Planet Host Star Properties as Probes of Planet Formation and Evolution

Thesis by

Aida Behmard

In Partial Fulfillment of the Requirements for the
Degree of
Ph.D. in Planetary Science

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2023
Defended May 8, 2023

Aida Behmard
ORCID: 0000-0003-0012-9093

پنجره

یک پنجره برای دیدن
یک پنجره برای شنیدن
یک پنجره که مثل حلقه های چاهی
در انتهای خود به قلب زمین می رسد
و باز می شود به سوی وسعت این مهربانی مکرر ابی رنگ
یک پنجره که دستهای کوچک تنهایی را
از بخشش شبانهٔ عطر ستاره های کریم
سرشار میکند
و میشود از انجا
خورشید را به غربت گلهای شمعدانی مهمان کرد
یک پنجره برای من کافیست

A window for seeing
A window for hearing
A window like an unending well
that ends deep in the heart of the earth
and opens out into this expanse of recurring blue kindness
A window that overfills the tiny hands of loneliness
with its nightly gift: the perfume of generous stars
and thereof, one can invite the sun
to the alienation of geraniums
One window suffices me

– *Window*, Forough Farrokhzad

The high tower is a hundred feet tall
From here one's hand could pluck the stars
I do not dare to speak in a loud voice
I fear to disturb the people in heaven

夜宿山寺
唐（李白）
危楼高百尺，
手可摘星辰。
不敢高声语，
恐惊天上人。

– *Staying the Night at a Mountain Temple*, Li Bai

ACKNOWLEDGEMENTS

I am immensely grateful to my advisor, Heather Knutson, for her unwavering support and guidance throughout my time at Caltech. Heather's perspective and encouragement enabled me to navigate many research problems and quandaries, while bolstering my scientific curiosity and excitement for the future of our field. I consider myself very lucky to have her as a role model. She embodies the kind of mentor I hope to become one day.

Thank you to the other members of my thesis committee, Jim Fuller, Geoff Blake, and Konstantin Batygin, for their illuminating scientific discussions and helpful advice over the years. An especially big thanks to Jim Fuller and Jason Sevilla for their collaboration on two exciting projects that now contribute to the bedrock of this thesis. Thank you to Fei Dai, who also worked with me on several projects presented herein. Immense thanks to Melissa Ness, Emily Cunningham, and Megan Bedell, who mentored me throughout the final project presented in this thesis. Though they were all the way in NYC, Melissa, Emily, and Megan fostered an incredibly warm and supportive research environment that stretched across the country to my Pasadena apartment. I am so excited to continue working with them over the next few years.

Everything going on throughout the GPS and PMA Divisions is made possible by the amazing staff members. A huge thanks to Gita Kantiben, Margaret Carlos, Ulrika Terrones, Ruth Loisel, and Julie Lee. Thanks to all the staff members currently and previously at the Caltech Center for Teaching, Learning, and Outreach, and the Caltech Center for Inclusion and Diversity, particularly Kitty Cahalan, Mitch Aiken, and Erin-Kate Escobar. Thank you for running essential outreach and education programs that affirmed my identity as a scientist, and enabled me to do my part in making science accessible and inspiring for everyone.

Thank you to my friends in Planetary Science, those who came before me (Mike Wong, Elizabeth Bailey, Dana Anderson, Ellen Leask, Nathan Stein, Nancy Thomas, Steve Markham, Nicole Wallack), and after (Jess Spake, Maria Camarca, Mike Greklek-McKeon, Marguerite Epstein-Martin, Julie Inglis, Morgan Saidel, Zac Milby, Kim Paragas). Because of them, Planetary Science has always felt like a family to me. A special thanks to members of my own cohort, particularly Harriet Brettle, Qicheng Zhang, Yayaati Chachan, and Shreyas Vissapragada.

And thanks to my other friends across Caltech. A big thank you to Jim Barry for the art and fun times and wisdom. Also regarding wisdom, a big thanks to Lee Coleman. Thank you to my officemates and friends from the Blake group, Olivia Wilkins, Cam Buzard, and Kyle Virgil. Thank you to other GPS friends, Freya Morris, Emma Sosa, Olivia Pardo, Newton Nguyen, and Renée Wang. Thank you to Ryan Rubenzahl and Sarah Blunt, the best academic siblings imaginable. Thank you to all my friends who passed through "Planet Haus"−Harriet Brettle, Janosz Dewberry, Mia de los Reyes, Marguerite Epstein-Martin, Ben Idini, Jackie Pezzato, and Shreyas Vissapragada. A very special thanks to Shreyas and Jackie, who supported me when I could not carry myself. Thank you to my cat Interstellar Medium ("Izzy"), the best homie a girl could ask for. Thank you to Austin Chadwick, for everything.

This is for my family−my parents and Kian. Mom, dad, your achievements and sacrifices are what have enabled me to do everything I have ever done. Thank you for putting the nature documentaries on, for helping me with my math homework, and for never thinking that my dreams were too big. I owe you everything. I love you. Thank you.

ABSTRACT

Over the past three decades, we have discovered more than five thousand exoplanets with astonishingly diverse sizes, orbital configurations, and compositions. This broad range of characteristics underscores that most exoplanetary systems are unlike the Solar System, thus challenging our preconceptions of how planets form and evolve. The stars that exoplanets orbit hold clues to understanding their diversity. A star and its planets are born from a single cloud of gas and dust, so planet properties are intimately linked to those of the host star.

This thesis presents five studies that leverage observational techniques to characterize exoplanet host stars, and subsequently illuminate the mechanisms underpinning planet formation and evolution. In the first study, we modified a machine learning framework (*The Cannon*, Ness et al., 2015) to measure iron abundances in cool (<5200 K) dwarfs, the most common type of star in our galaxy. Our data-driven approach circumvents the need for physical models, making it ideal for cool star characterization. We tested our implementation with a sample of cool dwarfs that have well-measured iron abundances from empirical methods, and achieved iron abundance precisions of 0.08 dex. Iron is a proxy for the amount of protoplanetary disk solids that can contribute to planet cores, so well-constrained iron abundances are essential for vetting planet formation theories. Thus, our adaptation of *The Cannon* is a valuable tool for examining planet formation in cool star systems.

In the second study, we utilized a Bayesian statistical framework to identify stellar companions to planet host stars using astrometric data from the Gaia spacecraft, which tracks the movements of stars. We constrained the angle of alignment between the orbits of host stars and their companions, and found evidence for an excess of misaligned close-in giant planet systems that can shed light on dynamical processes (e.g., obliquity excitation) that shape planetary system architectures.

In the third study, we conducted a survey with the Keck High Resolution Echelle Spectrometer (HIRES, $R \approx 50{,}000$) to investigate planet engulfment signatures imprinted on host star compositions. We collected spectra for 36 planet host binary systems, and measured the abundances of 15 elements in each star. Binaries are excellent laboratories for probing planet engulfment because binary companions share the same molecular birth cloud, and are thus born chemically homogeneous. Planet-related processes can alter the birth compositions of each star in different

ways, so binary differential abundances may bear the fingerprint of engulfment. We did not identify any robust engulfment signatures, but did find a trend of increasing abundance differences between companions as a function of binary separation. This new result provides important context for investigations of planet formation that utilize binary star chemistry.

The fourth study is a natural extension of the third, in which we used the stellar evolution code `MESA` to quantify the strength and duration of refractory element enhancements that constitute engulfment signatures. We found that engulfment signatures are depleted by stellar interior mixing processes, with the depletion rate affected by the amount of planetary mass engulfed, the properties (i.e., mass and metallicity) of the engulfing star, and when the engulfment event occurred within the lifetime of the system. We found that engulfment signatures remain observable on short timescales ($<2-3$ Gyr) for all scenarios concerning solar-like stars, which may explain the lack of robust engulfment detections in our Keck-HIRES survey of binary systems. These results indicate that it may be difficult to observe engulfment signatures in solar-like stars that are several Gyr old.

In the fifth and final study, we extended our elemental abundance methodology to the 17th data release of the Apache Point Observatory Galactic Evolution Experiment (APOGEE DR17), which provides abundances for >650,000 stars that span the entirety of our galaxy. We used the APOGEE DR17 sample to quantify the abundance precision needed for discerning large-scale abundance patterns related to planet formation ($\lesssim$0.04 dex)−valuable information for investigating planet formation and evolution via stellar compositions on a galactic scale.

PUBLISHED CONTENT AND CONTRIBUTIONS

Behmard, Aida, Fei Dai, John M. Brewer, et al. (May 2023). "Planet engulfment detections are rare according to observations and stellar modelling". In: *Monthly Notices of the Royal Astronomical Society* 521.2, pp. 2969–2987. DOI: 10.1093/mnras/stad745. arXiv: 2210.12121 [astro-ph.EP].
A.B. conceived the project with input from coauthors, performed observations, reduced data, and wrote the manuscript.

Behmard, Aida, Melissa K. Ness, et al. (Apr. 2023). "Elemental Abundances of Kepler Objects of Interest in APOGEE DR17". In: *The Astronomical Journal* 165.4, 178, p. 178. DOI: 10.3847/1538-3881/acc32a.
A.B. conceived the project with input from coauthors and wrote the manuscript.

Behmard, Aida, Jason Sevilla, and Jim Fuller (Feb. 2023). "Planet engulfment signatures in twin stars". In: *Monthly Notices of the Royal Astronomical Society* 518.4, pp. 5465–5474. DOI: 10.1093/mnras/stac3435. arXiv: 2210.11679 [astro-ph.SR].
A.B. conceived the project with input from coauthors and wrote the manuscript.

Behmard, Aida, Fei Dai, and Andrew W. Howard (Apr. 2022). "Stellar Companions to TESS Objects of Interest: A Test of Planet-Companion Alignment". In: *The Astronomical Journal* 163.4, 160, p. 160. DOI: 10.3847/1538-3881/ac53a7. arXiv: 2202.01798 [astro-ph.EP].
A.B. conceived the project with input from coauthors and wrote the manuscript.

Behmard, Aida, Erik A. Petigura, and Andrew W. Howard (May 2019). "Data-driven Spectroscopy of Cool Stars at High Spectral Resolution". In: *The Astrophysical Journal* 876.1, 68, p. 68. DOI: 10.3847/1538-4357/ab14e0. arXiv: 1904.00094 [astro-ph.EP].
A.B. conceived the project with input from coauthors and wrote the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

INTRODUCTION

1.1 The Star-Planet Connection

Nearly three decades have passed since the first exoplanet discovery (Mayor and Queloz, 1995), and the count of confirmed exoplanets has now surpassed five thousand. This huge planetary census has shown us that most exoplanetary systems exhibit properties that diverge from those of the Solar System (Figure 1.1), indicating that the mechanisms underlying planet formation and evolution are much more complex than previously thought.

Because planets and the stars they orbit are born from the same cloud of gas and dust, the chemical and dynamical conditions of planet formation are intimately linked to the attributes of host stars. This planet-star connection is considered so essential that it has given rise to the phrase, "know thy star, know thy planet"[1], and has been leveraged in landmark studies to shed light on planet formation. For example, it is well established that planet occurrence increases around host stars with higher amounts of iron (e.g., Gonzalez 1997; Heiter and Luck 2003; Santos, Israelian, and Mayor 2004; Fischer and Valenti 2005; Johnson et al. 2010) (Figure 1.2). This "planet-metallicity" correlation grows stronger with increasing planet mass/radius (e.g., Sousa et al. 2008; Schlaufman and Laughlin 2011; Buchhave and Latham 2015; Wang, Fischer, Horch, et al. 2015). Because iron is regarded as a proxy for the solid content of protoplanetary disks, this indicates that large amounts of solid, rocky material facilitate rapid growth of massive planetary cores, which enables accretion of substantial gaseous envelopes. Thus, the planet-metallicity correlation constitutes strong evidence for the core accretion model of planet formation (Rice and Armitage, 2003; Ida and Lin, 2004; Alibert, Mordasini, and Benz, 2011; Mordasini et al., 2012; Maldonado et al., 2019).

Stellar abundances of elements beyond iron can also yield insights into planet formation by constraining formation locations and interior compositions. For example, comparing abundance ratios of volatile elements such as C/O in host star versus planetary atmospheres can reveal where planets formed in relation to relevant ice lines (e.g., CO, CO_2, H_2O) within protoplanetary disks (Öberg, Murray-Clay, and

[1] https://www.nasa.gov/feature/ames/kepler/know-thy-star-know-thy-planet

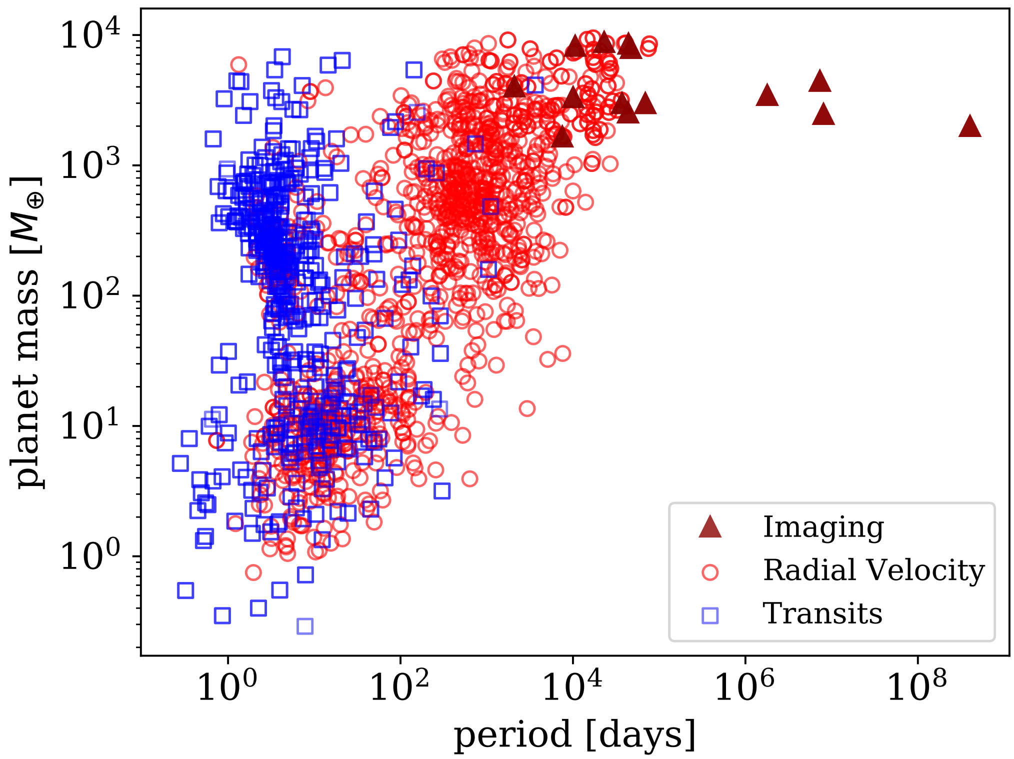


Figure 1.1: The distribution of masses and orbital periods of known planets from the Exoplanet Archive (Akeson et al., 2013).

