p. L13. ISSN: 0004-637X. DOI: 10.1086/595001. (Visited on 01/21/2020).

Koch, Andreas and R. Michael Rich (Oct. 2014). "A Chemical Confirmation of the Faint Boötes II Dwarf Spheroidal Galaxy". In: *The Astrophysical Journal* 794.1, p. 89. DOI: 10.1088/0004-637X/794/1/89. (Visited on 10/07/2019).

Koch, Andreas, Mark I. Wilkinson, Jan T. Kleyna, Gerard F. Gilmore, et al. (Mar. 2007). "Stellar Kinematics and Metallicities in the Leo I Dwarf Spheroidal Galaxy-Wide-Field Implications for Galactic Evolution". In: *The Astrophysical Journal* 657, pp. 241–261. ISSN: 0004-637X. DOI: 10.1086/510879. (Visited on 01/21/2020).

Koch, Andreas, Mark I. Wilkinson, Jan T. Kleyna, Mike Irwin, et al. (Jan. 2009). "A Spectroscopic Confirmation of the Bootes II Dwarf Spheroidal". In: *The Astrophysical Journal* 690, pp. 453–462. ISSN: 0004-637X. DOI: 10.1088/0004-637X/690/1/453. (Visited on 01/21/2020).

Kollmeier, Juna A. et al. (Nov. 2017). *SDSS-V: Pioneering Panoptic Spectroscopy*. (Visited on 09/13/2022).

Koposov, Sergey E. et al. (June 2015). "Beasts of the Southern Wild: Discovery of Nine Ultra Faint Satellites in the Vicinity of the Magellanic Clouds." In: *The Astrophysical Journal* 805, p. 130. ISSN: 0004-637X. DOI: 10.1088/0004-637X/805/2/130. (Visited on 10/17/2022).

Kovalev, Mikhail et al. (Aug. 2019). "Non-LTE Chemical Abundances in Galactic Open and Globular Clusters". In: *Astronomy and Astrophysics* 628, A54. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201935861.

Kroupa, Pavel (Apr. 2001). "On the Variation of the Initial Mass Function". In: *Monthly Notices of the Royal Astronomical Society* 322, pp. 231–246. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2001.04022.x. (Visited on 08/30/2022).

Kurucz, R. L. (Feb. 1970). "Atlas: A Computer Program for Calculating Model Stellar Atmospheres". In: *SAO Special Report* 309.

Kurucz, Robert L. (2005). "ATLAS12, SYNTHE, ATLAS9, WIDTH9, et Cetera". In: *Memorie della Societa Astronomica Italiana Supplementi* 8, p. 14.

— (Mar. 2013). "ATLAS12: Opacity Sampling Model Atmosphere Program". In: *Astrophysics Source Code Library*, ascl:1303.024. (Visited on 10/07/2019).

— (Oct. 2017). "ATLAS9: Model Atmosphere Program with Opacity Distribution Functions". In: *Astrophysics Source Code Library*, ascl:1710.017.

— (1993). *SYNTHE Spectrum Synthesis Programs and Line Data*. CD-ROM. (Visited on 10/07/2019).

Kurucz, Robert L. and Eugene H. Avrett (May 1981). "Solar Spectrum Synthesis. I. A Sample Atlas from 224 to 300 Nm". In: *SAO Special Report* 391.

Kustner, Friedrich (Jan. 1921). "Der Kugelfoermige Sternhaufen Messier 15". In: *Veroeffentlichungen des Astronomisches Institute der Universitaet Bonn* 15, p. 1.

Lacchin, E. et al. (July 2020). "Chemical Evolution of Ultrafaint Dwarf Galaxies: Testing the IGIMF". In: *Monthly Notices of the Royal Astronomical Society* 495, pp. 3276–3294. ISSN: 0035-8711. DOI: 10.1093/mnras/staa585. (Visited on 06/06/2023).

Lai, David K. et al. (Sept. 2011). "The [Fe/H], [C/Fe], and [α/Fe] Distributions of the Boötes I Dwarf Spheroidal Galaxy". In: *The Astrophysical Journal* 738.1, p. 51. DOI: 10.1088/0004-637X/738/1/51. (Visited on 10/04/2019).

Lanfranchi, G. A. and F. Matteucci (Mar. 2010). "Chemical Evolution Models for the Dwarf Spheroidal Galaxies Leo 1 and Leo 2". In: *Astronomy and Astrophysics* 512, A85. ISSN: 0004-6361. DOI: 10.1051/0004-6361/200913045. (Visited on 06/06/2023).

— (June 2007). "Effects of the Galactic Winds on the Stellar Metallicity Distribution of Dwarf Spheroidal Galaxies". In: *Astronomy and Astrophysics* 468, pp. 927–936. ISSN: 0004-6361. DOI: 10.1051/0004-6361:20066576. (Visited on 06/06/2023).

Lanfranchi, G. A., F. Matteucci, and G. Cescutti (July 2006). "Detailed Chemical Evolution of Carina and Sagittarius Dwarf Spheroidal Galaxies". In: *Astronomy and Astrophysics* 453, pp. 67–75. ISSN: 0004-6361. DOI: 10.1051/0004-6361:20054627. (Visited on 06/07/2023).

Lanfranchi, Gustavo A. and Francesca Matteucci (July 2004). "The Predicted Metallicity Distribution of Stars in Dwarf Spheroidal Galaxies". In: *Monthly Notices of the Royal Astronomical Society* 351, pp. 1338–1348. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2004.07877.x. (Visited on 06/06/2023).

Larsen, S. S. et al. (Apr. 2022). "The Chemical Composition of Globular Clusters in the Local Group". In: *A&A* 660, A88. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202142243. (Visited on 03/02/2023).

Latour, M. et al. (Sept. 2019). "A Stellar Census in Globular Clusters with MUSE: Multiple Populations Chemistry in NGC 2808". In: *arXiv e-prints*, arXiv:1909.04959. (Visited on 10/08/2019).

Laureijs, R. et al. (Oct. 2011). *Euclid Definition Study Report*. DOI: 10.48550/arXiv.1110.3193. (Visited on 07/17/2023).

Laverick, M. et al. (Apr. 2018). "The Belgian Repository of Fundamental Atomic Data and Stellar Spectra (BRASS). I. Cross-matching Atomic Databases of Astrophysical Interest". In: *Astronomy and Astrophysics* 612, A60. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201731933. (Visited on 01/03/2020).

Lawler, J. E. et al. (Apr. 2013). "Improved Log(Gf) Values for Lines of Ti I and Abundance Determinations in the Photospheres of the Sun and Metal-Poor Star HD 84937 (Accurate Transition Probabilities for Ti I)". In: *The Astrophysical Journal Supplement Series* 205, p. 11. ISSN: 0067-0049. DOI: 10.1088/0067-0049/205/2/11. (Visited on 09/13/2022).

Leaman, Ryan et al. (July 2009). "Stellar Metallicities and Kinematics in a Gas-rich Dwarf Galaxy: First Calcium Triplet Spectroscopy of Red Giant Branch Stars in WLM". In:

The Astrophysical Journal 699, pp. 1–14. ISSN: 0004-637X. DOI: 10.1088/0004-637X/699/1/1. (Visited on 01/21/2020).

Leep, E. M., G. Wallerstein, and J. B. Oke (May 1986). "High-Resolution CCD Spectra of Stars in Globular Clusters. I - Oxygen in M13". In: *The Astronomical Journal* 91, pp. 1117–1120. ISSN: 0004-6256. DOI: 10.1086/114088. (Visited on 01/21/2020).

Leep, E. Myckky, J. B. Oke, and George Wallerstein (Feb. 1987). "High-Resolution CCD Spectra of Stars in Globular Clusters. II - Metals and CNO in M71". In: *The Astronomical Journal* 93, pp. 338–346. ISSN: 0004-6256. DOI: 10.1086/114318. (Visited on 01/21/2020).

Leroy, Adam K. et al. (Dec. 2008). "The Star Formation Efficiency in Nearby Galaxies: Measuring Where Gas Forms Stars Effectively". In: *The Astronomical Journal* 136, pp. 2782–2845. ISSN: 0004-6256. DOI: 10.1088/0004-6256/136/6/2782. (Visited on 10/19/2022).