Bergin, 2011). This can illuminate system dynamical histories; if planets are located far from their likely formation areas, they may have undergone migration. Host star C/O can also indicate if rocky planet interiors are composed of carbonates or silicates. In the latter case, refractory element abundance ratios such as Mg/Si can further characterize the silicate mineralogy, e.g., pyroxene versus olivine-dominated minerals (Grevesse, Asplund, and Sauval, 2007; Carter-Bond et al., 2012; Brewer and Fischer, 2017).

Planetary system dynamical evolution is also shaped by the presence of stellar companions to host stars. Stellar companions can perturb the structures of protoplanetary disks and induce misalignment between planetary orbits and the spin-axes of host stars (Wu, Murray, and Ramsahai, 2007; Batygin, 2012; Storch, Anderson, and Lai, 2014). Companions can also shorten disk lifetimes, thereby suppressing planet formation (e.g., Jang-Condell, Mugrauer, and Schmidt 2008). Once the disk has dissipated, companions can then excite the orbits of existing planets by driving them to high eccentricities, leading to planet migration or even ejection from the system (e.g., Kaib, Raymond, and Duncan 2013; Naoz 2016). Some observational studies

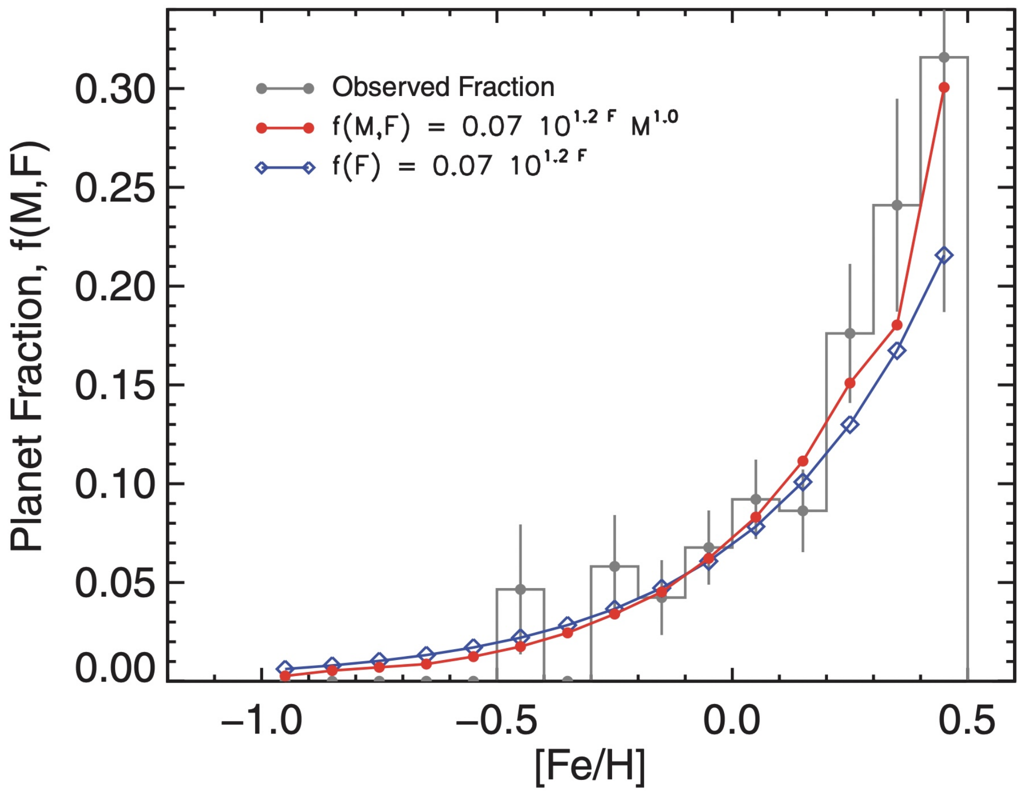


Figure 1.2: The planet fraction ($f = N_{\text{planets}}/N_{\text{stars}}$) as a function of host star metallicity for the Johnson et al. (2010) stellar sample (gray histogram). The red filled circles show the planet fraction for the masses and metallicities of the stars in each bin. The blue open diamonds show the best-fitting relationship between planet fraction and stellar metallicity assuming solar-mass stars. Figure from Johnson et al. (2010).

have reported that close binary companions are less common for planet-hosting stars in the *Kepler* field (e.g., Wang, Fischer, Xie, et al. 2014; Kraus et al. 2016), but others are less conclusive (e.g., Horch et al. 2014; Matson et al. 2018). Recently, Moe and Kratter (2021) combined various observational surveys to consistently characterize the stellar multiplicity of *Kepler* planet hosts. They found comparable occurrence rates for planets in wide (>200 AU separation) binaries and single star systems, and a much lower planet occurrence rate for close (<1 AU) binary systems. However, the effect of these stellar companions on planetary orbital evolution is less well-understood. For example, observations of misaligned hot Jupiters have not uncovered correlations between orbit misalignment and the presence of a directly imaged stellar companion (Ngo et al., 2015). Nearly half of solar-like stars harbor stellar companions (Raghavan et al., 2010), so understanding how companions affect planet formation is vital for fully understanding the star-planet connection.

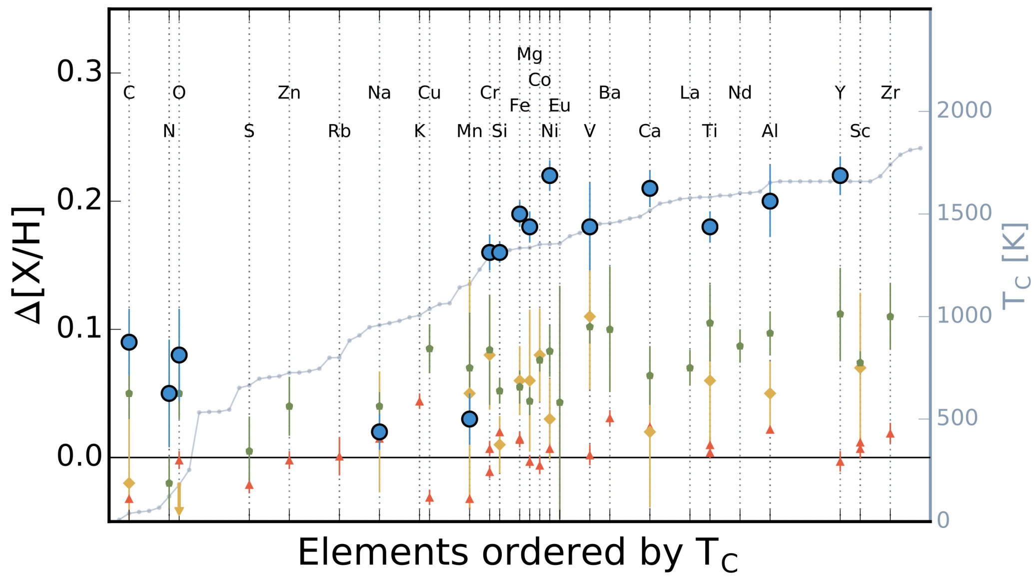


Figure 1.3: The differential abundance measurements between HD 240430 (Kronos) and HD 240429 (Krios) (blue circles). The abundance differences between the two stars (Kronos−Krios) indicate engulfment of rocky planetary material by the star Kronos. Figure from Oh et al. (2018).

Stellar companions are also vital for examining planet formation processes that alter host star chemistry. Host stars and their stellar companions, or "siblings", are born chemically homogeneous by virtue of sharing the same parent molecular cloud (e.g., De Silva, Freeman, Asplund, et al. 2007; De Silva, Freeman, and Bland-Hawthorn 2009; Bland-Hawthorn, Krumholz, and Freeman 2010). However, the process of planet formation can alter the birth compositions of each star in different ways, so differential abundances between stellar companions may bear the fingerprint of planet formation. For example, differential abundance patterns that decrease with element condensation temperature T_c (the temperature at which an element sublimates from the solid to gas phase) may indicate depletion of rocky, planet core-forming material. The Sun exhibits a subtle depletion of rocky material compared to >80% of 79 solar twins (Bedell et al., 2018); how this is related to the formation of our own Solar System (e.g., late-stage accretion of planetary material, accretion of dust-depleted gas after planet formation) has yet to be understood (e.g., Ramírez et al. 2011; Oh et al. 2018; Booth and Owen 2020). Conversely, abundance patterns that increase according to T_c indicate rocky enrichment, potentially from ingestion of planetary material. There are several observations of T_c enrichment patterns reported in the literature that appear to indicate planet engulfment, notably HD

240429−30 (Oh et al. 2018, Figure 1.3). In general, observational investigations of stars with similar bulk metal content find that their detailed elemental abundance patterns vary substantially (e.g., Adibekyan et al. 2012; Bensby, Feltzing, and Oey 2014; Brewer, Fischer, et al. 2016). This suggests that host star compositions are altered by a wide range of chemical/dynamical processes that sculpt the observed population of planetary systems.

1.2 High Resolution Spectroscopy

High resolution spectroscopy ($R > 10{,}000$) is one of the most powerful tools available for characterizing planet host stars. Spectra at these resolution levels are capable of resolving individual absorption lines for measuring stellar elemental abundances to high precision (typical uncertainties of $\lesssim$0.1 dex[2], e.g., Asplund et al. 2009; Brewer and Fischer 2018). There are various methods for deriving elemental abundances from stellar spectra. One common technique relies on modeling stellar spectra by constructing model stellar photospheres, then modeling the radiative transfer of photos as they pass through those photospheres. Spectral synthesis codes like `SME` (Valenti and Piskunov, 1996) and `MOOG` Sneden, 1973 can generate such synthetic spectra and match them to their observational counterparts. Unfortunately, stellar models become less reliable as stellar spectral types diverge from solar, particularly for cooler stars of type K4 and later whose spectra include molecular transitions with unconstrained quantum properties (e.g., Yee, Petigura, and von Braun 2017). Other commonly used methods involve empirical relations between equivalent width (EW) measurements of certain spectral lines and other stellar parameters, but so far they only measure the abundances of Fe and Ti (e.g., Mann et al. 2014; Newton, Charbonneau, Irwin, Berta-Thompson, et al. 2014; Newton, Charbonneau, Irwin, and Mann 2015; Veyette et al. 2017). Other methods include data-driven approaches, such as *The Cannon* (Ness et al., 2015), which uses training sets of spectra from stars with well-constrained elemental abundances to construct models that predict abundances from new spectra. Such data-driven methods are automated and not dependent on stellar models, making them valuable options for measuring abundances of large stellar samples that diverge from solar spectral types. All of these methods generally require stellar spectra with $R > 10{,}000$. Once measured, host star abundances can be used to constrain the chemistry of planet-forming environments, and/or elucidate planet formation processes that may alter host star chemistry beyond primordial

[2] In this work, we adopt the standard “bracket” chemical abundance notation [X/H] = $A(\mathrm{X})$ - $A(\mathrm{X})_{\odot}$, where $A(\mathrm{X}) = \log(n_{\mathrm{X}}/n_H) + 12$ and n_{X} is the number density of species X in the star’s photosphere.

compositions (e.g., Ramírez et al. 2011; Mack et al. 2014; Biazzo et al. 2015; Teske, Khanal, and Ramírez 2016; Saffe et al. 2017; Tucci Maia et al. 2019; Jofré et al. 2021).

Much of the work in this thesis makes use of the High Resolution Echelle Spectrometer (HIRES) on the Keck I 10 m telescope (Vogt et al., 1994). The Keck-HIRES resolution of $R \approx 50{,}000$ enables stellar abundance measurements at precisions of 0.01–0.05 dex for various refractory and volatile elements at typical signal-to-noise levels of SNR > 40/pixel (Brewer and Fischer, 2018). These abundance uncertainties are smaller than the upper end of chemical dispersion observed in stellar clusters and associations (∼0.05 dex, e.g., Bovy 2016), making it possible to tease out chemical signatures of planet formation in host stars from comparison with their stellar companions (i.e., "siblings"). Surveys carried out with Keck-HIRES can be used to individually target stars for measuring stellar parameters and elemental abundances (see Chapters 1 and 3). However, carrying out observations for large stellar samples is not feasible with instruments like Keck-HIRES, and instead require large-scale spectroscopic surveys.

1.3 Large Stellar Surveys

Luckily, large-scale stellar surveys are becoming increasingly common. These surveys are motivated by studies of galactic evolution which rely heavily upon understanding the nature of our own galaxy. Thus, they are extensive and homogeneous chemodynamical surveys that span all regions of the Milky Way. They include the Large Sky Area Multi-Object Fibre Spectroscopic Telescope (LAMOST, Zhao et al. 2012) survey, the GALactic Archaeology with HERMES (GALAH, Buder et al. 2018) survey, and the Sloan Digital Sky Survey's Apache Point Observatory Galactic Evolution Experiment (APOGEE, Majewski et al. 2017), to name a few. Because these large-scale spectroscopic surveys provide high-precision elemental abundances for large stellar samples, they are valuable tools for examining the chemistry of planet formation. For example, the LAMOST-*Kepler* project seeks to observe as many stars as possible in the *Kepler* field to construct a large, unbiased sample that can be used for population studies of planet host stars (De Cat et al., 2015; Ren et al., 2016; Fu et al., 2020). Dong et al. (2018) used the LAMOST-*Kepler* project to examine the distribution of short-period (P < 10 days) *Kepler* planets as a function of host star metallicity, and discovered a population of close-in Neptune-sized planets that, like hot Jupiters, occur more frequently around metal-rich hosts. The LAMOST-*Kepler* catalog has since been expanded to include data from the *Gaia* mission, and was used by Chen, Xie, Zhou, Dong, et al. (2021) to examine planet multiplicity across different galactic populations. They found that the fraction of thin-to-thick disk stars (classified via kinematic properties and metallicity) increases/decreases with transiting planet multiplicity. In a follow-up paper, Chen, Xie, Zhou, Yang, et al. (2022) investigated the "radius valley" that separates super-Earths and sub-Neptunes (Fulton et al., 2017), and found that valley morphology varies as a function of iron and α-element abundance (e.g., Mg, Si, Ca, Ti). The GALAH survey provides abundances for elements beyond iron, measured with the spectral synthesis pipeline `SME`. Clark, Clerté, et al. (2021) and Clark, Wright, et al. (2022) used GALAH to examine the detailed chemical properties of TESS planet host stars, and tag planet hosts to thin vs. thick disk populations via their chemokinetic properties.

The APOGEE survey presents the largest (>650,000 stars, Abdurro'uf et al. 2022) spectroscopic catalog yet that includes abundance measurements for a wide set of elements beyond iron. APOGEE provides high resolution ($R \approx 22{,}500$) and high signal-to-noise ratio (S/N > 100) spectroscopy, with abundances measured by its accompanying Automated Stellar Parameters and Chemical Abundances Pipeline (ASPCAP, García Pérez et al. 2016). Wilson, Teske, et al. (2018) used APOGEE DR14 to examine the metallicity distribution of *Kepler* Objects of Interest (KOIs), and found a correlation between orbital period and host star [Fe/H] characterized by a critical period ($P = 8.3 \pm 0.1$ days) below which small planets systematically orbit more metal-rich stars. In a follow-up paper, Wilson, Cañas, et al. (2022) examined correlations between planet occurrence and APOGEE abundances of elements beyond iron, and found that higher abundances of any refractory elements correspond to enhanced planet occurrence, but the correlations weaken as a function of increasing orbital period and decreasing planet radii. Additionally, as described in Chapter 6, Nibauer et al. (2021) used APOGEE DR16 to examine abundance trends of solar analog stars, and found evidence for two distinct populations of stars−one with abundance vs. T_c trends similar to that of the Sun, and another with abundance trends that increase with T_c. Such trends may be linked to planet formation, indicating that large stellar surveys like APOGEE could provide a means to examine chemical abundance patterns of planet hosts stars across the Milky Way. We utilized APOGEE DR17 in the final project of this thesis, presented in Chapter 6.