Letarte, B. et al. (July 2006). "VLT/UVES Spectroscopy of Individual Stars in Three Globular Clusters in the Fornax Dwarf Spheroidal Galaxy". In: *Astronomy & Astrophysics* 453.2, pp. 547–554. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361:20054439.

Li, T. S. et al. (Mar. 2017). "Farthest Neighbor: The Distant Milky Way Satellite Eridanus II". In: *The Astrophysical Journal* 838, p. 8. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aa6113. (Visited on 01/24/2020).

Liu, Ming-Yu, Thomas Breuel, and Jan Kautz (Mar. 2017). *Unsupervised Image-to-Image Translation Networks*. DOI: 10.48550/arXiv.1703.00848. (Visited on 07/18/2023).

Longeard, Nicolas et al. (Jan. 2020). "The Pristine Dwarf-Galaxy Survey - II. In-depth Observational Study of the Faint Milky Way Satellite Sagittarius II". In: *Monthly Notices of the Royal Astronomical Society* 491, pp. 356–377. ISSN: 0035-8711. DOI: 10.1093/mnras/stz2854. (Visited on 01/24/2020).

Lucchesi, R. et al. (Jan. 2020). "Homogeneity in the Early Chemical Evolution of the Sextans Dwarf Spheroidal Galaxy". In: *arXiv e-prints* 2001, arXiv:2001.11033. (Visited on 02/26/2020).

Luo, A.-Li et al. (Aug. 2015). "The First Data Release (DR1) of the LAMOST Regular Survey". In: *Research in Astronomy and Astrophysics* 15.8, p. 1095. ISSN: 1674-4527. DOI: 10.1088/1674-4527/15/8/002. (Visited on 11/21/2019).

Majewski, Steven R. et al. (Sept. 2017). "The Apache Point Observatory Galactic Evolution Experiment (APOGEE)". In: *The Astronomical Journal* 154.3, p. 94. ISSN: 0004-6256. DOI: 10.3847/1538-3881/aa784d. (Visited on 11/21/2019).

Maoz, Dan and Or Graur (Oct. 2017). "Star Formation, Supernovae, Iron, and α: Consistent Cosmic and Galactic Histories". In: *The Astrophysical Journal* 848, p. 25. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aa8b6e. (Visited on 10/17/2022).

Maoz, Dan, Filippo Mannucci, and Timothy D. Brandt (Nov. 2012). "The Delay-Time Distribution of Type Ia Supernovae from Sloan II". In: *Monthly Notices of the Royal Astronomical Society* 426, pp. 3282–3294. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2012.21871.x. (Visited on 08/27/2022).

Martin, N. F., R. A. Ibata, S. C. Chapman, et al. (Sept. 2007). "A Keck/DEIMOS Spectroscopic Survey of Faint Galactic Satellites: Searching for the Least Massive Dwarf

Galaxies". In: *Monthly Notices of the Royal Astronomical Society* 380.1, pp. 281–300. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2007.12055.x. (Visited on 10/07/2019).

Martin, Nicolas F., Marla Geha, et al. (May 2016). "Is Draco II One of the Faintest Dwarf Galaxies? First Study from Keck/DEIMOS Spectroscopy". In: *Monthly Notices of the Royal Astronomical Society* 458.1, p. L59. DOI: 10.1093/mnrasl/slw013. (Visited on 10/07/2019).

Martin, Nicolas F., Rodrigo A. Ibata, Michelle L. M. Collins, et al. (Feb. 2016). "Triangulum II: A Very Metal-poor and Dynamically Hot Stellar System". In: *The Astrophysical Journal* 818, p. 40. ISSN: 0004-637X. DOI: 10.3847/0004-637X/818/1/40. (Visited on 01/24/2020).

Martínez-Vázquez, C. E. et al. (Nov. 2021). "Variable Stars in Local Group Galaxies - V. The Fast and Early Evolution of the Low-Mass Eridanus II dSph Galaxy". In: *Monthly Notices of the Royal Astronomical Society* 508, pp. 1064–1083. ISSN: 0035-8711. DOI: 10.1093/mnras/stab2493. (Visited on 03/09/2023).

Martins, Lucimara P. and Paula Coelho (Nov. 2007). "Testing the Accuracy of Synthetic Stellar Libraries". In: *Monthly Notices of the Royal Astronomical Society* 381, pp. 1329–1346. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2007.11954.x. (Visited on 08/07/2023).

Mashonkina, L. and T. Gehren (Dec. 2000). "Barium and Europium Abundances in Cool Dwarf Stars and Nucleosynthesis of Heavy Elements". In: *Astronomy and Astrophysics* 364, pp. 249–264. ISSN: 0004-6361. (Visited on 10/08/2022).

Mashonkina, L., A. J. Korn, and N. Przybilla (Jan. 2007). "A Non-LTE Study of Neutral and Singly-Ionized Calcium in Late-Type Stars". In: *Astronomy and Astrophysics* 461, pp. 261–275. ISSN: 0004-6361. DOI: 10.1051/0004-6361:20065999. (Visited on 12/18/2019).

Masseron, T. et al. (Feb. 2019). "Homogeneous Analysis of Globular Clusters from the APOGEE Survey with the BACCHUS Code - I. The Northern Clusters". In: *A&A* 622, A191. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201834550. (Visited on 03/02/2023).

Mateo, Mario et al. (Sept. 2012). "M2FS: The Michigan/Magellan Fiber System". In: *Ground-Based and Airborne Instrumentation for Astronomy IV*. Vol. 8446. International Society for Optics and Photonics, 84464Y. DOI: 10.1117/12.926448. (Visited on 10/07/2019).

Matteucci, F. and P. Francois (Aug. 1989). "Galactic Chemical Evolution : Abundance Gradients of Individual Elements." In: *Monthly Notices of the Royal Astronomical Society* 239, pp. 885–904. ISSN: 0035-8711. DOI: 10.1093/mnras/239.3.885. (Visited on 06/07/2023).

Matteucci, Francesca (Dec. 2021). "Modelling the Chemical Evolution of the Milky Way". In: *Astronomy and Astrophysics Review* 29, p. 5. ISSN: 0935-4956. DOI: 10.1007/s00159-021-00133-8. (Visited on 08/07/2023).

McLeod, Anna F. et al. (Mar. 2020). "Stellar Feedback and Resolved Stellar IFU Spectroscopy in the Nearby Spiral Galaxy NGC 300". In: *The Astrophysical Journal* 891, p. 25. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab6d63. (Visited on 06/30/2023).

McQuinn, Kristen. B. W., Liese van Zee, and Evan D. Skillman (Nov. 2019). "Galactic Winds in Low-mass Galaxies". In: *The Astrophysical Journal* 886, p. 74. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab4c37. (Visited on 08/30/2022).

McWilliam, Andrew (Jan. 1997). "Abundance Ratios and Galactic Chemical Evolution". In: *Annual Review of Astronomy and Astrophysics* 35, pp. 503–556. ISSN: 0066-4146. DOI: 10.1146/annurev.astro.35.1.503. (Visited on 07/25/2023).

Mészáros, Szabolcs et al. (May 2015). "Exploring Anticorrelations and Light Element Variations in Northern Globular Clusters Observed by the APOGEE Survey". In: *The Astronomical Journal* 149, p. 153. ISSN: 0004-6256. DOI: 10.1088/0004-6256/149/5/153. (Visited on 08/30/2022).

Minchev, I., C. Chiappini, and M. Martig (Oct. 2013). "Chemodynamical Evolution of the Milky Way Disk. I. The Solar Vicinity". In: *Astronomy and Astrophysics* 558, A9.

Minnaert, M. (Nov. 1934). "The Measurement of Equivalent Widths as a Method for the Investigation of Line Profiles". In: *The Observatory* 57, pp. 328–342. ISSN: 0029-7704. (Visited on 01/21/2020).

Mitchell, Peter D. et al. (May 2020). "Galactic Outflow Rates in the EAGLE Simulations". In: *Monthly Notices of the Royal Astronomical Society* 494, pp. 3971–3997. ISSN: 0035-8711. DOI: 10.1093/mnras/staa938. (Visited on 10/19/2022).