References

Abdurro'uf et al. (Apr. 2022). "The Seventeenth Data Release of the Sloan Digital Sky Surveys: Complete Release of MaNGA, MaStar, and APOGEE-2 Data". In: *The Astrophysical Journal Supplement Series* 259.2, 35, p. 35. DOI: 10.3847/1538-4365/ac4414. arXiv: 2112.02026 [astro-ph.GA].

Adibekyan, V. Zh. et al. (Nov. 2012). "Exploring the α-enhancement of metal-poor planet-hosting stars. The Kepler and HARPS samples". In: *Astronomy and Astrophysics* 547, A36, A36. DOI: 10.1051/0004-6361/201220167. arXiv: 1209.6272 [astro-ph.EP].

Akeson, R. L. et al. (Aug. 2013). "The NASA Exoplanet Archive: Data and Tools for Exoplanet Research". In: *Publications of the Astronomical Society of the Pacific* 125.930, p. 989. DOI: 10.1086/672273. arXiv: 1307.2944 [astro-ph.IM].

Alibert, Y., C. Mordasini, and W. Benz (Feb. 2011). "Extrasolar planet population synthesis. III. Formation of planets around stars of different masses". In: *Astronomy and Astrophysics* 526, A63, A63. DOI: 10.1051/0004-6361/201014760. arXiv: 1101.0513 [astro-ph.EP].

Asplund, Martin et al. (Sept. 2009). "The Chemical Composition of the Sun". In: *Annual Review of Astronomy and Astrophysics* 47.1, pp. 481–522. DOI: 10.1146/annurev.astro.46.060407.145222. arXiv: 0909.0948 [astro-ph.SR].

Batygin, Konstantin (Nov. 2012). "A primordial origin for misalignments between stellar spin axes and planetary orbits". In: *Nature* 491.7424, pp. 418–420. DOI: 10.1038/nature11560.

Bedell, Megan et al. (Sept. 2018). "The Chemical Homogeneity of Sun-like Stars in the Solar Neighborhood". In: *The Astrophysical Journal* 865.1, 68, p. 68. DOI: 10.3847/1538-4357/aad908. arXiv: 1802.02576 [astro-ph.SR].

Bensby, T., S. Feltzing, and M. S. Oey (Feb. 2014). "Exploring the Milky Way stellar disk. A detailed elemental abundance study of 714 F and G dwarf stars in the solar neighbourhood". In: *Astronomy and Astrophysics* 562, A71, A71. DOI: 10.1051/0004-6361/201322631. arXiv: 1309.2631 [astro-ph.GA].

Biazzo, K. et al. (Nov. 2015). "The GAPS programme with HARPS-N at TNG. X. Differential abundances in the XO-2 planet-hosting binary". In: *Astronomy and Astrophysics* 583, A135, A135. DOI: 10.1051/0004-6361/201526375. arXiv: 1506.01614 [astro-ph.SR].

Bland-Hawthorn, Joss, Mark R. Krumholz, and Ken Freeman (Apr. 2010). "The Long-term Evolution of the Galactic Disk Traced by Dissolving Star Clusters". In: *The Astrophysical Journal* 713.1, pp. 166–179. DOI: 10.1088/0004-637X/713/1/166. arXiv: 1002.4357 [astro-ph.GA].

Booth, Richard A. and James E. Owen (Apr. 2020). "Fingerprints of giant planets in the composition of solar twins". In: *Monthly Notices of The Royal Astronomical*

Society 493.4, pp. 5079–5088. DOI: 10.1093/mnras/staa578. arXiv: 2002.11135 [astro-ph.EP].

Bovy, Jo (Jan. 2016). “The Chemical Homogeneity of Open Clusters”. In: *The Astrophysical Journal* 817.1, 49, p. 49. DOI: 10.3847/0004-637X/817/1/49. arXiv: 1510.06745 [astro-ph.GA].

Brewer, J. M. and D. A. Fischer (Aug. 2018). “Spectral Properties of Cool Stars: Extended Abundance Analysis of Kepler Objects of Interest”. In: *The Astrophysical Journal Supplement Series* 237, 38, p. 38. DOI: 10.3847/1538-4365/aad501. arXiv: 1804.00673 [astro-ph.SR].

Brewer, J. M., D. A. Fischer, et al. (Aug. 2016). “Spectral Properties of Cool Stars: Extended Abundance Analysis of 1,617 Planet-search Stars”. In: *The Astrophysical Journal Supplement Series* 225, 32, p. 32. DOI: 10.3847/0067-0049/225/2/32. arXiv: 1606.07929 [astro-ph.SR].

Brewer, John M. and Debra A. Fischer (May 2017). “Erratum: “C/O and Mg/Si Ratios of Stars in the Solar Neighborhood” (<A href=“http://doi.org/10.3847/0004-637x/831/1/20”>2016, ApJ, 831, 20</A>)”. In: *The Astrophysical Journal* 840.2, 121, p. 121. DOI: 10.3847/1538-4357/aa6d53. arXiv: 1608.06286 [astro-ph.SR].

Buchhave, Lars A. and David W. Latham (Aug. 2015). “The Metallicities of Stars with and without Transiting Planets”. In: *The Astrophysical Journal* 808.2, 187, p. 187. DOI: 10.1088/0004-637X/808/2/187. arXiv: 1507.03557 [astro-ph.EP].

Buder, Sven et al. (Aug. 2018). “The GALAH Survey: second data release”. In: *Monthly Notices of the Royal Astronomical Society* 478.4, pp. 4513–4552. DOI: 10.1093/mnras/sty1281. arXiv: 1804.06041 [astro-ph.SR].

Carter-Bond, Jade C. et al. (Mar. 2012). “Low Mg/Si Planetary Host Stars and Their Mg-depleted Terrestrial Planets”. In: *The Astrophysical Journal Letters* 747.1, L2, p. L2. DOI: 10.1088/2041-8205/747/1/L2. arXiv: 1201.1939 [astro-ph.EP].

Chen, Di-Chang, Ji-Wei Xie, Ji-Lin Zhou, Subo Dong, et al. (Mar. 2021). “Planets Across Space and Time (PAST). I. Characterizing the Memberships of Galactic Components and Stellar Ages: Revisiting the Kinematic Methods and Applying to Planet Host Stars”. In: *The Astrophysical Journal* 909.2, 115, p. 115. DOI: 10.3847/1538-4357/abd5be. arXiv: 2102.09424 [astro-ph.EP].

Chen, Di-Chang, Ji-Wei Xie, Ji-Lin. Zhou, Jia-Yi Yang, et al. (June 2022). “Planets Across Space and Time (PAST). III. Morphology of the Planetary Radius Valley as a Function of Stellar Age and Metallicity in the Galactic Context Revealed by the LAMOST-Gaia-Kepler Sample”. In: *The Astronomical Journal* 163.6, 249, p. 249. DOI: 10.3847/1538-3881/ac641f. arXiv: 2204.01940 [astro-ph.EP].

Clark, Jake T., Mathieu Clerté, et al. (July 2021). "The GALAH Survey: using galactic archaeology to refine our knowledge of TESS target stars". In: *Monthly Notices of the Royal Astronomical Society* 504.4, pp. 4968–4989. DOI: 10.1093/mnras/stab1052. arXiv: 2008.05372 [astro-ph.EP].

Clark, Jake T., Duncan J. Wright, et al. (Feb. 2022). "The GALAH Survey: improving our understanding of confirmed and candidate planetary systems with large stellar surveys". In: *Monthly Notices of the Royal Astronomical Society* 510.2, pp. 2041–2060. DOI: 10.1093/mnras/stab3498. arXiv: 2111.14883 [astro-ph.EP].

De Cat, P. et al. (Sept. 2015). "Lamost Observations in the Kepler Field. I. Database of Low-resolution Spectra". In: *The Astrophysical Journal Supplement Series* 220.1, 19, p. 19. DOI: 10.1088/0067-0049/220/1/19. arXiv: 1508.06391 [astro-ph.SR].

De Silva, G. M., K. C. Freeman, M. Asplund, et al. (Mar. 2007). "Chemical Homogeneity in Collinder 261 and Implications for Chemical Tagging". In: *The Astronomical Journal* 133.3, pp. 1161–1175. DOI: 10.1086/511182. arXiv: astro-ph/0611832 [astro-ph].

De Silva, G. M., K. C. Freeman, and J. Bland-Hawthorn (Apr. 2009). "Reconstructing Fossil Sub-structures of the Galactic Disk: Clues from Abundance Patterns of Old Open Clusters and Moving Groups". In: *Publications of the Astronomical Society of the Pacific* 26, pp. 11–16. DOI: 10.1071/AS08019. arXiv: 0810.2287.

Dong, Subo et al. (Jan. 2018). "LAMOST telescope reveals that Neptunian cousins of hot Jupiters are mostly single offspring of stars that are rich in heavy elements". In: *Proceedings of the National Academy of Science* 115.2, pp. 266–271. DOI: 10.1073/pnas.1711406115. arXiv: 1706.07807 [astro-ph.EP].

Fischer, Debra A. and Jeff Valenti (Apr. 2005). "The Planet-Metallicity Correlation". In: *The Astrophysical Journal* 622.2, pp. 1102–1117. DOI: 10.1086/428383. URL: https://doi.org/10.1086/428383.

Fu, Jian-Ning et al. (Oct. 2020). "Overview of the LAMOST-Kepler project". In: *Research in Astronomy and Astrophysics* 20.10, 167, p. 167. DOI: 10.1088/1674-4527/20/10/167. arXiv: 2008.10776 [astro-ph.SR].

Fulton, Benjamin J. et al. (Sept. 2017). "The California-Kepler Survey. III. A Gap in the Radius Distribution of Small Planets". In: *The Astronomical Journal* 154.3, 109, p. 109. DOI: 10.3847/1538-3881/aa80eb. arXiv: 1703.10375 [astro-ph.EP].

García Pérez, Ana E. et al. (2016). "ASPCAP: The APOGEE Stellar Parameter and Chemical Abundances Pipeline". In: *The Astronomical Journal*.

Gonzalez, Guillermo (Feb. 1997). "The stellar metallicity-giant planet connection". In: *Monthly Notices of The royal* 285.2, pp. 403–412. DOI: 10.1093/mnras/285.2.403.

Grevesse, N., M. Asplund, and A. J. Sauval (June 2007). "The Solar Chemical Composition". In: *Space Science Reviews* 130.1-4, pp. 105–114. DOI: 10.1007/s11214-007-9173-7.

Heiter, U. and R. E. Luck (Oct. 2003). "Abundance Analysis of Planetary Host Stars. I. Differential Iron Abundances". In: *The Astronomical Journal* 126.4, pp. 2015–2036. DOI: 10.1086/378366. arXiv: astro-ph/0307321 [astro-ph].

Horch, Elliott P. et al. (Nov. 2014). "Most Sub-arcsecond Companions of Kepler Exoplanet Candidate Host Stars are Gravitationally Bound". In: *The Astrophysical Journal* 795.1, 60, p. 60. DOI: 10.1088/0004-637X/795/1/60. arXiv: 1409.1249 [astro-ph.SR].

Ida, Shigeru and D. N. C. Lin (Nov. 2004). "Toward a Deterministic Model of Planetary Formation. II. The Formation and Retention of Gas Giant Planets around Stars with a Range of Metallicities". In: *The Astrophysical Journal* 616.1, pp. 567–572. DOI: 10.1086/424830. URL: https://doi.org/10.1086/424830.

Jang-Condell, H., M. Mugrauer, and T. Schmidt (Aug. 2008). "Disk Truncation and Planet Formation in γ Cephei". In: *The Astrophysical Journal Letters* 683.2, p. L191. DOI: 10.1086/591791. arXiv: 0807.4472 [astro-ph].

Jofré, Emiliano et al. (Dec. 2021). "The Peculiar Chemical Pattern of the WASP-160 Binary System: Signatures of Planetary Formation and Evolution?" In: *The Astronomical Journal* 162.6, 291, p. 291. DOI: 10.3847/1538-3881/ac25ef. arXiv: 2109.04590 [astro-ph.EP].

Johnson, John Asher et al. (Aug. 2010). "Giant Planet Occurrence in the Stellar Mass-Metallicity Plane". In: *Publications of the Astronomical Society of the Pacific* 122.894, p. 905. DOI: 10.1086/655775. arXiv: 1005.3084 [astro-ph.EP].

Kaib, Nathan A., Sean N. Raymond, and Martin Duncan (Jan. 2013). "Planetary system disruption by Galactic perturbations to wide binary stars". In: *Nature* 493.7432, pp. 381–384. DOI: 10.1038/nature11780. arXiv: 1301.3145 [astro-ph.EP].

Kraus, Adam L. et al. (July 2016). "The Impact of Stellar Multiplicity on Planetary Systems. I. The Ruinous Influence of Close Binary Companions". In: *The Astronomical Journal* 152.1, 8, p. 8. DOI: 10.3847/0004-6256/152/1/8. arXiv: 1604.05744 [astro-ph.EP].

Mack Claude E., III et al. (June 2014). "Detailed Abundances of Planet-hosting Wide Binaries. I. Did Planet Formation Imprint Chemical Signatures in the Atmospheres of HD 20782/81?" In: *The Astrophysical Journal* 787.2, 98, p. 98. DOI: 10.1088/0004-637X/787/2/98. arXiv: 1404.1967 [astro-ph.EP].

Majewski, Steven R. et al. (Sept. 2017). "The Apache Point Observatory Galactic Evolution Experiment (APOGEE)". In: *The Astronomical Journal* 154.3, 94, p. 94. DOI: 10.3847/1538-3881/aa784d. arXiv: 1509.05420 [astro-ph.IM].

Maldonado, J. et al. (Apr. 2019). "Connecting substellar and stellar formation: the role of the host star's metallicity". In: *Astronomy and Astrophysics* 624, A94, A94. DOI: 10.1051/0004-6361/201833827. arXiv: 1903.01141 [astro-ph.SR].

Mann, Andrew W. et al. (June 2014). "Prospecting in Ultracool Dwarfs: Measuring the Metallicities of Mid- and Late-M Dwarfs". In: *The Astronomical Journal* 147.6, 160, p. 160. DOI: 10.1088/0004-6256/147/6/160. arXiv: 1403.5560 [astro-ph.SR].

Matson, Rachel et al. (Jan. 2018). "Stellar Companions of Exoplanet Host Stars in K2". In: *American Astronomical Society Meeting Abstracts #231*. Vol. 231. American Astronomical Society Meeting Abstracts, 109.02, p. 109.02.

Mayor, Michel and Didier Queloz (Nov. 1995). "A Jupiter-mass companion to a solar-type star". In: *Nature* 378.6555, pp. 355–359. DOI: 10.1038/378355a0.