Moore, E.H. (1920). "On the Reciprocal of the General Algebraic Matrix". In: *Bulletin of the American Mathematical Society* 26, pp. 394–395. DOI: 10.1090/S0002-9904-1920-03322-7..

MSE Science Team et al. (Apr. 2019). "The Detailed Science Case for the Maunakea Spectroscopic Explorer, 2019 Edition". In: *arXiv e-prints*, arXiv:1904.04907. (Visited on 11/21/2019).

Muñoz, Ricardo R. et al. (Sept. 2006). "Exploring Halo Substructure with Giant Stars. XI. The Tidal Tails of the Carina Dwarf Spheroidal Galaxy and the Discovery of Magellanic Cloud Stars in the Carina Foreground". In: *The Astrophysical Journal* 649, pp. 201–223. ISSN: 0004-637X. DOI: 10.1086/505620. (Visited on 01/21/2020).

Muratov, Alexander L. et al. (Dec. 2015). "Gusty, Gaseous Flows of FIRE: Galactic Winds in Cosmological Simulations with Explicit Stellar Feedback". In: *Monthly Notices of the Royal Astronomical Society* 454, pp. 2691–2713. ISSN: 0035-8711. DOI: 10.1093/mnras/stv2126. (Visited on 10/19/2022).

Murray, Norman, Eliot Quataert, and Todd A. Thompson (Jan. 2005). "On the Maximum Luminosity of Galaxies and Their Central Black Holes: Feedback from Momentum-driven Winds". In: *The Astrophysical Journal* 618, pp. 569–585. ISSN: 0004-637X. DOI: 10.1086/426067. (Visited on 10/24/2022).

— (Jan. 2010). "The Disruption of Giant Molecular Clouds by Radiation Pressure & the Efficiency of Star Formation in Galaxies". In: *The Astrophysical Journal* 709, pp. 191–209. ISSN: 0004-637X. DOI: 10.1088/0004-637X/709/1/191. (Visited on 06/07/2023).

Mutlu-Pakdil, Burçin et al. (Sept. 2021). "Resolved Dwarf Galaxy Searches within 5 Mpc with the Vera Rubin Observatory and Subaru Hyper Suprime-Cam". In: *The Astrophys-*

ical Journal 918, p. 88. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ac0db8. (Visited on 07/17/2023).

Nelson, Benjamin, Eric B. Ford, and Matthew J. Payne (Jan. 2014). "RUN DMC: An Efficient, Parallel Code for Analyzing Radial Velocity Observations Using N-body Integrations and Differential Evolution Markov Chain Monte Carlo". In: *The Astrophysical Journal Supplement Series* 210, p. 11. ISSN: 0067-0049. DOI: 10.1088/0067-0049/210/1/11. (Visited on 08/30/2022).

Ness, M., K. Freeman, et al. (Apr. 2013). "ARGOS - III. Stellar Populations in the Galactic Bulge of the Milky Way". In: *Monthly Notices of the Royal Astronomical Society* 430, pp. 836–857. ISSN: 0035-8711. DOI: 10.1093/mnras/sts629. (Visited on 07/25/2023).

Ness, M., David W. Hogg, et al. (July 2015). "The Cannon: A Data-Driven Approach to Stellar Label Determination". In: *The Astrophysical Journal* 808.1, p. 16. DOI: 10.1088/0004-637X/808/1/16. (Visited on 09/25/2019).

Nissen, Poul Erik and Bengt Gustafsson (Oct. 2018). "High-Precision Stellar Abundances of the Elements: Methods and Applications". In: *Astronomy and Astrophysics Review* 26, p. 6. ISSN: 0935-4956. DOI: 10.1007/s00159-018-0111-3. (Visited on 01/03/2020).

Nomoto, Ken'ichi, Chiaki Kobayashi, and Nozomu Tominaga (Aug. 2013). "Nucleosynthesis in Stars and the Chemical Enrichment of Galaxies". In: *Annual Review of Astronomy and Astrophysics* 51, pp. 457–509. ISSN: 0066-4146. DOI: 10.1146/annurev-astro-082812-140956. (Visited on 07/25/2023).

Norris, John E. et al. (Dec. 2008). "The Abundance Spread in the Boötes I Dwarf Spheroidal Galaxy". In: *The Astrophysical Journal Letters* 689, p. L113. ISSN: 0004-637X. DOI: 10.1086/595962. (Visited on 01/21/2020).

O'Briain, Teaghan et al. (Jan. 2021). "Cycle-StarNet: Bridging the Gap between Theory and Data by Leveraging Large Data Sets". In: *The Astrophysical Journal* 906, p. 130. ISSN: 0004-637X. DOI: 10.3847/1538-4357/abca96. (Visited on 07/17/2023).

Oke, J. B. et al. (Apr. 1995). "The Keck Low-Resolution Imaging Spectrometer". In: *Publications of the Astronomical Society of the Pacific* 107, p. 375. DOI: 10.1086/133562. (Visited on 10/04/2019).

Olszewski, Edward W. et al. (Feb. 1991). "Spectroscopy of Giants in LMC Clusters. I - Velocities, Abundances, and the Age-Metallicity Relation". In: *The Astronomical Journal* 101, pp. 515–537. ISSN: 0004-6256. DOI: 10.1086/115701. (Visited on 01/21/2020).

Otsuki, Kaori et al. (Apr. 2006). "Neutron-Capture Elements in the Metal-poor Globular Cluster M15". In: *The Astrophysical Journal Letters* 641, pp. L117–L120. ISSN: 0004-637X. DOI: 10.1086/504106. (Visited on 11/30/2020).

Pagel, Bernard E. J. (Oct. 1997). *Nucleosynthesis and Chemical Evolution of Galaxies*. (Visited on 07/25/2023).

Pandya, Viraj et al. (Dec. 2021). "Characterizing Mass, Momentum, Energy, and Metal Outflow Rates of Multiphase Galactic Winds in the FIRE-2 Cosmological Simulations". In: *Monthly Notices of the Royal Astronomical Society* 508, pp. 2979–3008. ISSN: 0035-8711. DOI: 10.1093/mnras/stab2714. (Visited on 06/06/2023).

Pasquini, L. et al. (Dec. 2002). "Installation and Commissioning of FLAMES, the VLT Multifibre Facility". In: *The Messenger* 110, p. 1. ISSN: 0722-6691. (Visited on 10/07/2019).

Paxton, Bill, Lars Bildsten, et al. (Jan. 2011). "Modules for Experiments in Stellar Astrophysics (MESA)". In: *The Astrophysical Journal Supplement Series* 192, p. 3. DOI: 10.1088/0067-0049/192/1/3. (Visited on 10/04/2019).

Paxton, Bill, Matteo Cantiello, et al. (Sept. 2013). "Modules for Experiments in Stellar Astrophysics (MESA): Planets, Oscillations, Rotation, and Massive Stars". In: *The Astrophysical Journal Supplement Series* 208, p. 4. DOI: 10.1088/0067-0049/208/1/4. (Visited on 10/04/2019).

Paxton, Bill, Pablo Marchant, et al. (Sept. 2015). "Modules for Experiments in Stellar Astrophysics (MESA): Binaries, Pulsations, and Explosions". In: *The Astrophysical Journal Supplement Series* 220, p. 15. DOI: 10.1088/0067-0049/220/1/15. (Visited on 10/04/2019).

Paxton, Bill, Josiah Schwab, et al. (Feb. 2018). "Modules for Experiments in Stellar Astrophysics (MESA): Convective Boundaries, Element Diffusion, and Massive Star Explosions". In: *The Astrophysical Journal Supplement Series* 234, p. 34. ISSN: 0067-0049. DOI: 10.3847/1538-4365/aaa5a8. (Visited on 11/30/2020).

Payne, Cecilia Helena (1925). "Stellar Atmospheres; a Contribution to the Observational Study of High Temperature in the Reversing Layers of Stars." PhD thesis. Radcliffe College. (Visited on 01/21/2020).

Pazder, John S. et al. (June 2006). "WFOS: A Wide Field Optical Spectrograph for the Thirty Meter Telescope". In: *Ground-Based and Airborne Instrumentation for Astronomy*. Vol. 6269. International Society for Optics and Photonics, p. 62691X. DOI: 10.1117/12.672712. (Visited on 11/21/2019).

Pearson, Sarah et al. (Dec. 2022). "Mapping Dark Matter with Extragalactic Stellar Streams: The Case of Centaurus A". In: *The Astrophysical Journal* 941, p. 19. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ac9bfb. (Visited on 07/17/2023).

Peeples, Molly S. and Francesco Shankar (Nov. 2011). "Constraints on Star Formation Driven Galaxy Winds from the Mass-Metallicity Relation at z= 0". In: *Monthly Notices of the Royal Astronomical Society* 417, pp. 2962–2981. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2011.19456.x. (Visited on 10/19/2022).

Pehlivan Rhodin, A. et al. (Feb. 2017). "Experimental and Theoretical Oscillator Strengths of Mg I for Accurate Abundance Analysis". In: *Astronomy and Astrophysics* 598, A102. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201629849. (Visited on 01/03/2020).

Penrose, R. (July 1955). "A Generalized Inverse for Matrices". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 51.3, pp. 406–413. ISSN: 1469-8064, 0305-0041. DOI: 10.1017/S0305004100030401. (Visited on 10/07/2019).

Pickering, J. C., A. P. Thorne, and R. Perez (Feb. 2001). "Oscillator Strengths of Transitions in Ti II in the Visible and Ultraviolet Regions". In: *The Astrophysical Journal Supplement Series* 132, pp. 403–409. ISSN: 0067-0049. DOI: 10.1086/318958. (Visited on 01/03/2020).

Pogge, R. W. et al. (July 2010). "The Multi-Object Double Spectrographs for the Large Binocular Telescope". In: *Ground-based and Airborne Instrumentation for Astronomy III* 7735, 77350A. DOI: 10.1117/12.857215. (Visited on 10/07/2019).

Pont, Frédéric et al. (Feb. 2004). "The Chemical Enrichment History of the Fornax Dwarf Spheroidal Galaxy from the Infrared Calcium Triplet". In: *The Astronomical Journal* 127, pp. 840–860. ISSN: 0004-6256. DOI: 10.1086/380608. (Visited on 01/21/2020).

Prochaska, J. et al. (Dec. 2020). "PypeIt: The Python Spectroscopic Data Reduction Pipeline". In: *The Journal of Open Source Software* 5, p. 2308. DOI: 10.21105/joss.02308. (Visited on 06/29/2021).

Rao, Calyampudi Radhakrishna (1945). "Information and Accuracy Attainable in the Estimation of Statistical Parameters". In: *Bulletin of the Calcutta Mathematical Society* 37, pp. 81–91. ISSN: 0008-0659. (Visited on 10/04/2019).

Recio-Blanco, A. et al. (June 2022). *Gaia Data Release 3: Analysis of RVS Spectra Using the General Stellar Parametriser from Spectroscopy.*

Rix, Hans-Walter and Jo Bovy (May 2013). "The Milky Way's Stellar Disk. Mapping and Modeling the Galactic Disk". In: *Astronomy and Astrophysics Review* 21, p. 61. ISSN: 0935-4956. DOI: 10.1007/s00159-013-0061-8. (Visited on 07/25/2023).

Robertson, J. Gordon (Aug. 2017). "Detector Sampling of Optical/IR Spectra: How Many Pixels per FWHM?" In: *Publications of the Astronomical Society of Australia* 34, e035. ISSN: 1323-3580. DOI: 10.1017/pasa.2017.29. arXiv: 1707.06455. (Visited on 05/20/2020).

Rodríguez, Ósmar, Dan Maoz, and Ehud Nakar (Sept. 2022). *The Iron Yield of Core-collapse Supernovae.* DOI: 10.48550/arXiv.2209.05552. (Visited on 03/09/2023).

Roederer, Ian (May 2019). "The Potential of Ultraviolet Spectroscopy to Open New Frontiers to Study the First Stars". In: *Bulletin of the American Astronomical Society* 51, p. 49. (Visited on 10/09/2019).

Roederer, Ian U. and Paul S. Barklem (Apr. 2018). "A New Test of Copper and Zinc Abundances in Late-type Stars Using Ultraviolet Cu II and Zn II Lines". In: *The Astrophysical Journal* 857, p. 2. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aab71f. (Visited on 08/30/2022).

Roederer, Ian U., Karl-Ludwig Kratz, et al. (June 2009). "The End of Nucleosynthesis: Production of Lead and Thorium in the Early Galaxy". In: *The Astrophysical Journal* 698, pp. 1963–1980. ISSN: 0004-637X. DOI: 10.1088/0004-637X/698/2/1963. (Visited on 08/30/2022).

Roederer, Ian U., James E. Lawler, et al. (Mar. 2008). "Europium, Samarium, and Neodymium Isotopic Fractions in Metal-Poor Stars". In: *The Astrophysical Journal* 675, pp. 723–745. ISSN: 0004-637X. DOI: 10.1086/526452. (Visited on 10/05/2022).

Romano, Donatella et al. (Feb. 2015). "Chemical Enrichment in Very Low Metallicity Environments: Boötes I". In: *Monthly Notices of the Royal Astronomical Society* 446, pp. 4220–4231. ISSN: 0035-8711. DOI: 10.1093/mnras/stu2427. (Visited on 06/06/2023).

Roth, Martin M. et al. (Oct. 2018). "MUSE Crowded Field 3D Spectroscopy in NGC 300. I. First Results from Central Fields". In: *Astronomy and Astrophysics* 618, A3. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201833007. (Visited on 06/30/2023).

Ruchti, G. R. et al. (Sept. 2016). "A New Algorithm for Optimizing the Wavelength Coverage for Spectroscopic Studies: Spectral Wavelength Optimization Code (SWOC)". In: *Monthly Notices of the Royal Astronomical Society* 461, pp. 2174–2191. ISSN: 0035-8711. DOI: 10.1093/mnras/stw1351. (Visited on 01/15/2020).

Russell, Henry Norris (July 1929). "On the Composition of the Sun's Atmosphere". In: *The Astrophysical Journal* 70, p. 11. ISSN: 0004-637X. DOI: 10.1086/143197. (Visited on 12/15/2019).

Rutledge, G. A., J. E. Hesser, and P. B. Stetson (Aug. 1997). "Galactic Globular Cluster Metallicity Scale from the Ca II Triplet II. Rankings, Comparisons, and Puzzles". In: *Publications of the Astronomical Society of the Pacific* 109, pp. 907–919. ISSN: 0004-6280. DOI: 10.1086/133959. (Visited on 01/21/2020).

Ryabchikova, T. et al. (May 2015). "A Major Upgrade of the VALD Database". In: *Physica Scripta* 90, p. 054005. ISSN: 0031-8949. DOI: 10.1088/0031-8949/90/5/054005. (Visited on 09/13/2022).

Sandage, A. (Dec. 1970). "Main-Sequence Photometry, Color-Magnitude Diagrams, and Ages for the Globular Clusters M3, M13, M15 and M92." In: *The Astrophysical Journal* 162, p. 841. ISSN: 0004-637X. DOI: 10.1086/150715. (Visited on 06/29/2021).

Sandford, Nathan (June 2020). "Chem-I-Calc: The Chemical Information Calculator". In: *Zenodo*. DOI: 10.5281/zenodo.3924672. (Visited on 11/30/2020).

Sandford, Nathan R., David H. Weinberg, et al. (Oct. 2022). *Strong Outflows and Inefficient Star Formation in the Reionization-era Ultra-faint Dwarf Galaxy Eridanus II*. DOI: 10.48550/arXiv.2210.17045. (Visited on 11/12/2022).

Sandford, Nathan R., Daniel R. Weisz, and Yuan-Sen Ting (Aug. 2020). "Forecasting Chemical Abundance Precision for Extragalactic Stellar Archaeology". In: *The Astrophysical Journal Supplement Series* 249, p. 24. ISSN: 0067-0049. DOI: 10.3847/1538-4365/ab9cb0.