Moe, Maxwell and Kaitlin M. Kratter (Nov. 2021). "Impact of binary stars on planet statistics - I. Planet occurrence rates and trends with stellar mass". In: *Monthly Notices of the Royal Astronomical Society* 507.3, pp. 3593–3611. DOI: 10.1093/mnras/stab2328. arXiv: 1912.01699 [astro-ph.EP].

Mordasini, C. et al. (Nov. 2012). "Characterization of exoplanets from their formation. II. The planetary mass-radius relationship". In: *Astronomy and Astrophysics* 547, A112, A112. DOI: 10.1051/0004-6361/201118464. arXiv: 1206.3303 [astro-ph.EP].

Naoz, Smadar (Oct. 2016). "Dynamical Effects of Stellar Companions". In: *IAU Focus Meeting* 29A, pp. 65–70. DOI: 10.1017/S1743921316002416.

Ness, M. et al. (July 2015). "The Cannon: A data-driven approach to Stellar Label Determination". In: *The Astrophysical Journal* 808, 16, p. 16. DOI: 10.1088/0004-637X/808/1/16. arXiv: 1501.07604 [astro-ph.SR].

Newton, E. R., D. Charbonneau, J. Irwin, Z. K. Berta-Thompson, et al. (Jan. 2014). "Near-infrared Metallicities, Radial Velocities, and Spectral Types for 447 Nearby M Dwarfs". In: *The Astronomical Journal* 147, 20, p. 20. DOI: 10.1088/0004-6256/147/1/20. arXiv: 1310.1087 [astro-ph.SR].

Newton, E. R., D. Charbonneau, J. Irwin, and A. W. Mann (Feb. 2015). "An Empirical Calibration to Estimate Cool Dwarf Fundamental Parameters from H-band Spectra". In: *The Astrophysical Journal* 800, 85, p. 85. DOI: 10.1088/0004-637X/800/2/85. arXiv: 1412.2758 [astro-ph.SR].

Ngo, Henry et al. (Feb. 2015). "Friends of Hot Jupiters. II. No Correspondence between Hot-jupiter Spin-Orbit Misalignment and the Incidence of Directly Imaged Stellar Companions". In: *The Astrophysical Journal* 800.2, 138, p. 138. DOI: 10.1088/0004-637X/800/2/138. arXiv: 1501.00013 [astro-ph.EP].

Nibauer, Jacob et al. (Feb. 2021). “Statistics of the Chemical Composition of Solar Analog Stars and Links to Planet Formation”. In: *The Astrophysical Journal* 907.2, 116, p. 116. DOI: 10.3847/1538-4357/abd0f1. arXiv: 2010.07241 [astro-ph.SR].

Öberg, Karin I., Ruth Murray-Clay, and Edwin A. Bergin (Dec. 2011). “The Effects of Snowlines on C/O in Planetary Atmospheres”. In: *The Astrophysical Journal Letters* 743.1, L16, p. L16. DOI: 10.1088/2041-8205/743/1/L16. arXiv: 1110.5567 [astro-ph.GA].

Oh, Semyeong et al. (Feb. 2018). “Kronos and Krios: Evidence for Accretion of a Massive, Rocky Planetary System in a Comoving Pair of Solar-type Stars”. In: *The Astrophysical Journal* 854.2, 138, p. 138. DOI: 10.3847/1538-4357/aaab4d. arXiv: 1709.05344 [astro-ph.SR].

Raghavan, Deepak et al. (Sept. 2010). “A Survey of Stellar Families: Multiplicity of Solar-type Stars”. In: *The Astrophysical Journal Supplement Series* 190.1, pp. 1–42. DOI: 10.1088/0067-0049/190/1/1. arXiv: 1007.0414 [astro-ph.SR].

Ramírez, I. et al. (Oct. 2011). “Elemental Abundance Differences in the 16 Cygni Binary System: A Signature of Gas Giant Planet Formation?” In: *The Astrophysical Journal* 740.2, 76, p. 76. DOI: 10.1088/0004-637X/740/2/76. arXiv: 1107.5814 [astro-ph.SR].

Ren, Anbing et al. (Aug. 2016). “LAMOST Observations in the Kepler Field. Analysis of the Stellar Parameters Measured with LASP Based on Low-resolution Spectra”. In: *The Astrophysical Journal Supplement Series* 225.2, 28, p. 28. DOI: 10.3847/0067-0049/225/2/28. arXiv: 1607.01453 [astro-ph.SR].

Rice, W. K. M. and Philip J. Armitage (Oct. 2003). “On the Formation Timescale and Core Masses of Gas Giant Planets”. In: *The Astrophysical Journal* 598.1, pp. L55–L58. DOI: 10.1086/380390. URL: https://doi.org/10.1086/380390.

Saffe, C. et al. (July 2017). “Signatures of rocky planet engulfment in HAT-P-4. Implications for chemical tagging studies”. In: *Astronomy and Astrophysics* 604, L4, p. L4. DOI: 10.1051/0004-6361/201731430. arXiv: 1707.02180 [astro-ph.SR].

Santos, N. C., G. Israelian, and M. Mayor (Mar. 2004). “Spectroscopic [Fe/H] for 98 extra-solar planet-host stars. Exploring the probability of planet formation”. In: *Astronomy and Astrophysics* 415, pp. 1153–1166. DOI: 10.1051/0004-6361:20034469. arXiv: astro-ph/0311541 [astro-ph].

Schlaufman, Kevin C. and Gregory Laughlin (Sept. 2011). “Kepler Exoplanet Candidate Host Stars Are Preferentially Metal Rich”. In: *The Astrophysical Journal* 738.2, 177, p. 177. DOI: 10.1088/0004-637X/738/2/177. arXiv: 1106.6043 [astro-ph.EP].

Sneden, C. A. (1973). “Carbon and Nitrogen Abundances in Metal-Poor Stars.” PhD thesis. THE UNIVERSITY OF TEXAS AT AUSTIN.

Sousa, S. G. et al. (Aug. 2008). "Spectroscopic parameters for 451 stars in the HARPS GTO planet search program. Stellar [Fe/H] and the frequency of exo-Neptunes". In: *Astronomy and Astrophysics* 487.1, pp. 373–381. DOI: 10.1051/0004-6361:200809698. arXiv: 0805.4826 [astro-ph].

Storch, Natalia I., Kassandra R. Anderson, and Dong Lai (Sept. 2014). "Chaotic dynamics of stellar spin in binaries and the production of misaligned hot Jupiters". In: *Science* 345.6202, pp. 1317–1321. DOI: 10.1126/science.1254358. arXiv: 1409.3247 [astro-ph.EP].

Teske, Johanna K., Sandhya Khanal, and Ivan Ramírez (Mar. 2016). "The Curious Case of Elemental Abundance Differences in the Dual Hot Jupiter Hosts WASP-94A and B". In: *The Astrophysical Journal* 819.1, 19, p. 19. DOI: 10.3847/0004-637X/819/1/19. arXiv: 1601.01731 [astro-ph.SR].

Tucci Maia, M. et al. (Aug. 2019). "Revisiting the 16 Cygni planet host at unprecedented precision and exploring automated tools for precise abundances". In: *Astronomy and Astrophysics* 628, A126, A126. DOI: 10.1051/0004-6361/201935952. arXiv: 1906.04195 [astro-ph.SR].

Valenti, J. A. and N. Piskunov (Sept. 1996). "Spectroscopy made easy: A new tool for fitting observations with synthetic spectra." In: *Astronomy and Astrophysics, Supplement* 118, pp. 595–603.

Veyette, Mark J. et al. (Dec. 2017). "A Physically Motivated and Empirically Calibrated Method to Measure the Effective Temperature, Metallicity, and Ti Abundance of M Dwarfs". In: *The Astrophysical Journal* 851.1, 26, p. 26. DOI: 10.3847/1538-4357/aa96aa. arXiv: 1710.10259 [astro-ph.SR].

Vogt, S. S. et al. (June 1994). "HIRES: the high-resolution echelle spectrometer on the Keck 10-m Telescope". In: *Instrumentation in Astronomy VIII*. Ed. by D. L. Crawford and E. R. Craine. Vol. 2198. Proceedings of SPIE, p. 362. DOI: 10.1117/12.176725.

Wang, Ji, Debra A. Fischer, Elliott P. Horch, et al. (June 2015). "Influence of Stellar Multiplicity On Planet Formation. III. Adaptive Optics Imaging of Kepler Stars With Gas Giant Planets". In: *The Astrophysical Journal* 806.2, 248, p. 248. DOI: 10.1088/0004-637X/806/2/248. arXiv: 1505.05363 [astro-ph.EP].

Wang, Ji, Debra A. Fischer, Ji-Wei Xie, et al. (Aug. 2014). "Influence of Stellar Multiplicity on Planet Formation. II. Planets are Less Common in Multiple-star Systems with Separations Smaller than 1500 AU". In: *The Astrophysical Journal* 791.2, 111, p. 111. DOI: 10.1088/0004-637X/791/2/111. arXiv: 1407.3344 [astro-ph.EP].

Wilson, Robert F., Caleb I. Cañas, et al. (Mar. 2022). "The Influence of 10 Unique Chemical Elements in Shaping the Distribution of Kepler Planets". In: *The Astronomical Journal* 163.3, 128, p. 128. DOI: 10.3847/1538-3881/ac3a06. arXiv: 2111.01753 [astro-ph.EP].

Wilson, Robert F., Johanna Teske, et al. (Feb. 2018). "Elemental Abundances of Kepler Objects of Interest in APOGEE. I. Two Distinct Orbital Period Regimes Inferred from Host Star Iron Abundances". In: *The Astronomical Journal* 155.2, 68, p. 68. DOI: 10.3847/1538-3881/aa9f27. arXiv: 1712.01198 [astro-ph.EP].

Wu, Yanqin, Norman W. Murray, and J. Michael Ramsahai (Nov. 2007). "Hot Jupiters in Binary Star Systems". In: *The Astrophysical Journal* 670.1, pp. 820–825. DOI: 10.1086/521996. arXiv: 0706.0732 [astro-ph].

Yee, S. W., E. A. Petigura, and K. von Braun (Feb. 2017). "Precision Stellar Characterization of FGKM Stars using an Empirical Spectral Library". In: *The Astrophysical Journal* 836, 77, p. 77. DOI: 10.3847/1538-4357/836/1/77. arXiv: 1701.00922 [astro-ph.SR].

Zhao, Gang et al. (July 2012). "LAMOST spectral survey — An overview". In: *Research in Astronomy and Astrophysics* 12.7, pp. 723–734. DOI: 10.1088/1674-4527/12/7/002.

Chapter 2

DATA-DRIVEN SPECTROSCOPY OF COOL STARS AT HIGH SPECTRAL RESOLUTION

Behmard, Aida, Erik A. Petigura, and Andrew W. Howard (May 2019). "Data-driven Spectroscopy of Cool Stars at High Spectral Resolution". In: *The Astrophysical Journal* 876.1, 68, p. 68. DOI: 10.3847/1538-4357/ab14e0. arXiv: 1904.00094 [astro-ph.EP].

Abstract

The advent of large-scale spectroscopic surveys underscores the need to develop robust techniques for determining stellar properties ("labels", i.e., physical parameters and elemental abundances). However, traditional spectroscopic methods that utilize stellar models struggle to reproduce cool (<4700 K) stellar atmospheres due to an abundance of unconstrained molecular transitions, making modeling via synthetic spectral libraries difficult. Because small, cool stars such as K and M dwarfs are both common and good targets for finding small, cool planets, establishing precise spectral modeling techniques for these stars is of high priority. To address this, we apply *The Cannon*, a data-driven method of determining stellar labels, to Keck High Resolution Echelle Spectrometer (HIRES) spectra of 141 cool (<5200 K) stars from the California Planet Search. Our implementation is capable of predicting labels for small (<1 $R_\odot$) stars of spectral types K and later with accuracies of 68 K in effective temperature (T_{eff}), 5% in stellar radius (R_*), and 0.08 dex in bulk metallicity ([Fe/H]), and maintains this performance at low spectral resolutions ($R < 5000$). As M-dwarfs are the focus of many future planet-detection surveys, this work can aid efforts to better characterize the cool star population and uncover correlations between cool star abundances and planet occurrence for constraining planet formation theories.

2.1 Introduction

Precise determination of stellar properties (e.g., masses, radii, effective temperatures, elemental abundances) is a challenging, yet essential component of stellar and planetary astrophysics. Accurate measurements of masses (M_*), radii (R_*), and temperatures (T_{eff}) are crucial for vetting models of stellar structure and evolution, and the chemical compositions of stellar photospheres reflect formation histories and can link stars to their parent molecular clouds, providing a window into galactic chemical evolution. The burgeoning field of exoplanets also calls for robust methods of determining stellar properties as characterization of planets is predicated on thorough characterization of their stellar hosts.

Stellar spectroscopy has a rich history, beginning with Annie Jump Cannon and her colleagues at Harvard College Observatory who developed the current stellar classification system based upon visual inspection of spectral features. Modern spectroscopic methods involve matching information-rich portions of empirical spectra to benchmark or synthetic spectra generated from model stellar photospheres. Two commonly-used spectral modeling tools are `SME` and `MOOG` (Valenti and Piskunov, 1996; Sneden, 1973), both of which have undergone significant evolution since their inception (e.g., Valenti and Fischer 2005; Valenti, Fischer, et al. 2009; Deen 2013; Brewer, Fischer, Basu, et al. 2015; Piskunov and Valenti 2017). However, current model photospheres are limited by an incomplete knowledge of the physics behind stellar attributes; they suffer from poorly constrained atomic and molecular opacities, often assume local thermodynamic equilibrium (LTE), and inaccurately model dynamical effects such as convection or stellar winds, if at all. While three-dimensional hydrodynamic models have been created that allow for non-LTE conditions, they still suffer from the other aforementioned drawbacks and are computationally expensive. Laboratory studies have refined atomic and molecular data and improved line lists, but departures from solar type atmospheres still present significant modeling challenges.

Stars of spectral types K4 ($T_{\rm eff} \lesssim 4700$ K) and later are particularly difficult to model with synthetic spectral techniques as their optical and NIR spectra feature dense clusters of molecular lines that lack reliable opacity data. In the optical regions of K and M dwarf spectra, TiO and VO bands are prominent, as well as hydride bands such as MgH, CaH, and FeH. The NIR regions of M dwarf spectra often feature H_2O (e.g., Rojas-Ayala et al. 2012). Characterization of late-type stars such as M dwarfs is important because they are common, representing $\sim$75% of stars in the solar neighborhood (Henry et al., 2006). Small, cool stars are also popular targets for exoplanet surveys as their low M_* and R_* result in deeper transit signals and larger Doppler shifts, increasing the probability of detecting and characterizing small planets.

Empirical methods offer alternative routes for predicting K and M dwarf parameters and abundances. Common proper motion pairs of M dwarfs and F-, G-, K-type stars of known metallicities ([Fe/H]) can be used to calibrate M dwarf metallicities with equivalent widths (EWs) of NIR spectral features (Newton, Charbonneau, Irwin, Berta-Thompson, et al., 2014; Mann, Deacon, et al., 2014). Similarly, temperatures (T_{eff}) and stellar radii (R_*) can be calibrated with EWs of K and M dwarf NIR spectra (Newton, Charbonneau, Irwin, and Mann, 2015), and parallaxes can provide further constraints on stellar properties (Mann, Feiden, et al., 2015; Mann, Dupuy, et al., 2017). Empirical as opposed to synthetic spectral libraries composed of touchstone stars with well-measured properties are also capable of predicting accurate parameters for stars of mid-K spectral types and later (Yee, Petigura, and von Braun, 2017).

Another promising method for modeling cool stars is offered by *The Cannon*, a data-driven approach to modeling spectroscopic data (Ness et al., 2015). In brief, *The Cannon* predicts stellar parameters and elemental abundances from spectroscopic data via a two-step process: a "training step" where the spectra for a set of reference objects with well-determined parameters and/or abundances are used to construct a predictive model of the flux, and a "test step" where the model is used to infer those of objects given their spectra. Unlike traditional spectroscopic modeling methods, *The Cannon* makes no use of physical stellar models, and does not require an accompanying library of synthetic spectra for reference. Here, we modify *The Cannon* to optimize parameter and elemental abundance predictions for K and M dwarfs with HIRES spectra.

Throughout this work, we refer to stellar parameters *and* elemental abundances (T_{eff}, R_*, and [Fe/H]) as "labels" to be consistent with previous literature on *The Cannon* (e.g., Ness et al. 2015; Casey et al. 2016; Ho et al. 2017) and to adhere to machine learning/supervised methods terminology. We evaluate *The Cannon*'s ability to predict stellar labels in our cool star sample with cross-validation experiments. Cross-validation was carried out by dividing a reference set of cool stars with well-determined labels into training and validation sets. The reference set is pulled from a library compiled by Yee, Petigura, and von Braun (2017) (see Section 2.2 for more details). Performance was evaluated by examining how well *Cannon*-predicted labels for the validation set matched those reported in the library. In Section 2.3, we present *The Cannon*, and outline our implementation and its performance on our cool star sample in Section 2.4. We find that *The Cannon* can predict labels with precisions of 68 K in T_{eff}, 5% in R_*, and 0.08 in [Fe/H] (dex). Discussion of the results is presented in Section 2.5.

2.2 Cool Star Sample

Our spectral library was compiled by Yee, Petigura, and von Braun (2017) and consists of 404 touchstone stars originating from several source catalogs that span the spectral types ~M5–F1 ($T_{\rm eff} \approx 3000$–7000 K, $R_* \approx 0.1$–$16\ R_\odot$) (Figure 2.1). The stars have spectra obtained from HIRES at the Keck-I 10-m telescope (Vogt et al., 1994) as part of the California Planet Search (CPS). For more details on CPS, see Howard et al. (2010). The HIRES spectra are high-resolution ($R \approx 60{,}000$) and high signal-to-noise (S/N > 40/pixel, with ~80% having S/N > 100/pixel). The spectra originate from the middle HIRES detector CCD chip and contain 16 spectral orders. The HIRES blaze function has been removed and the spectra registered onto a common wavelength scale (λ = 4990–6410 Å) uniform in $\Delta\log\lambda$ to ensure that linear velocity shifts correspond to linear pixel shifts (Yee, Petigura, and von Braun, 2017). We confined the wavelength range to 13 orders (λ = 4990–6095 Å) to avoid redder portions of the middle HIRES CCD chip that are more affected by tellurics.

To isolate a cool star sample composed of K and M dwarfs, we employed radius and temperature cuts of $T_{\rm eff}$ <5200 K and $R_* < 1\ R_\odot$, leaving 141 stars. These cool stars are primarily drawn from the catalog described in Mann, Feiden, et al. (2015) with $T_{\rm eff}$, R_*, and [Fe/H] determined from a combination of spectrophotometry, SED modeling, *Gaia* parallaxes, and EW empirical relations (quoted uncertainties of 60 K, 3.8%, and 0.08 dex, respectively). A smaller subset originate from the catalog compiled by von Braun et al. (2014), and have interferometrically-determined R_* (quoted uncertainties of <5%). Many of the early K dwarfs in the sample have $T_{\rm eff}$ and [Fe/H] determined from LTE spectral synthesis carried out by Brewer, Fischer, Valenti, et al. (2016) with `SME` (quoted uncertainties of 60 K and 0.05 dex, respectively), while the sample mid to late K dwarfs have $T_{\rm eff}$, R_*, and [Fe/H] determined from a combination of spectrophotometry, SED modeling, parallaxes, and `SME` analysis carried out by Yee, Petigura, and von Braun (2017) (quoted uncertainties of 5%, 7.4%, and 0.1 dex, respectively). Because most of these catalogs do not provide a complete set of $T_{\rm eff}$, R_*, and [Fe/H] values, Yee, Petigura, and von Braun (2017) conducted an isochrone analysis using Dartmouth stellar models (Dotter et al., 2008) to obtain a homogeneous label set, and took uncertainties as the 5th and 95th percentiles of the MCMC distributions that resulted from fitting to the stellar model grids. For more details on any of the library catalogs or the isochrone analysis procedure, see Yee, Petigura, and von Braun (2017).

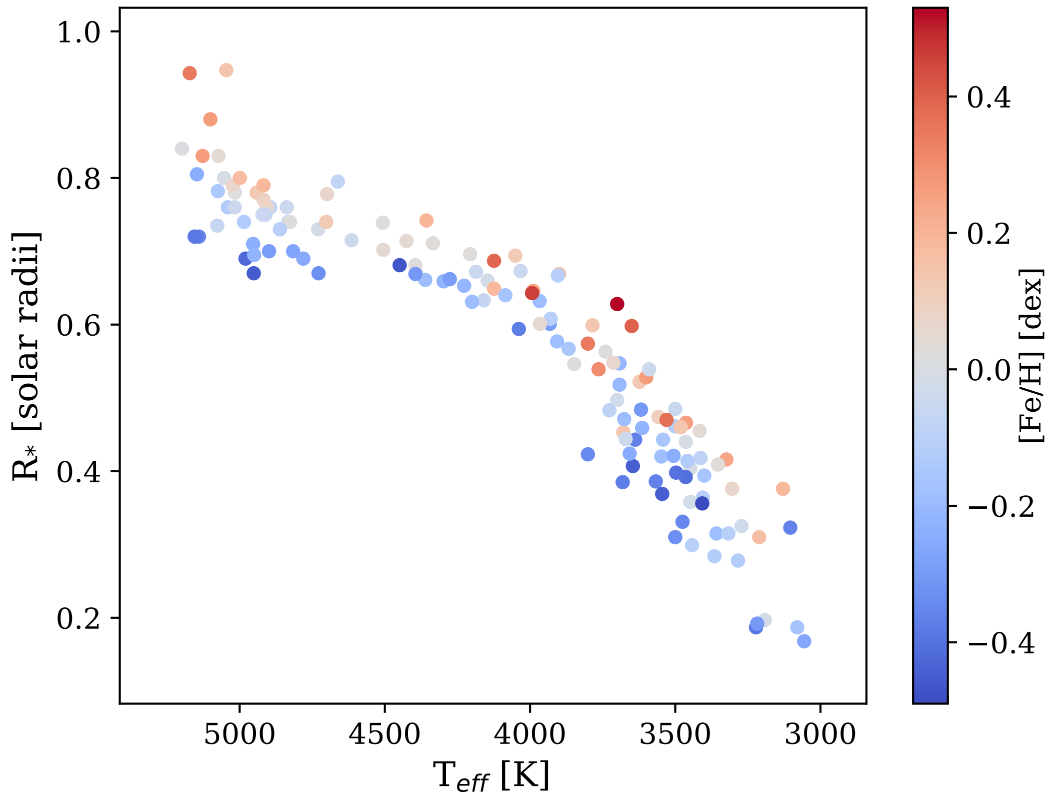


Figure 2.1: The domain of $T_{\rm eff}$, R_*, and [Fe/H] for our reference sample of 141 cool stars pulled from the library outlined in Yee, Petigura, and von Braun (2017). The cool stars have temperatures and radii that satisfy $T_{\rm eff} < 5200$ K and $R_* < 1\ R_\odot$.

2.3 *The Cannon*

Preparing HIRES spectra for *The Cannon*

To prepare the spectral library for *The Cannon*, we must ensure that the spectra satisfy certain conditions; the spectra must share a common wavelength grid, be shifted onto the rest wavelength frame, share a common line-spread function, and be continuum-normalized via a method independent of S/N (Ness et al., 2015). The first two conditions are already satisfied for the library spectra, and we can assume that they effectively share a line-spread function, though there may be negligible variation due to variable observation seeing conditions. To carry out normalization, we applied error-weighted, broad Gaussian smoothing with

$$\bar{f}(\lambda_0) = \frac{\sum_j (f_j \sigma_j^{-2} w_j(\lambda_0))}{\sum_j (\sigma_j^{-2} w_j(\lambda_0))}, \tag{2.1}$$

where f_j is the flux at pixel j of the wavelength range, σ_j is the uncertainty at pixel

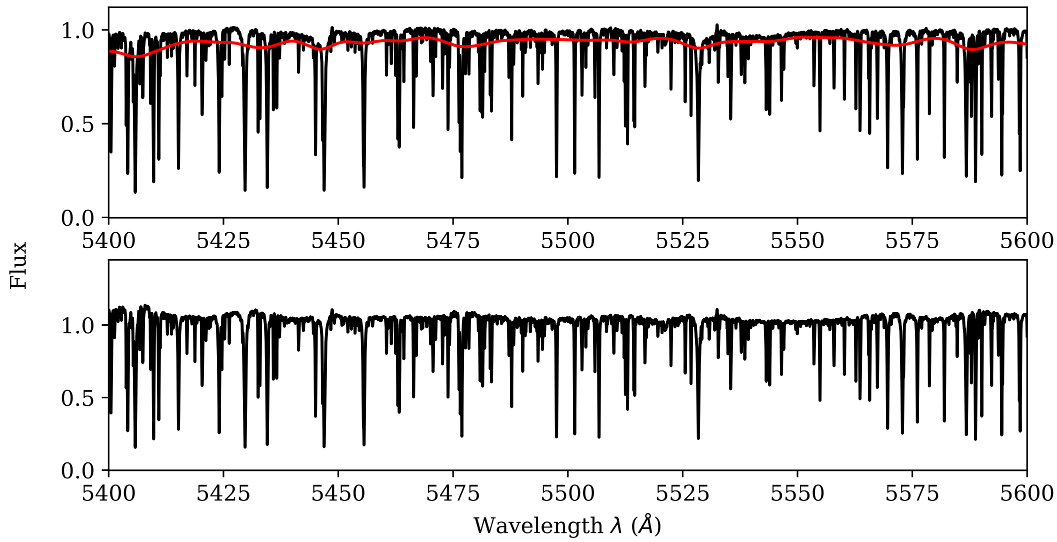


Figure 2.2: HIRES spectrum of a reference sample star (HD100623) before and after normalization. The top panel shows the pre-normalized spectrum overlaid with the Gaussian-smoothed version of itself in red, while the bottom panel shows the normalized spectrum after the Gaussian-smoothed signal was divided out. The displayed wavelength region (λ = 5400-5600 Å) is a subset of the full wavelength range and was chosen for better visualization of the spectrum and accompanying Gaussian-smoothed curve.

j, and the weight $w_j(\lambda_0)$ is drawn from a Gaussian:

$$w_j(\lambda_0) = e^{-\frac{(\lambda_0-\lambda_j)^2}{L^2}}, \tag{2.2}$$

where L was chosen to be 3 Å. If larger L values are chosen for HIRES spectra, continuum-normalization begins to remove high resolution features. For reference, Ho et al. (2017) used a width of L = 50 Å to normalize low-resolution Large Sky Area Multi-Object Fibre Spectroscopic Telescope (LAMOST) spectra ($R \approx 1800$). The Gaussian smoothing procedure is illustrated in Figure 2.2.

Training Step

We used *The Cannon 2*, the second implementation of *The Cannon* developed by Casey et al. (2016). Hereafter, we will refer to *The Cannon 2* simply as *The Cannon*. This version builds upon the original with additional features that are designed to aid prediction of a larger label set including elemental abundances that go beyond bulk metallicity ([Fe/H]), such as regularization.

As outlined in Section 2.1, in the training step, *The Cannon* uses a set of reference objects with well-determined labels to construct a predictive model of the flux at every pixel of the wavelength range that is a function of the stellar labels. Model construction is based on two assumptions: that continuum-normalized spectra with identical labels look identical at every pixel, and that the flux at every pixel in a spectrum changes continuously as a function of the stellar labels. While *The Cannon* can be trained on any set of empirical spectra and their labels, the resultant model will only be capable of predicting labels for spectra with properties that are represented in the training set. In other words, *The Cannon* is not able to accurately extrapolate outside the training set parameter space, so the training set spectra must be representative of the test set spectra in order to predict accurate label values. It is also important to note that the *Cannon*-predicted labels will only be as accurate as those of the training set.

The flux model f_{jn} for a spectrum n at pixel j can be written as

$$f_{jn} = \boldsymbol{v}(l_n) \cdot \boldsymbol{\theta}_j + e_{jn} \, , \tag{2.3}$$

where $\boldsymbol{\theta}_j$ is the set of spectral model coefficients at each pixel j and $\boldsymbol{v}(l_n)$ is a function of the label list l_n that is unique for each spectrum n. The function $\boldsymbol{v}(l_n)$ is referred to as the "vectorizer" which can accommodate functions that are linear in the coefficients $\boldsymbol{\theta}_j$, but not necessarily simple polynomial expansions of the label list l_n. The noise term is described by e_{jn} and can be taken as sampled from a Gaussian with zero mean and variance $\sigma_{jn}^2 + s_j^2$ where σ_{jn}^2 is the uncertainty reported on the input HIRES spectra (flux variance) and s_j^2 is the intrinsic scatter of the model at each pixel j. This intrinsic scatter can be likened to the expected deviation of the model from the spectrum at j.

To determine the optimal model labels ($\boldsymbol{\theta}_j$,s_j^2), we can relate the flux model to a single-pixel log-likelihood function:

$$\begin{aligned} \ln p(f_{jn}|\boldsymbol{\theta}_j, \boldsymbol{v}(l_n), s_j^2) = -\frac{[f_{jn} - \boldsymbol{v}(l_n) \cdot \boldsymbol{\theta}]^2}{\sigma_{jn}^2 + s_j^2} - \\ \ln(\sigma_{jn}^2 + s_j^2) - \Lambda Q(\boldsymbol{\theta}) \, , \end{aligned} \tag{2.4}$$

where Λ is a regularization parameter and $Q(\boldsymbol{\theta})$ is a regularizing function that encourages the model coefficients $\boldsymbol{\theta}_j$ to take on zero values, resulting in a simpler model that is less prone to overfitting. In the case of L1 regularization implemented within *The Cannon*, the regularizing function takes the form

$$Q(\boldsymbol{\theta}) = \sum_{j=0}^{J-1} |\theta_j| \,. \tag{2.5}$$

L1 regularization was chosen because *The Cannon* is designed for predicting large sets of elemental abundances, and it is reasonable to assume that only one or a few elemental abundances will affect the flux at a single pixel of the wavelength range. For more details on regularization or the model itself, see (Casey et al., 2016).