— (July 2023). "Validating Stellar Abundance Measurements from Multiresolution Spectroscopy". In: *ApJS* 267.1, p. 18. ISSN: 0067-0049. DOI: 10.3847/1538-4365/acd37b. (Visited on 07/13/2023).

Savino, A. et al. (May 2023). *The Hubble Space Telescope Survey of M31 Satellite Galaxies II. The Star Formation Histories of Ultra-Faint Dwarf Galaxies*. DOI: 10.48550/arXiv.2305.13360. (Visited on 07/25/2023).

Schlafly, Edward F. and Douglas P. Finkbeiner (Aug. 2011). "Measuring Reddening with Sloan Digital Sky Survey Stellar Spectra and Recalibrating SFD". In: *The Astrophysical Journal* 737, p. 103. ISSN: 0004-637X. DOI: 10.1088/0004-637X/737/2/103. (Visited on 08/30/2022).

Schönrich, Ralph and James Binney (June 2009). "Chemical Evolution with Radial Mixing". In: *Monthly Notices of the Royal Astronomical Society* 396, pp. 203–222. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2009.14750.x. (Visited on 03/09/2023).

Sharma, Sanjib, Michael R. Hayden, and Joss Bland-Hawthorn (Nov. 2021). "Chemical Enrichment and Radial Migration in the Galactic Disc - the Origin of the [αFe] Double Sequence". In: *Monthly Notices of the Royal Astronomical Society* 507, pp. 5882–5901. ISSN: 0035-8711. DOI: 10.1093/mnras/stab2015. (Visited on 06/07/2023).

Shetrone, M., D. Bizyaev, et al. (Dec. 2015). "The SDSS-III APOGEE Spectral Line List for H-band Spectroscopy". In: *The Astrophysical Journal Supplement Series* 221, p. 24. ISSN: 0067-0049. DOI: 10.1088/0067-0049/221/2/24. (Visited on 01/03/2020).

Shetrone, Matthew, Kim A. Venn, et al. (Feb. 2003). "VLT/UVES Abundances in Four Nearby Dwarf Spheroidal Galaxies. I. Nucleosynthesis and Abundance Ratios". In: *The Astronomical Journal* 125, p. 684. ISSN: ['0004-6256']. DOI: 10.1086/345966. (Visited on 11/21/2019).

Shetrone, Matthew D., Michael Bolte, and Peter B. Stetson (May 1998). "Keck HIRES Abundances in the Dwarf Spheroidal Galaxy Draco". In: *The Astronomical Journal* 115, p. 1888. DOI: 10.1086/300341. (Visited on 10/07/2019).

Shetrone, Matthew D., Patrick Côté, and W. L. W. Sargent (Feb. 2001). "Abundance Patterns in the Draco, Sextans, and Ursa Minor Dwarf Spheroidal Galaxies". In: *The Astrophysical Journal* 548, pp. 592–608. ISSN: 0004-637X. DOI: 10.1086/319022. (Visited on 01/21/2020).

Shetrone, Matthew D., Michael H. Siegel, et al. (Jan. 2009). "Chemical Abundances of the Leo II Dwarf Galaxy". In: *The Astronomical Journal* 137.1, p. 62. ISSN: 0004-6256. DOI: 10.1088/0004-6256/137/1/62. (Visited on 11/21/2019).

Simon, J. D., A. Drlica-Wagner, et al. (July 2015). "Stellar Kinematics and Metallicities in the Ultra-faint Dwarf Galaxy Reticulum II". In: *The Astrophysical Journal* 808, p. 95. ISSN: 0004-637X. DOI: 10.1088/0004-637X/808/1/95. (Visited on 01/09/2020).

Simon, J. D., T. S. Li, et al. (Mar. 2017). "Nearest Neighbor: The Low-mass Milky Way Satellite Tucana III". In: *The Astrophysical Journal* 838, p. 11. ISSN: 0004-637X. DOI: 10.3847/1538-4357/aa5be7. (Visited on 01/09/2020).

Simon, Joshua D. (Aug. 2019). "The Faintest Dwarf Galaxies". In: *Annual Review of Astronomy and Astrophysics* 57, pp. 375–415. ISSN: 0066-4146. DOI: 10.1146/annurev-astro-091918-104453.

Simon, Joshua D., Thomas M. Brown, et al. (Feb. 2021). "Eridanus II: A Fossil from Reionization with an Off-center Star Cluster". In: *The Astrophysical Journal* 908, p. 18. ISSN: 0004-637X. DOI: 10.3847/1538-4357/abd31b. (Visited on 03/09/2023).

Simon, Joshua D. and Marla Geha (Nov. 2007). "The Kinematics of the Ultra-faint Milky Way Satellites: Solving the Missing Satellite Problem". In: *The Astrophysical Journal* 670, p. 313. DOI: 10.1086/521816. (Visited on 10/07/2019).

Sitnova, T. M., L. I. Mashonkina, and T. A. Ryabchikova (Feb. 2013). "Influence of Departures from LTE on Oxygen Abundance Determination in the Atmospheres of A-K Stars". In: *Astronomy Letters* 39, pp. 126–140. ISSN: 0320-01081063-7737. DOI: 10.1134/S1063773713020084. (Visited on 07/21/2021).

Skidmore, Warren, TMT International Science Development Teams, and TMT Science Advisory Committee (Dec. 2015). "Thirty Meter Telescope Detailed Science Case: 2015".

In: *Research in Astronomy and Astrophysics* 15, p. 1945. ISSN: 1674-4527. DOI: 10.1088/1674-4527/15/12/001. (Visited on 09/13/2022).

Skrutskie, M. F. et al. (Feb. 2006). "The Two Micron All Sky Survey (2MASS)". In: *The Astronomical Journal* 131, pp. 1163–1183. ISSN: 0004-6256. DOI: 10.1086/498708. (Visited on 06/29/2021).

Slater, Colin T. et al. (June 2015). "A Deep Study of the Dwarf Satellites Andromeda XXVIII and Andromeda XXIX". In: *The Astrophysical Journal* 806, p. 230. ISSN: 0004-637X. DOI: 10.1088/0004-637X/806/2/230. (Visited on 01/24/2020).

Smiljanic, R. et al. (Oct. 2014). "The Gaia-ESO Survey: The Analysis of High-Resolution UVES Spectra of FGK-type Stars". In: *A&A* 570, A122. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201423937. (Visited on 03/02/2023).

Smith, Verne V. et al. (June 2021). "The APOGEE Data Release 16 Spectral Line List". In: *The Astronomical Journal* 161, p. 254. ISSN: 0004-6256. DOI: 10.3847/1538-3881/abefdc. (Visited on 09/13/2022).

Sneden, C. (May 1974). "Carbon and Nitrogen Abundances in Metal-Poor Stars." In: *The Astrophysical Journal* 189, pp. 493–507. ISSN: 0004-637X. DOI: 10.1086/152828. (Visited on 01/21/2020).

— (Sept. 1973). "The Nitrogen Abundance of the Very Metal-Poor Star HD 122563." In: *The Astrophysical Journal* 184, pp. 839–849. ISSN: 0004-637X. DOI: 10.1086/152374. (Visited on 01/21/2020).

Sneden, C., R. G. Gratton, and D. A. Crocker (June 1991). "Trends in Copper and Zinc Abundances for Disk and Halo Stars." In: *Astronomy and Astrophysics, Vol. 246, p. 354 (1991)* 246, p. 354. ISSN: 0004-6361. (Visited on 08/30/2022).

Sneden, C. and M. Parthasarathy (Apr. 1983). "The r- and s-Process Nuclei in the Early History of the Galaxy - HD 122563". In: *The Astrophysical Journal* 267, pp. 757–778. DOI: 10.1086/160913. (Visited on 01/01/2020).

Sneden, Christopher, John J. Cowan, et al. (July 2003). "The Extremely Metal-poor, Neutron Capture-rich Star CS 22892-052: A Comprehensive Abundance Analysis". In: *The Astrophysical Journal* 591, pp. 936–953. ISSN: 0004-637X. DOI: 10.1086/375491. (Visited on 01/10/2020).