In the training step, the log-likelihood is maximized via the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm to derive the best-fit model coefficients $\boldsymbol{\theta}_j$ and intrinsic scatter s_j^2:

$$\boldsymbol{\theta}_j, s_j^2 \leftarrow \underset{\boldsymbol{\theta}, s_j}{a} rgmax \Big[\sum_{n=0}^{N-1} \ln p(f_{jn} | \boldsymbol{\theta}_j, \boldsymbol{v}(l_n), s_j^2) \Big] \,. \tag{2.6}$$

Plugging in the explicit form of the log-likelihood (Equation 2.4) leads to

$$\begin{aligned} \boldsymbol{\theta}_j, s_j^2 \leftarrow \underset{\boldsymbol{\theta}, s_j}{a} rgmax \Big[\sum_{n=0}^{N-1} - \frac{[f_{jn} - \boldsymbol{v}(l_n) \cdot \boldsymbol{\theta}]^2}{\sigma_{jn}^2 + s_j^2} \\ - \sum_{n=0}^{N-1} \ln(\sigma_{jn}^2 + s_j^2) \,-\, \Lambda Q(\boldsymbol{\theta}) \Big] \,. \end{aligned} \tag{2.7}$$

Test Step

In the test step, we set the model labels ($\boldsymbol{\theta}_j$,s_j^2) to the optimized values determined during the training step at every pixel j, and fit for the label list l_n for each star n that minimizes the log-likelihood:

$$l_n \leftarrow \underset{l_n}{argmin} \Big[\sum_{j=0}^{J-1} - \frac{[f_{jn} - \boldsymbol{v}(l_n) \cdot \boldsymbol{\theta}]^2}{\sigma_{jn}^2 + s_j^2} \Big] \,. \tag{2.8}$$

Optimization of the log-likelihood in the test step is carried out via weighted least squares.

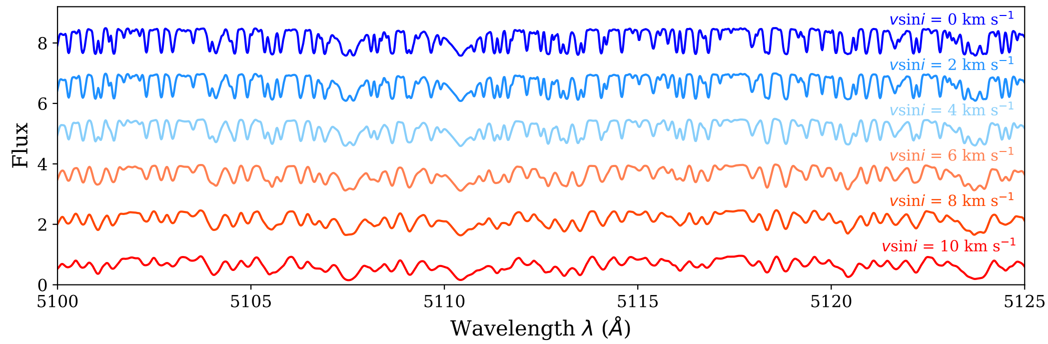


Figure 2.3: Synthetic spectra generated via `SpecMatch-Syn` under the same $T_{\rm eff}$, $\log g$, and [Fe/H] conditions ($T_{\rm eff}$ = 4000 K, $\log g$ = 4.5 cm s^{-2}, [Fe/H] = 0.2 dex) but with varying amounts of additional broadening. From top to bottom, the spectra have $v\sin i$ = 0–10 km s^{-1} in increments of +2 km s^{-1}.

2.4 *The Cannon* Performance

Building Intuition with Synthetic Spectra

Before running *The Cannon* on our cool star sample, we sought to establish a measure of baseline performance. We did this by constructing a sample of synthetic spectra that mimics the cool star sample: 141 "stars" with the same label values as the true cool sample. Because the labels of the synthetic spectra, by definition, lack uncertainty, they can provide a sense of how well *The Cannon* performs under perfect conditions. The synthetic spectra are generated from the publicly-available code `SpecMatch-Syn` which fits five regions of optical spectra by interpolating within a grid of model spectra from the library described in Coelho et al. (2005). For more details on `SpecMatch-Syn`, see Petigura (2015). See Figure 2.3 for examples of synthetic spectra.

We tested the validity of *Cannon*-predicted labels through a bootstrap leave-one-out cross-validation scheme where we trained the spectral model on all objects in the synthetic spectral sample but one, and predicted labels for the object that was left out. We carried out this scheme iteratively to pass through the entire sample and predict labels for every object. Following the work of Ness et al. (2015) and Casey et al. (2016), we began with a spectral model in which the label list l_n was quadratic in the labels, resulting in the following label list:

$$\begin{aligned} l_n \equiv [1, T_{\rm eff}, R_*, [Fe/H], T_{\rm eff}^2, T_{\rm eff}\cdot R_*, \\ T_{\rm eff}\cdot[Fe/H], R_*^2, R_*\cdot[Fe/H], [Fe/H]^2], \end{aligned} \tag{2.9}$$

where 1, the first element in the label list, is there to allow for a linear offset in the fitting (Ness et al., 2015). We found that modeling the projected rotational velocity $v\sin i$ as a fitted-for parameter in addition to $T_{\rm eff}$, R_*, and [Fe/H] resulted in more accurate labels predictions; a second order model without $v\sin i$ achieves accuracies of 40 K in $T_{\rm eff}$, 13% in R_*, and 0.06 dex in [Fe/H], while a second order model with $v\sin i$ achieves accuracies of 32 K in $T_{\rm eff}$, 13% in R_*, and 0.03 dex in [Fe/H].

Using a third order rather than second order (quadratic-in-label) model with $v\sin i$ further improves label predictions; a third order model achieves accuracies of 22 K in $T_{\rm eff}$, 8% in R_*, and 0.03 dex in [Fe/H]. Thus, these tests with synthetic spectra motivate a third order *Cannon* model with $v\sin i$ included as a label. The third order model results in a label list composed of additional third order cross terms, bringing the total number of terms up to 20.

Cool Star Sample

To run *The Cannon* on the cool star sample, we employed the same bootstrap leave-one-out cross-validation scheme. As in the case of synthetic spectra, the cool star HIRES spectra are best described by a third order model, which is unsurprising given their resolution of $R \approx 60{,}000$ (~3 times the resolution of APOGEE spectra). The more flexible model may also better-describe our more diverse training set, composed of stars with a wider $T_{\rm eff}$ range (APOGEE stars are confined to $T_{\rm eff}$ = 3500–5500 K). A third order model fitting for $T_{\rm eff}$, R_*, and [Fe/H] achieves precisions of 80 K in $T_{\rm eff}$, 6% in R_*, and 0.1 dex in [Fe/H].

We found that *The Cannon* predicted anomalously poor label values for one source (GL896A). Upon inspection, we found that the spectrum of GL896A exhibits significantly broader features than any other source in our sample. Because GL896A is not well-represented in the training set, *The Cannon* is unable to construct a model that well-describes GL896A (Figure 2.5, top panel). While such fast rotators are rare amongst K and M dwarfs, the presence of this target indicates that our implementation of *The Cannon* must still take them into consideration. We modified our implementation by augmenting and diversifying the training set; we created x copies of each spectrum in the cool star sample (exploring different values of x to see which resulted in the best performance), and applied differential values of artificial broadening to simulate faster stellar rotation. Artificial broadening was carried out by convolving the spectra with a rotational-macroturbulent broadening kernel described in Hirano et al. (2011).

In order for this scheme to work, $v\sin i$ must be specified as a fitted-for label as in the tests with synthetic spectra. This is problematic because more than half of the cool stars do not have reported $v\sin i$ values. We dealt with this by assigning all sources in the augmented sample a new label that describes general broadening, taken to be the FWHM of a Gaussian fitted to the spectral autocorrelation peaks (Figure 2.4). This resulted in better flux predictions for the spectrum of GL896A (Figure 2.5), and better label predictions overall. The most precise labels are achieved when the cool star sample is augmented by $x = 5$ (5 copies generated for each spectrum), and the copies are artificially broadened by 0–5 km s^{-1} as the cool star sample does not appear to include any significantly rapid rotators ($v\sin i > 5$ km s^{-1}). We ultimately achieved precisions of 68 K in $T_{\rm eff}$, 5% in R_*, and 0.08 dex in [Fe/H], and verified that these label predictions vary within the reported precisions for different *Cannon* runs. While it may seem surprising that *The Cannon* achieves better predictions in R_* for empirical spectra compared to synthetic spectra (5% versus 8% in R_*), it should be noted that the synthetic spectra may not accurately reflect the input R_* values as a conversion to $\log g$ was required, which in turn required M_*. We do not have M_* values for the cool star sample, and instead assumed a linear relationship between R_* and M_* ($M_*/M_\odot = R_*/R_\odot$) to obtain mass estimates. This is a valid approximation for the main sequence, but is not perfect.

Because of the large number of terms in our model, we considered overfitting to be a potential issue. That is, overly precise modeling of the training set flux may lead to less accurate label predictions. To address this, we added regularization to our *Cannon* model and assessed whether label prediction improved. We explored a grid of regularization strengths from $\Lambda = 10^{-6}$ to $\Lambda = 10^{2}$ uniform in log space. We found that no matter the regularization strength, adding regularization to the model always resulted in less precise label predictions. It is possible that regularization does not lead to better predictions for the 3 labels ($T_{\rm eff}$, R_*, [Fe/H]) we are considering because all of these labels affect the flux at each wavelength point. Thus, we do not benefit from regularization that encourages sparsity (L1, encourages the model coefficients to go to zero). L1 regularization may lead to better label predictions if we expand our label set to include elemental abundances, but that is beyond the scope of this study.

Performance at Low S/N

To investigate the effect of photon shot noise on the precision of label predictions made with *The Cannon*, we carried out the same procedure employed by Yee, Petigura, and von Braun (2017) for the empirical spectroscopic tool `SpecMatch-Emp`;

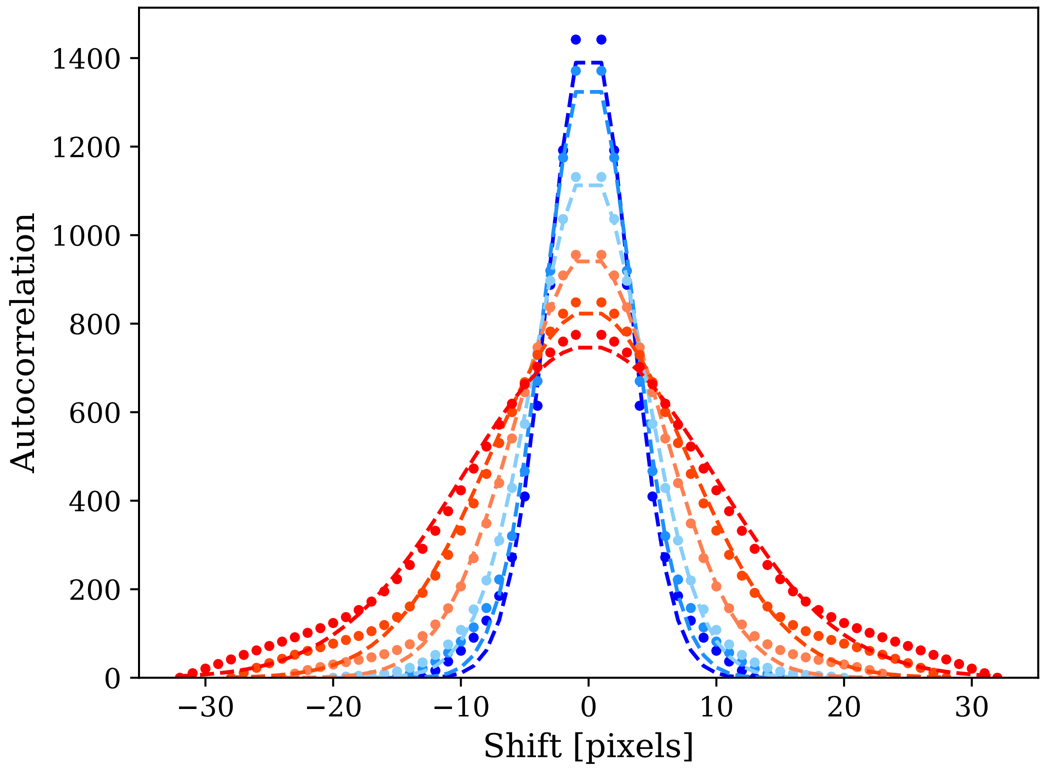


Figure 2.4: The autocorrelation functions of spectra displayed in Figure 2.3, marked in dotted lines. The overlaid Gaussian fits are displayed in dashed lines.

we isolated a subset of 20 stars from the cool star sample with S/N > 160/pixel and degraded their spectra by injecting Gaussian noise to simulate target S/N values of 120, 100, 80, 60, 40, 20, and 10 per pixel. We generated 20 S/N-degraded spectra for each spectrum in the subset and S/N target value, then compared the precision of the *Cannon* label predictions for the degraded spectra with those of the original S/N > 160/pixel spectra, which we took as ground truth. The results are summarized in Figure 2.7.

As expected, lower S/N leads to larger median scatter in label predictions made with *The Cannon*. However, the median scatter at S/N = 10/pixel is still low, with 3.5 K in T_{eff}, 0.4% in R_*, and 0.006 dex in [Fe/H]. This demonstrates that *The Cannon* is quite robust, even at low S/N values. This performance is better than that achieved by `SpecMatch-Emp`, which has median scatter values at S/N = 10/pixel of 10.4 K in T_{eff}, 1.7 % in R_*, and 0.008 dex in [Fe/H] (Yee, Petigura, and von Braun, 2017), though it should be noted that `SpecMatch-Emp` conducted this test with stars spanning the HR diagram while our sample is T_{eff}-limited.