Sneden, Christopher, Jennifer Johnson, et al. (June 2000). "Neutron-Capture Element Abundances in the Globular Cluster M15". In: *The Astrophysical Journal* 536, pp. L85–L88. ISSN: 0004-637X. DOI: 10.1086/312742. (Visited on 08/30/2022).

Sneden, Christopher, Robert P. Kraft, et al. (Nov. 1997). "Star-To-Star Abundance Variations Among Bright Giants in the Metal-Poor Globular Cluster M15". In: *The Astronomical Journal* 114, p. 1964. ISSN: 0004-6256. DOI: 10.1086/118618. (Visited on 08/30/2022).

Sneden, Christopher, Catherine A. Pilachowski, and Robert P. Kraft (Sept. 2000). "Barium and Sodium Abundances in the Globular Clusters M15 and M92". In: *The Astronomical Journal* 120, pp. 1351–1363. ISSN: 0004-6256. DOI: 10.1086/301509. (Visited on 08/30/2022).

Sobeck, Jennifer S., Inese I. Ivans, et al. (June 2006). "Manganese Abundances in Cluster and Field Stars". In: *The Astronomical Journal* 131, pp. 2949–2958. ISSN: 0004-6256. DOI: 10.1086/503106. (Visited on 08/30/2022).

Sobeck, Jennifer S., Robert P. Kraft, et al. (June 2011). "The Abundances of Neutron-capture Species in the Very Metal-poor Globular Cluster M15: A Uniform Analysis of Red Giant Branch and Red Horizontal Branch Stars". In: *The Astronomical Journal* 141, p. 175. ISSN: 0004-6256. DOI: 10.1088/0004-6256/141/6/175. (Visited on 08/30/2022).

Somerville, Rachel S. and Romeel Davé (Aug. 2015). "Physical Models of Galaxy Formation in a Cosmological Framework". In: *Annual Review of Astronomy and Astrophysics* 53, pp. 51–113. ISSN: 0066-4146. DOI: 10.1146/annurev-astro-082812-140951. (Visited on 10/19/2022).

Spencer, Meghin E. et al. (Feb. 2017). "A Multi-epoch Kinematic Study of the Remote Dwarf Spheroidal Galaxy Leo II". In: *The Astrophysical Journal* 836, p. 202. ISSN: ['0004-637X']. DOI: 10.3847/1538-4357/836/2/202. (Visited on 11/21/2019).

Spergel, D. et al. (Mar. 2015). *Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report*. DOI: 10.48550/arXiv.1503.03757. (Visited on 07/17/2023).

Spite, M. et al. (Sept. 2018). "A CEMP-no Star in the Ultra-Faint Dwarf Galaxy Pisces II". In: *Astronomy and Astrophysics* 617, A56. DOI: 10.1051/0004-6361/201833548. (Visited on 10/07/2019).

Starkenburg, E., V. Hill, E. Tolstoy, P. François, et al. (Jan. 2013). "The Extremely Low-Metallicity Tail of the Sculptor Dwarf Spheroidal Galaxy". In: *Astronomy and Astrophysics* 549, A88. DOI: 10.1051/0004-6361/201220349. (Visited on 10/07/2019).

Starkenburg, E., V. Hill, E. Tolstoy, J. I. González Hernández, et al. (Apr. 2010). "The NIR Ca Ii Triplet at Low Metallicity. Searching for Extremely Low-Metallicity Stars in Classical Dwarf Galaxies". In: *Astronomy and Astrophysics* 513, A34. ISSN: 0004-6361. DOI: 10.1051/0004-6361/200913759. (Visited on 12/18/2019).

Steinmetz, M. et al. (Oct. 2006). "The Radial Velocity Experiment (RAVE): First Data Release". In: *The Astronomical Journal* 132.4, p. 1645. ISSN: 0004-6256. DOI: 10.1086/506564. (Visited on 11/21/2019).

Straumit, Ilya et al. (May 2022). "Zeta-Payne: A Fully Automated Spectrum Analysis Algorithm for the Milky Way Mapper Program of the SDSS-V Survey". In: *The Astronomical Journal* 163, p. 236. ISSN: 0004-6256. DOI: 10.3847/1538-3881/ac5f49. (Visited on 01/31/2023).

Strömgren, B. (1940). "On the Chemical Composition of the Solar Atmosphere". In: *Festschrift Für Elis Strömgren*, p. 218. (Visited on 01/21/2020).

Suda, Takuma et al. (Oct. 2017). "Stellar Abundances for Galactic Archaeology Database. IV. Compilation of Stars in Dwarf Galaxies". In: *Publications of the Astronomical Society of Japan* 69, p. 76. ISSN: 0004-6264. DOI: 10.1093/pasj/psx059. (Visited on 01/20/2020).

Suntzeff, N. B. (Sept. 1981). "Carbon and Nitrogen Abundances in the Giant Stars of the Globular Clusters M3 and M13". In: *The Astrophysical Journal Supplement Series* 47, pp. 1–32. ISSN: 0067-0049. DOI: 10.1086/190750. (Visited on 01/21/2020).

Suntzeff, Nicholas B. et al. (Nov. 1993). "Spectroscopy of Giants in the Sextans Dwarf Spheroidal Galaxy". In: *The Astrophysical Journal* 418, p. 208. ISSN: 0004-637X. DOI: 10.1086/173383. (Visited on 01/21/2020).

Swan, J. et al. (Mar. 2016). "Ca II Triplet Spectroscopy of RGB Stars in NGC 6822: Kinematics and Metallicities". In: *Monthly Notices of the Royal Astronomical Society* 456, pp. 4315–4327. ISSN: 0035-8711. DOI: 10.1093/mnras/stv2774. (Visited on 01/24/2020).

Szentgyorgyi, Andrew, Daniel Baldwin, et al. (Aug. 2016). "The GMT-Consortium Large Earth Finder (G-CLEF): An Optical Echelle Spectrograph for the Giant Magellan Telescope (GMT)". In: *Ground-Based and Airborne Instrumentation for Astronomy VI*. Vol. 9908. International Society for Optics and Photonics, p. 990822. DOI: 10.1117/12.2233506. (Visited on 11/21/2019).

Szentgyorgyi, Andrew, Gabor Furesz, et al. (Sept. 2011). "Hectochelle: A Multiobject Optical Echelle Spectrograph for the MMT". In: *Publications of the Astronomical Society of the Pacific* 123.908, p. 1188. ISSN: 1538-3873. DOI: 10.1086/662209. (Visited on 10/07/2019).

Taibi, S. et al. (Sept. 2022). "Stellar Metallicity Gradients of Local Group Dwarf Galaxies". In: *Astronomy and Astrophysics* 665, A92. ISSN: 0004-6361. DOI: 10.1051/0004-6361/202243508. (Visited on 06/07/2023).

Takada, Masahiro et al. (Feb. 2014). "Extragalactic Science, Cosmology, and Galactic Archaeology with the Subaru Prime Focus Spectrograph". In: *Publications of the Astronomical Society of Japan* 66, R1. ISSN: 0004-6264. DOI: 10.1093/pasj/pst019. (Visited on 01/20/2020).

Tamura, Naoyuki et al. (July 2018). "Prime Focus Spectrograph (PFS) for the Subaru Telescope: Ongoing Integration and Future Plans". In: *Ground-Based and Airborne Instrumentation for Astronomy VII*. Vol. 10702. International Society for Optics and Photonics, p. 107021C. DOI: 10.1117/12.2311871. (Visited on 10/07/2019).

Taylor, William et al. (July 2018). "Rising MOONS: An Update on the VLT's next Multi-Object Spectrograph as It Begins to Grow". In: *Ground-Based and Airborne Instrumentation for Astronomy VII*. Vol. 10702. SPIE, pp. 421–440. DOI: 10.1117/12.2313403. (Visited on 09/13/2022).

ter Braak, Cajo J. F. and Jasper A. Vrugt (Dec. 2008). "Differential Evolution Markov Chain with Snooker Updater and Fewer Chains". In: *Stat Comput* 18.4, pp. 435–446. ISSN: 1573-1375. DOI: 10.1007/s11222-008-9104-9. (Visited on 08/30/2022).