Motivated by the small observed scatter in [Fe/H], we attempted to estimate the min-

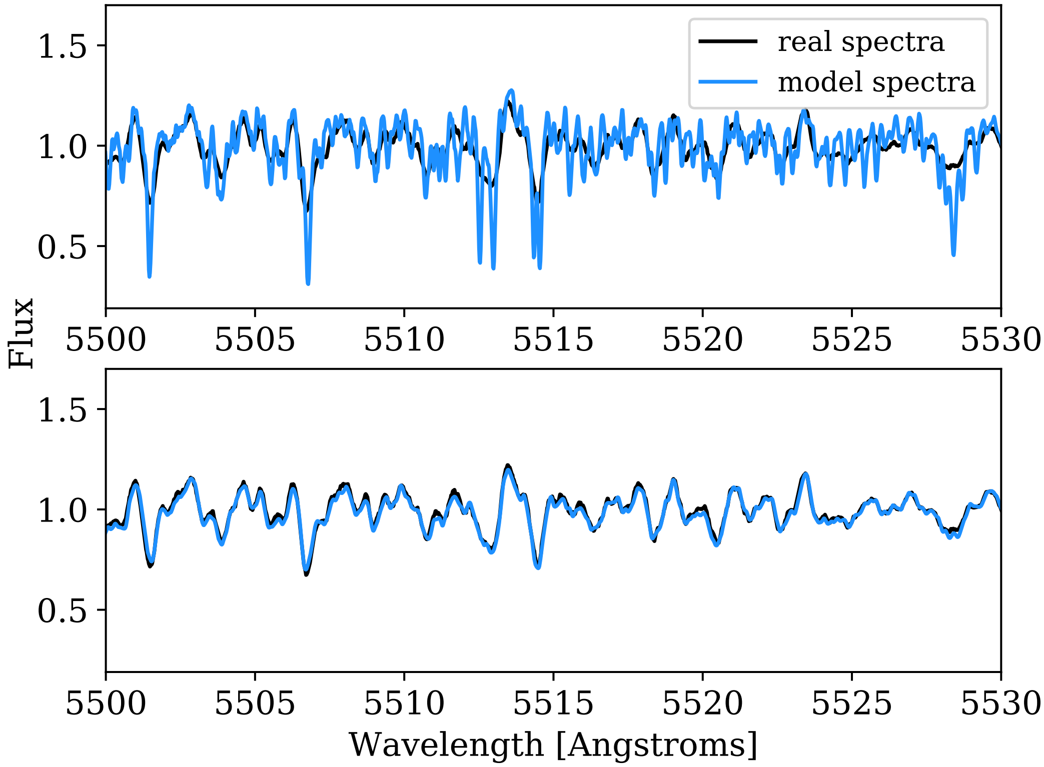


Figure 2.5: The spectrum of GL896A overlaid with the *Cannon* model before augmenting the library with broadened copies of the spectra (top), and after (bottom). The true spectrum is plotted in black while the *Cannon* models are plotted in blue.

imum change in [Fe/H] that is theoretically detectable. To do so, we considered the difference between two spectra corresponding to stars with slightly different metallicities (Δ[Fe/H]). We defined a quantity $\mathcal{S}$ that relates three quantities: Δ[Fe/H]; the derivative of the spectrum with changing metallicity, $\delta f/\delta[Fe/H]$; and the flux uncertainty σ_f. $\mathcal{S}$ can be thought of as analogous to S/N. For the jth pixel, the relation is

$$\mathcal{S}_j = \frac{\left(\frac{\delta f}{\delta[Fe/H]}\right)_j \Delta[Fe/H]}{\sigma_{f,j}} \ . \tag{2.10}$$

This equation can be rewritten as

$$\mathcal{S}_j = \frac{\left(\frac{\delta f}{\delta[Fe/H]}\right)_j \Delta[Fe/H]}{c_j\langle\sigma_f\rangle} \ , \tag{2.11}$$

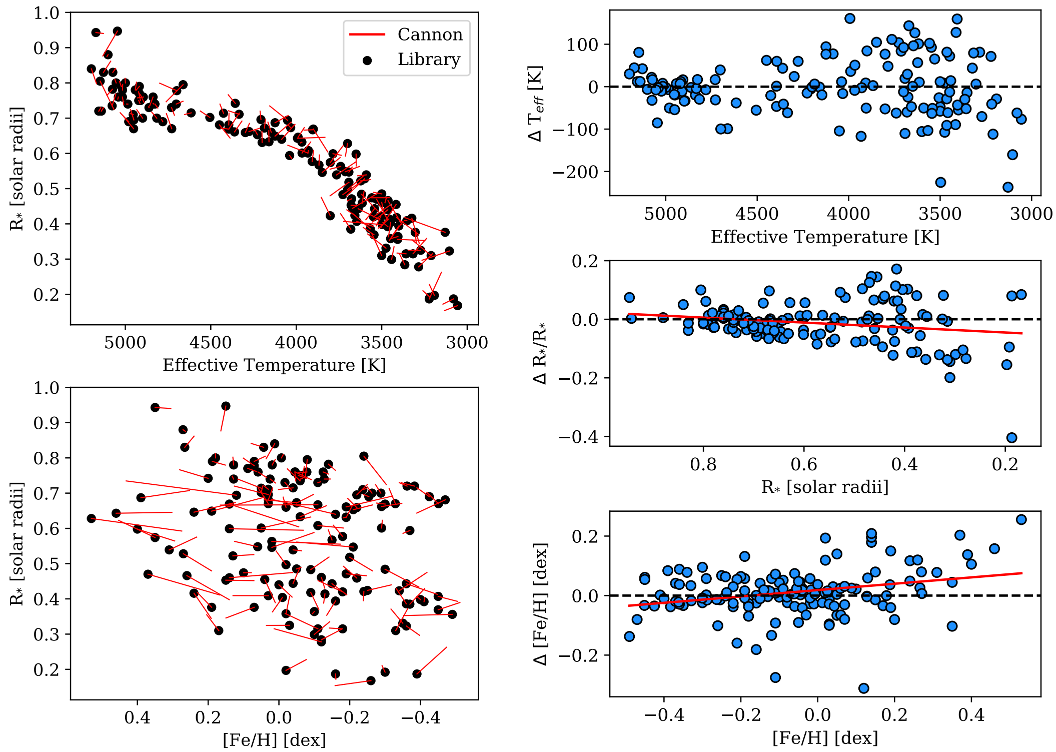


Figure 2.6: Comparison of the cool star sample library labels with the *Cannon*-predicted labels (T_{eff}, R_*, [Fe/H]). In the left panel plots, the black points represent the library labels and the red lines represent the *Cannon* labels. The right panel plots display the label residuals, with the red lines denoting possible trends. We note that the slope values of these linear trends are much lower than those of residuals from labels predicted via techniques that make use of empirical spectral libraries (Yee, Petigura, and von Braun, 2017).

where $\langle \sigma_f \rangle$ is the average flux uncertainty and c_j is directly related to the blaze function. The total $\mathcal{S}$ (summing over pixels) of the metallicity measurement can be written as

$$\begin{aligned} \mathcal{S} &= \sqrt{\sum_j (\mathcal{S}_j)^2} \\ &= \frac{\Delta[Fe/H]}{\langle \sigma_f \rangle} \sqrt{\sum_j \left[\left(\frac{\delta f}{\delta[Fe/H]} \right) / c_j \right]^2} \, . \end{aligned} \tag{2.12}$$

Rearranging terms to solve for the minimum theoretically detectable change in metallicity yields

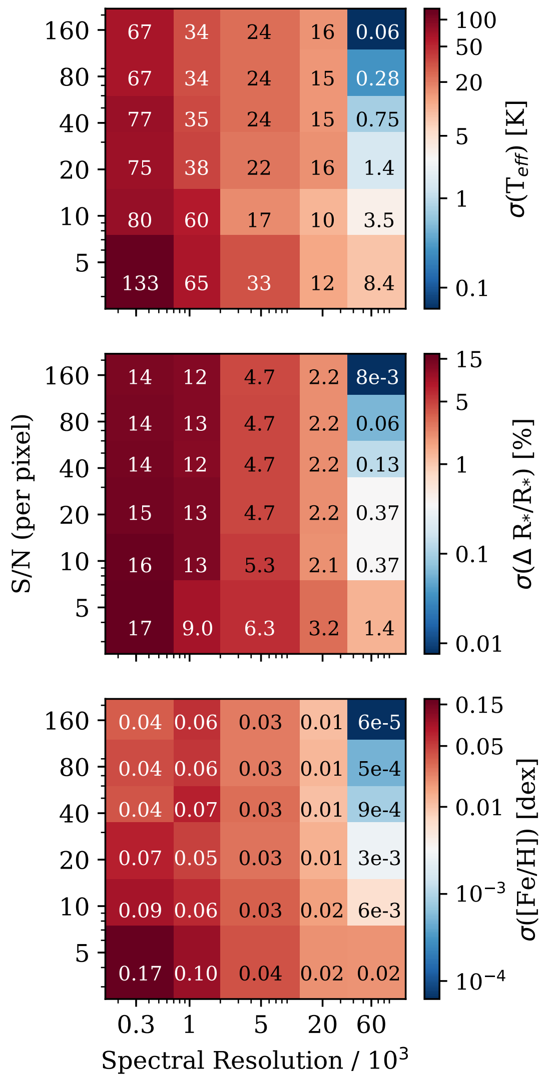


Figure 2.7: Log-log plots showing the median scatter of *Cannon*-derived labels as a function of both S/N and resolution. Each colored block within the subplots represents the median RMS difference in T_{eff} (top), R_* (middle), and [Fe/H] (bottom) predictions from the cool star subset with spectra satisfying S/N > 160/pixel when degraded to lower S/N and resolution. The median RMS difference is also explicitly provided within each block in units of K (T_{eff}), solar radii (R_*), and dex ([Fe/H). The median scatter increases as the S/N and resolution decreases, which is representative of the effect photon shot noise and lower resolution would have on the precision of *Cannon* label predictions for HIRES spectra.

$$\Delta[Fe/H] = \mathcal{S}\langle\sigma_f\rangle \Bigg/ \sqrt{\sum_j \left[\left(\frac{\delta f}{\delta[Fe/H]}\right)/c_j\right]^2} \ . \tag{2.13}$$

A metallicity change is detectable at 1σ if $\mathcal{S} = 1$. For a S/N =10/pixel as considered above, i.e., $\langle \sigma_f \rangle = 0.1$, we find Δ[Fe/H] = 0.001 dex. This is much smaller than the median scatter in [Fe/H] predictions made with *The Cannon* at S/N = 10/pixel (0.006 dex). Therefore, the sensitivity of *The Cannon* lies within theoretical bounds.

Performance at Low Spectral Resolution

While HIRES spectra are observed at $R \approx 60{,}000$, many large spectroscopic surveys are observed at lower spectral resolutions. Thus, it is valuable to quantify how spectral resolution affects the accuracies of label predictions with *The Cannon*. We expected performance to decrease as spectral resolution decreases because lines will blend together, resulting in less spectral information for *The Cannon* to work with.

To investigate spectral resolution dependence, we followed the same procedure used for the S/N degradation test; we used the same subset of 20 stars with S/N > 160 and simulated lower resolution by convolving their spectra with a Gaussian kernel. We again treated the label predictions of the original high resolution ($R \approx 60{,}000$) spectra as ground truth. We simulated spectra with target resolution values of $R = 50{,}000$, 40,000, 30,000, 20,000, 10,000, 7500, and 5000. The results are summarized in Figure 2.7, which also illustrates how the precisions of label predictions are affected when both S/N and resolution are degraded.

As in the case of degraded S/N, the accuracy of *Cannon* label predictions decrease with spectral resolution. At $R = 30{,}000$, median scatter in the labels is 6.7 K in $T_{\rm eff}$, 0.2% in R_*, and 0.009 dex in [Fe/H]. This performance is better than that of `SpecMatch-Emp`'s at equivalent resolution, with median scatter values of 10.1 K in $T_{\rm eff}$, 1.3% in R_*, and 0.014 dex in [Fe/H] (Yee, Petigura, and von Braun, 2017).

The Cannon also exhibits a much slower reduction in label accuracy as resolution continues to decrease; at $R = 5000$, the median scatter in *Cannon* predictions is 24.1 K in $T_{\rm eff}$, 4.7% in R_*, and 0.026 dex in [Fe/H], while `SpecMatch-Emp`'s is 962 K in $T_{\rm eff}$, 228% in R_*, and 0.094 dex in [Fe/H] (Yee, Petigura, and von Braun, 2017). This suggests that *The Cannon* would be a favorable method for predicting labels for spectra from many large, lower resolution spectroscopic surveys (e.g., SEGUE (Beers et al., 2006), $R \approx 2000$, RAVE (Steinmetz et al., 2006), $R \approx 7000$, LAMOST (Newberg et al., 2012), $R \approx 1800$).

Performance with Label Errors

To investigate the effect of errors in the library labels on predictions made with *The Cannon*, we followed the same procedure used for the S/N and resolution degradation tests; we used the same subset of 20 stars with S/N > 160 and injected Gaussian noise into the labels to simulate additional uncertainty up to 1x the achievable precisions (68 K in $T_{\rm eff}$, 5% in stellar radius R_*, and 0.08 dex in [Fe/H]). We found that the labels are quite robust to realistic random noise in the library labels; adding 1x uncertainty leads to an increase in label prediction uncertainties of 22 K in $T_{\rm eff}$, 4% in R_*, and 0.06 dex in [Fe/H]. The results are summarized in Table 2.1.

It is worth noting that the scatter in *Cannon*-predicted values of $T_{\rm eff}$ is lower than the original label uncertainty by more than 50%, suggesting that in the limit of a very large library with labels containing a certain amount of random noise, *The Cannon* can derive a model that yields a higher $T_{\rm eff}$ precision compared to that of the library spectra. We note that this result is insensitive to zero-point offsets; it is not possible to bootstrap to higher label precisions using *The Cannon*.

Table 2.1: Median RMS scatter in all *Cannon*-derived labels

Added uncertainty ($T_{\rm eff}$, R_*, [Fe/H])	$\sigma(T_{\rm eff})$ K	$\sigma(\Delta R_*/R_*)$ %	σ([Fe/H]) dex
68 K	22	2	0.02
5% R_*	17	4	0.02
0.08 dex	10	1	0.06

The label scatter was computed after adding an 1x the amount of uncertainty to all labels.

2.5 Discussion

We evaluated how well *The Cannon*, a data-driven spectroscopic tool, is able to predict stellar labels for cool stars ($T_{\rm eff}$ = 3000–5200 K) given high-resolution spectra. With adjustments to the spectral training set, it achieves precisions of 68 K in $T_{\rm eff}$, 5% in R_*, and 0.08 dex in [Fe/H]. Unlike traditional spectroscopic modeling techniques, *The Cannon* does not rely on stellar models that struggle to reproduce the complexities of cool star spectra. Rather, as a data-driven method, *The Cannon*'s performance improves as the input spectra become more information-rich.

In the case of spectra with perfect labels (no uncertainty) as simulated with synthetic spectra, *The Cannon* achieves label accuracies of 22 K in $T_{\rm eff}$, 8% in R_*, and 0.03 dex in [Fe/H]. *The Cannon* generally makes better label predictions for synthetic spectra because the labels of real spectra include uncertainties that are endemic to the catalogs from which the cool star sample originates. These catalogs are described in von Braun et al. (2014), Mann, Feiden, et al. (2015), Brewer, Fischer, Valenti, et al.

(2016), and Yee, Petigura, and von Braun (2017), and present labels derived from a combination of modified SME (Brewer, Fischer, Basu, et al., 2015), photometry, parallaxes, interferometry, and empirical relations between the labels and EWs of spectral features. Each of these techniques have associated uncertainties, resulting in less precise label predictions with *The Cannon* when compared to the case of spectra with perfect labels.

Compared to current synthetic spectral techniques (SME, MOOG, etc.), *The Cannon* is better-suited for predicting the labels of cool stars. While the latest iterations of spectral synthesis codes model cool stars more successfully than initial versions with additions such as more accurate radiative transfer algorithms, equations of state, and larger line lists, they still lack complete sets of molecular line opacities and sufficient constraints to fully disentangle the effects of T_{eff}, $\log g$, and abundances (e.g., Bean et al. 2006; Piskunov and Valenti 2017).