Theler, R. et al. (Nov. 2019). "The Chemical Evolution of the Dwarf Spheroidal Galaxy Sextans". In: *arXiv e-prints*, arXiv:1911.08627. (Visited on 11/21/2019).

Ting, Yuan-Sen, Charlie Conroy, and Hans-Walter Rix (July 2016). "Accelerated Fitting of Stellar Spectra". In: *The Astrophysical Journal* 826.1, p. 83. DOI: 10.3847/0004-637X/826/1/83. (Visited on 10/04/2019).

Ting, Yuan-Sen, Charlie Conroy, Hans-Walter Rix, and Martin Asplund (June 2018). "Measuring Oxygen Abundances from Stellar Spectra without Oxygen Lines". In: *The Astrophysical Journal* 860.2, p. 159. DOI: 10.3847/1538-4357/aac6c9.

Ting, Yuan-Sen, Charlie Conroy, Hans-Walter Rix, and Phillip Cargile (July 2017). "Prospects for Measuring Abundances of >20 Elements with Low-resolution Stellar Spec-

tra". In: *The Astrophysical Journal* 843.1, p. 32. DOI: 10.3847/1538-4357/aa7688. (Visited on 09/25/2019).

Ting, Yuan-Sen, Charlie Conroy, Hans-Walter Rix, and Phillip Cargile (July 2019). "The Payne: Self-consistent Ab Initio Fitting of Stellar Spectra". In: *The Astrophysical Journal* 879.2, p. 69. DOI: 10.3847/1538-4357/ab2331.

Ting, Yuan-Sen, Keith Hawkins, and Hans-Walter Rix (May 2018). "A Large and Pristine Sample of Standard Candles across the Milky Way: ∼100,000 Red Clump Stars with 3% Contamination". In: *The Astrophysical Journal* 858, p. L7. ISSN: 0004-637X. DOI: 10.3847/2041-8213/aabf8e. (Visited on 01/30/2023).

Ting, Yuan-Sen, Hans-Walter Rix, et al. (Nov. 2017). "Measuring 14 Elemental Abundances with R = 1800 LAMOST Spectra". In: *The Astrophysical Journal Letters* 849, p. L9. ISSN: 0004-637X. DOI: 10.3847/2041-8213/aa921c. (Visited on 12/31/2019).

Ting, Yuan-Sen and David H. Weinberg (Mar. 2022). "How Many Elements Matter?" In: *The Astrophysical Journal* 927, p. 209. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ac5023. (Visited on 02/08/2023).

Tinsley, B. M. (1980). "Evolution of the Stars and Gas in Galaxies". In: *Fundamentals of Cosmic Physics* 5, pp. 287–388. (Visited on 01/21/2020).

Tolstoy, Eline, Vanessa Hill, and Monica Tosi (Sept. 2009). "Star-Formation Histories, Abundances, and Kinematics of Dwarf Galaxies in the Local Group". In: *Annual Review of Astronomy and Astrophysics* 47, p. 371. ISSN: ['0066-4146']. DOI: 10.1146/annurev-astro-082708-101650. (Visited on 11/21/2019).

Tolstoy, Eline, M. J. Irwin, et al. (Dec. 2004). "Two Distinct Ancient Components in the Sculptor Dwarf Spheroidal Galaxy: First Results from the Dwarf Abundances and Radial Velocities Team". In: *The Astrophysical Journal Letters* 617, pp. L119–L122. ISSN: 0004-637X. DOI: 10.1086/427388. (Visited on 01/21/2020).

Tolstoy, Eline, Kim A. Venn, et al. (Feb. 2003). "VLT/UVES Abundances in Four Nearby Dwarf Spheroidal Galaxies. II. Implications for Understanding Galaxy Evolution". In: *The Astronomical Journal* 125, pp. 707–726. ISSN: 0004-6256. DOI: 10.1086/345967. (Visited on 01/21/2020).

Tumlinson, Jason (Jan. 2010). "Chemical Evolution in Hierarchical Models of Cosmic Structure. II. The Formation of the Milky Way Stellar Halo and the Distribution of the Oldest Stars". In: *The Astrophysical Journal* 708, pp. 1398–1418. ISSN: 0004-637X. DOI: 10.1088/0004-637X/708/2/1398. (Visited on 07/25/2023).

Unsöld, A. (1938). "Physik Der Sternamospharen, MIT Besonderer Berucksichtigung Der Sonne". In: *Berlin, Verlag von Julius Springer, 1938.* (Visited on 01/21/2020).

— (1942). "Quantitative Analyse Des B O-Sternes τ Scorpii. II. Teil: Deutung Des Linienspektrums. Kosmische Häufigkeit Der Leichten Elemente. Elektronendruck Pe, Temperatur T Und Schwerebeschleunigung g in Der Atmosphäre von τ Scorpii. Mit 8 Abbildungen." In: *Zeitschrift fur Astrophysik* 21, p. 22. ISSN: 0372-8331. (Visited on 01/21/2020).

Vargas, Luis C., Marla Geha, et al. (Apr. 2013). "The Distribution of Alpha Elements in Ultra-faint Dwarf Galaxies". In: *The Astrophysical Journal* 767, p. 134. ISSN: 0004-637X. DOI: 10.1088/0004-637X/767/2/134. (Visited on 01/21/2020).

Vargas, Luis C., Karoline M. Gilbert, et al. (Dec. 2014). "[α/Fe] Abundances of Four Outer M31 Halo Stars". In: *The Astrophysical Journal* 797, p. L2. ISSN: ['0004-637X']. DOI: 10.1088/2041-8205/797/1/L2. (Visited on 11/22/2019).

Vasiliev, Eugene and Holger Baumgardt (Aug. 2021). "Gaia EDR3 View on Galactic Globular Clusters". In: *Monthly Notices of the Royal Astronomical Society* 505, pp. 5978–6002. ISSN: 0035-8711. DOI: 10.1093/mnras/stab1475. (Visited on 08/30/2022).

Venn, K. A., D. J. Lennon, et al. (Feb. 2001). "First Stellar Abundances in NGC 6822 from VLT-UVES and Keck-HIRES Spectroscopy". In: *The Astrophysical Journal* 547, p. 765. ISSN: ['0004-637X']. DOI: 10.1086/318424. (Visited on 11/21/2019).

Venn, K. A., E. Starkenburg, et al. (Apr. 2017). "Gemini/GRACES Spectroscopy of Stars in Tri II". In: *Monthly Notices of the Royal Astronomical Society* 466, pp. 3741–3752. ISSN: 0035-8711. DOI: 10.1093/mnras/stw3198. (Visited on 01/24/2020).

Venn, Kim A., Mike Irwin, et al. (Sept. 2004). "Stellar Chemical Signatures and Hierarchical Galaxy Formation". In: *The Astronomical Journal* 128, pp. 1177–1195. ISSN: 0004-6256. DOI: 10.1086/422734. (Visited on 01/21/2020).

Vernet, J. et al. (Dec. 2011). "X-Shooter, the New Wide Band Intermediate Resolution Spectrograph at the ESO Very Large Telescope". In: *Astronomy and Astrophysics* 536, A105. DOI: 10.1051/0004-6361/201117752. (Visited on 10/07/2019).

Vincenzo, F. et al. (July 2014). "Chemical Evolution of Classical and Ultra-Faint Dwarf Spheroidal Galaxies". In: *Monthly Notices of the Royal Astronomical Society* 441, pp. 2815–2830. ISSN: 0035-8711. DOI: 10.1093/mnras/stu710. (Visited on 10/19/2022).

Voggel, Karina et al. (Aug. 2016). "Probing the Boundary between Star Clusters and Dwarf Galaxies: A MUSE View on the Dynamics of Crater/Laevens I". In: *Monthly Notices of the Royal Astronomical Society* 460, p. 3384. DOI: 10.1093/mnras/stw1132. (Visited on 10/08/2019).

Vogt, S. S. et al. (June 1994). "HIRES: The High-Resolution Echelle Spectrometer on the Keck 10-m Telescope". In: *Instrumentation in Astronomy VIII* 2198, p. 362. DOI: 10.1117/12.176725. (Visited on 10/04/2019).

Walker, M. G., V. Belokurov, et al. (Apr. 2009). "Leo V: Spectroscopy of a Distant and Disturbed Satellite". In: *The Astrophysical Journal Letters* 694, pp. L144–L147. ISSN: 0004-637X. DOI: 10.1088/0004-637X/694/2/L144. (Visited on 01/21/2020).

Walker, Matthew G., Mario Mateo, and Edward W. Olszewski (Feb. 2009). "Stellar Velocities in the Carina, Fornax, Sculptor, and Sextans dSph Galaxies: Data From the Magellan/MMFS Survey". In: *The Astronomical Journal* 137, p. 3100. DOI: 10.1088/0004-6256/137/2/3100. (Visited on 10/08/2019).

Walker, Matthew G., Mario Mateo, Edward W. Olszewski, John I. Bailey, et al. (Aug. 2015). "Magellan/M2FS Spectroscopy of the Reticulum 2 Dwarf Spheroidal Galaxy". In: *The Astrophysical Journal* 808, p. 108. DOI: 10.1088/0004-637X/808/2/108. (Visited on 10/08/2019).

Walker, Matthew G., Mario Mateo, Edward W. Olszewski, Oleg Y. Gnedin, et al. (Sept. 2007). "Velocity Dispersion Profiles of Seven Dwarf Spheroidal Galaxies". In: *The Astrophysical Journal* 667.1, p. L53. DOI: 10.1086/521998. (Visited on 10/07/2019).

Walker, Matthew G., Mario Mateo, Edward W. Olszewski, Sergey Koposov, et al. (Mar. 2016). "Magellan/M2FS Spectroscopy of Tucana 2 and Grus 1". In: *The Astrophysical Journal* 819.1, p. 53. DOI: 10.3847/0004-637X/819/1/53. (Visited on 10/08/2019).

Walker, Matthew G., Edward W. Olszewski, and Mario Mateo (Apr. 2015). "Bayesian Analysis of Resolved Stellar Spectra: Application to MMT/Hectochelle Observations of the Draco Dwarf Spheroidal". In: *Monthly Notices of the Royal Astronomical Society* 448, p. 2717. DOI: 10.1093/mnras/stv099. (Visited on 10/08/2019).

Wallerstein, George, E. Myckky Leep, and J. B. Oke (May 1987). "High-Resolution CCD Spectra of Stars in Globular Clusters. III - M4, M13, and M22". In: *The Astronomical Journal* 93, pp. 1137–1143. ISSN: 0004-6256. DOI: 10.1086/114396. (Visited on 01/21/2020).

Wang, Rui, A. -Li Luo, Jian-Jun Chen, et al. (Mar. 2020). "SPCANet: Stellar Parameters and Chemical Abundances Network for LAMOST-II Medium Resolution Survey". In: *The Astrophysical Journal* 891, p. 23. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab6dea. (Visited on 07/17/2023).

Wang, Rui, A. -Li Luo, Shuo Zhang, et al. (June 2023). "Stellar Parameters and Chemical Abundances Estimated from LAMOST-II DR8 MRS Based on Cycle-StarNet". In: *The Astrophysical Journal Supplement Series* 266, p. 40. ISSN: 0067-0049. DOI: 10.3847/1538-4365/acce36. (Visited on 07/17/2023).

Wang, Yun (2010). "Forecasts of Dark Energy Constraints from Baryon Acoustic Oscillations". In: *Modern Physics Letters A* 25, pp. 3093–3113. ISSN: 0217-7323. DOI: 10.1142/S0217732310034316. (Visited on 12/30/2019).

Wang, Zixian, Michael R. Hayden, et al. (July 2022). "Reliable Stellar Abundances of Individual Stars with the MUSE Integral-Field Spectrograph". In: *Monthly Notices of the Royal Astronomical Society* 514, pp. 1034–1053. ISSN: 0035-8711. DOI: 10.1093/mnras/stac1425. (Visited on 08/30/2022).

Weinberg, David H., Brett H. Andrews, and Jenna Freudenburg (Mar. 2017). "Equilibrium and Sudden Events in Chemical Evolution". In: *The Astrophysical Journal* 837, p. 183. ISSN: 0004-637X. DOI: 10.3847/1538-4357/837/2/183. (Visited on 03/09/2023).

Weisz, Daniel R. et al. (July 2014). "The Star Formation Histories of Local Group Dwarf Galaxies. I. Hubble Space Telescope/Wide Field Planetary Camera 2 Observations". In: *The Astrophysical Journal* 789, p. 147. ISSN: 0004-637X. DOI: 10.1088/0004-637X/789/2/147. (Visited on 02/16/2023).

Wheeler, Adam et al. (July 2020). "Abundances in the Milky Way across Five Nucleosynthetic Channels from 4 Million LAMOST Stars". In: *The Astrophysical Journal* 898, p. 58. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab9a46. (Visited on 07/17/2023).

Worley, C. C. et al. (May 2013). "Ba and Eu Abundances in M 15 Giant Stars". In: *A&A* 553, A47. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201321097. (Visited on 03/02/2023).

Wright, K. O. (1948). "A Study of Line Intensities in the Spectra of Four Solar-Type Stars." In: *Publications of the Dominion Astrophysical Observatory Victoria* 8. ISSN: 0078-6950. (Visited on 01/21/2020).

Xiang, Maosheng, Hans-Walter Rix, et al. (June 2022). "Stellar Labels for Hot Stars from Low-Resolution Spectra. I. The HotPayne Method and Results for 330 000 Stars from LAMOST DR6". In: *Astronomy and Astrophysics* 662, A66. ISSN: 0004-6361. DOI: 10.1051/0004-6361/202141570. (Visited on 01/31/2023).

Xiang, Maosheng, Yuan-Sen Ting, et al. (Dec. 2019). "Abundance Estimates for 16 Elements in 6 Million Stars from LAMOST DR5 Low-Resolution Spectra". In: *The Astrophysical Journal Supplement Series* 245, p. 34. ISSN: 0067-0049. DOI: 10.3847/1538-4365/ab5364.

Xiong, Jian-Ping et al. (Nov. 2021). "The Relative Calibration of Radial Velocity for LAMOST Medium Resolution Stellar Spectra". In: *Research in Astronomy and Astrophysics* 21, p. 265. ISSN: 1674-4527. DOI: 10.1088/1674-4527/21/10/265. (Visited on 08/30/2022).

Yanny, Brian et al. (May 2009). "SEGUE: A Spectroscopic Survey of 240,000 Stars with g = 14-20". In: *The Astronomical Journal* 137.5, p. 4377. ISSN: 0004-6256. DOI: 10.1088/0004-6256/137/5/4377. (Visited on 11/21/2019).

Yoon, J. et al. (Oct. 2019). "Identification of a Group III CEMP-no Star in the Dwarf Spheroidal Galaxy Canes Venatici I". In: *arXiv e-prints*, arXiv:1910.10038. arXiv: 1910.10038. (Visited on 12/31/2019).

Zhang, Bo, Chao Liu, and Li-Cai Deng (Jan. 2020). "Deriving the Stellar Labels of LAMOST Spectra with the Stellar LAbel Machine (SLAM)". In: *The Astrophysical Journal Supplement Series* 246, p. 9. ISSN: 0067-0049. DOI: 10.3847/1538-4365/ab55ef. (Visited on 07/17/2023).

Zhu, Jun-Yan et al. (Mar. 2017). *Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks*. DOI: 10.48550/arXiv.1703.10593. (Visited on 07/18/2023).

Zoutendijk, Sebastiaan L., Jarle Brinchmann, Leindert A. Boogaard, et al. (Mar. 2020). "The MUSE-Faint Survey - I. Spectroscopic Evidence for a Star Cluster in Eridanus 2 and Constraints on MACHOs as a Constituent of Dark Matter". In: *A&A* 635, A107. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201936155. (Visited on 03/09/2023).

Zoutendijk, Sebastiaan L., Jarle Brinchmann, Nicolas F. Bouché, et al. (July 2021). "The MUSE-Faint Survey - II. The Dark-Matter Density Profile of the Ultra-Faint Dwarf Galaxy Eridanus 2". In: *A&A* 651, A80. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202040239. (Visited on 03/09/2023).