It is more appropriate to compare *The Cannon* to other data-based techniques such as SpecMatch-Emp, a label-predicting spectroscopic tool developed by Yee, Petigura, and von Braun (2017) that utilizes an empirical spectral library. While SpecMatch-Emp achieves accuracies of 70 K in T_{eff}, 10% in R_*, and 0.12 dex in [Fe/H] for stars of spectral types ~K4 and later, these label predictions are slightly worse than those achieved by *The Cannon*. In addition, the residuals from label predictions with SpecMatch-Emp display linear trends where residuals are more negative for larger values in the label space, and more positive for smaller values in the label space (Yee, Petigura, and von Braun, 2017). These trends are partly explained by considering that the empirical spectral library spans a finite region (convex hull of the label values), and is inclined to pull spectral predictions at the edge of the region towards the center. While the residuals from label predictions with *The Cannon* also display slight linear trends, they are less pronounced and constitute a smaller source of systematic error (Fig. 3.6, right panel). This is because the choice of flux model coefficient values allows for some extrapolation outside the finite region spanned by the training set.

While *The Cannon* is a powerful tool for spectroscopic characterization, it has a number of drawbacks. For example, by individually treating each pixel within the spectral wavelength range, it assumes no covariance between flux values of any pixels. However, multiple spectral features can be affected by a single label, such as a particular elemental abundance or ionization state. This motivates converting *The Cannon* into a fully Bayesian framework through the inclusion of priors such as line lists to address covariance of different spectral features, or known correlations between labels such as the Stefan-Boltzman relation.

Although L1 regularization does not improve cool star label predictions for T_{eff}, R_*, and [Fe/H], L2 regularization may be better suited to such cases where labels do not include large sets of elemental abundances as L2 regularization does not encourage model coefficients to go to zero as rapidly. However, we are also interested in eventually using *The Cannon* to predict elemental abundances, in which case L1 regularization may become a useful feature. For example, we are interested in comparing the C/O ratios of K and M dwarfs to the characteristics of planets they host as such volatile ratios can probe planet formation histories. Ultimately, we will use *The Cannon* to conduct large demographic studies of cool stars with HIRES spectra with the goal of establishing correlations between small, cool stars such as K and M dwarfs and the planets they host. This work has wide potential application given that many future exoplanet surveys are focused on cool stars such as M dwarfs.

References

Bean, J. L. et al. (Dec. 2006). "Accurate M Dwarf Metallicities from Spectral Synthesis: A Critical Test of Model Atmospheres". In: *The Astrophysical Journal* 652, pp. 1604–1616. DOI: 10.1086/508321. eprint: astro-ph/0608093.

Beers, T. C. et al. (2006). "The SDSS-I Value-Added Catalog of stellar parameters and the SEGUE pipeline". In: *Memorie della Societa Astronomica Italiana* 77, p. 1171.

Brewer, J. M., D. A. Fischer, S. Basu, et al. (June 2015). "Accurate Gravities of F, G, and K stars from High Resolution Spectra Without External Constraints". In: *The Astrophysical Journal* 805, 126, p. 126. DOI: 10.1088/0004-637X/805/2/126. arXiv: 1503.07180 [astro-ph.SR].

Brewer, J. M., D. A. Fischer, J. A. Valenti, et al. (Aug. 2016). "Spectral Properties of Cool Stars: Extended Abundance Analysis of 1,617 Planet-search Stars". In: *The Astrophysical Journal Supplement Series* 225, 32, p. 32. DOI: 10.3847/0067-0049/225/2/32. arXiv: 1606.07929 [astro-ph.SR].

Casey, A. R. et al. (Mar. 2016). "The Cannon 2: A data-driven model of stellar spectra for detailed chemical abundance analyses". In: *ArXiv e-prints*. arXiv: 1603.03040 [astro-ph.SR].

Coelho, P. et al. (Nov. 2005). "A library of high resolution synthetic stellar spectra from 300 nm to 1.8 μm with solar and α-enhanced composition". In: *Astronomy and Astrophysics* 443, pp. 735–746. DOI: 10.1051/0004-6361:20053511. eprint: astro-ph/0505511.

Deen, C. P. (Sept. 2013). "Modification of the MOOG Spectral Synthesis Codes to Account for Zeeman Broadening of Spectral Lines". In: *The Astronomical Journal* 146, 51, p. 51. DOI: 10.1088/0004-6256/146/3/51. arXiv: 1306.6625 [astro-ph.SR].

Dotter, A. et al. (Sept. 2008). "The Dartmouth Stellar Evolution Database". In: *The Astrophysical Journal Supplement Series* 178, pp. 89–101. DOI: 10.1086/589654. arXiv: 0804.4473.

Henry, T. J. et al. (Dec. 2006). "The Solar Neighborhood. XVII. Parallax Results from the CTIOPI 0.9 m Program: 20 New Members of the RECONS 10 Parsec Sample". In: *The Astronomical Journal* 132, pp. 2360–2371. DOI: 10.1086/508233. eprint: astro-ph/0608230.

Hirano, T. et al. (Dec. 2011). "Improved Modeling of the Rossiter-McLaughlin Effect for Transiting Exoplanets". In: *The Astrophysical Journal* 742, 69, p. 69. DOI: 10.1088/0004-637X/742/2/69. arXiv: 1108.4430 [astro-ph.EP].

Ho, A. Y. Q. et al. (Feb. 2017). "Label Transfer from APOGEE to LAMOST: Precise Stellar Parameters for 450,000 LAMOST Giants". In: *The Astrophysical Journal* 836, 5, p. 5. DOI: 10.3847/1538-4357/836/1/5. arXiv: 1602.00303 [astro-ph.SR].

Howard, A. W. et al. (Oct. 2010). "The California Planet Survey. I. Four New Giant Exoplanets". In: *The Astrophysical Journal* 721, pp. 1467–1481. DOI: 10.1088/0004-637X/721/2/1467. arXiv: 1003.3488 [astro-ph.EP].

Mann, A. W., T. Dupuy, et al. (June 2017). "The Gold Standard: Accurate Stellar and Planetary Parameters for Eight Kepler M Dwarf Systems Enabled by Parallaxes". In: *The Astronomical Journal* 153, 267, p. 267. DOI: 10.3847/1538-3881/aa7140. arXiv: 1705.01545 [astro-ph.EP].

Mann, A. W., G. A. Feiden, et al. (May 2015). "How to Constrain Your M Dwarf: Measuring Effective Temperature, Bolometric Luminosity, Mass, and Radius". In: *The Astrophysical Journal* 804, 64, p. 64. DOI: 10.1088/0004-637X/804/1/64. arXiv: 1501.01635 [astro-ph.SR].

Mann, Andrew W., Niall R. Deacon, et al. (June 2014). "Prospecting in Ultracool Dwarfs: Measuring the Metallicities of Mid- and Late-M Dwarfs". In: *The Astronomical Journal* 147.6, 160, p. 160. DOI: 10.1088/0004-6256/147/6/160. arXiv: 1403.5560 [astro-ph.SR].

Ness, M. et al. (July 2015). "The Cannon: A data-driven approach to Stellar Label Determination". In: *The Astrophysical Journal* 808, 16, p. 16. DOI: 10.1088/0004-637X/808/1/16. arXiv: 1501.07604 [astro-ph.SR].

Newberg, H. J. et al. (Aug. 2012). "The LAMOST Spectroscopic Survey of Milky Way Stars (LEGUE)". In: *Galactic Archaeology: Near-Field Cosmology and the Formation of the Milky Way*. Ed. by W. Aoki et al. Vol. 458. Astronomical Society of the Pacific Conference Series, p. 405.

Newton, E. R., D. Charbonneau, J. Irwin, Z. K. Berta-Thompson, et al. (Jan. 2014). "Near-infrared Metallicities, Radial Velocities, and Spectral Types for 447 Nearby M Dwarfs". In: *The Astronomical Journal* 147, 20, p. 20. DOI: 10.1088/0004-6256/147/1/20. arXiv: 1310.1087 [astro-ph.SR].

Newton, E. R., D. Charbonneau, J. Irwin, and A. W. Mann (Feb. 2015). "An Empirical Calibration to Estimate Cool Dwarf Fundamental Parameters from H-band Spectra". In: *The Astrophysical Journal* 800, 85, p. 85. DOI: 10.1088/0004-637X/800/2/85. arXiv: 1412.2758 [astro-ph.SR].

Petigura, E. A. (2015). "Prevalence of Earth-size Planets Orbiting Sun-like Stars". PhD thesis. University of California, Berkeley.

Piskunov, N. and J. A. Valenti (Jan. 2017). "Spectroscopy Made Easy: Evolution". In: *Astronomy and Astrophysics* 597, A16, A16. DOI: 10.1051/0004-6361/201629124. arXiv: 1606.06073 [astro-ph.IM].

Rojas-Ayala, B. et al. (Apr. 2012). "Metallicity and Temperature Indicators in M Dwarf K-band Spectra: Testing New and Updated Calibrations with Observations of 133 Solar Neighborhood M Dwarfs". In: *The Astrophysical Journal* 748, 93, p. 93. DOI: 10.1088/0004-637X/748/2/93. arXiv: 1112.4567 [astro-ph.SR].

Sneden, C. A. (1973). "Carbon and Nitrogen Abundances in Metal-Poor Stars." PhD thesis. The University of Texas at Austin.

Steinmetz, M. et al. (Oct. 2006). "The Radial Velocity Experiment (RAVE): First Data Release". In: *The Astronomical Journal* 132, pp. 1645–1668. DOI: 10.1086/506564. eprint: astro-ph/0606211.

Valenti, J. A., D. Fischer, et al. (Sept. 2009). "Two Exoplanets Discovered at Keck Observatory". In: *The Astrophysical Journal* 702, pp. 989–997. DOI: 10.1088/0004-637X/702/2/989. arXiv: 0908.1612 [astro-ph.EP].

Valenti, J. A. and D. A. Fischer (July 2005). "Spectroscopic Properties of Cool Stars (SPOCS). I. 1040 F, G, and K Dwarfs from Keck, Lick, and AAT Planet Search Programs". In: *The Astrophysical Journal Supplement Series* 159, pp. 141–166. DOI: 10.1086/430500.

Valenti, J. A. and N. Piskunov (Sept. 1996). "Spectroscopy made easy: A new tool for fitting observations with synthetic spectra." In: *Astronomy and Astrophysics Reviews* 118, pp. 595–603.

Vogt, S. S. et al. (June 1994). "HIRES: the high-resolution echelle spectrometer on the Keck 10-m Telescope". In: *Instrumentation in Astronomy VIII*. Ed. by D. L. Crawford and E. R. Craine. Vol. 2198. Proceedings of the SPIE, p. 362. DOI: 10.1117/12.176725.

von Braun, K. et al. (Mar. 2014). “Stellar diameters and temperatures - V. 11 newly characterized exoplanet host stars”. In: *Monthly Notices of the Royal Astronomical Society* 438, pp. 2413–2425. DOI: 10.1093/mnras/stt2360. arXiv: 1312.1792 [astro-ph.EP].

Yee, S. W., E. A. Petigura, and K. von Braun (Feb. 2017). “Precision Stellar Characterization of FGKM Stars using an Empirical Spectral Library”. In: *The Astrophysical Journal* 836, 77, p. 77. DOI: 10.3847/1538-4357/836/1/77. arXiv: 1701.00922 [astro-ph.SR].

Chapter 3

STELLAR COMPANIONS TO TESS OBJECTS OF INTEREST: A TEST OF PLANET-COMPANION ALIGNMENT

Behmard, Aida, Fei Dai, and Andrew W. Howard (Apr. 2022). "Stellar Companions to TESS Objects of Interest: A Test of Planet-Companion Alignment". In: *The Astronomical Journal* 163.4, 160, p. 160. DOI: 10.3847/1538-3881/ac53a7. arXiv: 2202.01798 [astro-ph.EP].

Abstract

We present a catalog of stellar companions to host stars of TESS Objects of Interest (TOIs) identified from a marginalized likelihood ratio test that incorporates astrometric data from the *Gaia* Early Data Release 3 catalog (EDR3). The likelihood ratio is computed using a probabilistic model that incorporates parallax and proper motion covariances and marginalizes the distances and 3D velocities of stars in order to identify comoving stellar pairs. We find 172 comoving companions to 170 non-false positive TOI hosts, consisting of 168 systems with two stars and 2 systems with three stars. Amongst the 170 TOI hosts, 54 harbor confirmed planets that span a wide range of system architectures. We conduct an investigation of the mutual inclinations between the stellar companion and planetary orbits using *Gaia* EDR3, which is possible because transiting exoplanets must orbit within the line-of-sight, thus stellar companion kinematics can constrain mutual inclinations. While the statistical significance of the current sample is weak, we find that 73^{+14}_{-20}% of systems with *Kepler*-like architectures ($R_P \leq 4\ R_\oplus$ and $a < 1$ AU) appear to favor a non-isotropic orientation between the planetary and companion orbits with a typical mutual inclination α of $35 \pm 24°$. In contrast, 65^{+20}_{-35}% of systems with close-in giants ($P < 10$ days and $R_P > 4\ R_\oplus$) favor a perpendicular geometry ($\alpha = 89 \pm 21°$) between the planet and companion. Moreover, the close-in giants with large stellar obliquities (planet-host misalignment) are also those that favor significant planet-companion misalignment.

3.1 Introduction

Nearly half of FGK stars harbor stellar companions (Raghavan et al., 2010). If exoplanet hosts adhere to this trend, it is likely that many harbor yet undetected companions whose presence may instigate dynamical processes that drive planet migration, result in misaligned and/or eccentric planetary orbits, or shape system architectures in other ways. Thus, placing observational constraints on the properties of planet host systems with companions is essential for building a complete picture of planet formation and evolution.

Follow-up high-contrast imaging surveys have detected companions to many planet hosts from the *Kepler*, *K2*, and Transiting Exoplanet Survey Satellite (TESS) missions (e.g., Dressing et al. 2014; Wang et al. 2015; Kraus et al. 2016; Matson et al. 2018; Ziegler, Tokovinin, Briceño, et al. 2020; Ziegler, Tokovinin, Latiolais, et al. 2021). More recently, high-precision astrometric measurements from *Gaia* have made robust identification of comoving stars possible, resulting in large catalogs of stellar companions to planet host stars (e.g., Gaia Collaboration, Brown, Vallenari, Prusti, de Bruijne, Babusiaux, Bailer-Jones, et al. 2018; Gaia Collaboration, Brown, Vallenari, Prusti, de Bruijne, Babusiaux, Biermann, et al. 2021). For example, Mugrauer (2019), Mugrauer and Michel (2020), Michel and Mugrauer (2021), and Mugrauer and Michel (2021) identified companions to confirmed and candidate exoplanet hosts by establishing criteria for proper motions and parallaxes that when met constitute high confidence that a pair of stars is gravitationally bound.

Other studies have established probabilistic models to identify stellar companions; Oh et al. (2017) presented a model that calculates a likelihood ratio corresponding to the probability that two stars are comoving given their *Gaia* astrometric measurements, and applied the model to the first *Gaia* data release (DR1) to identify high-confidence comoving pairs. The likelihood ratio corresponds to the ratio of two hypotheses: (1) that a pair of stars share the same physical (3D) velocity, and (2) that the two stars have independent 3D velocities. The model treats uncertainties associated with *Gaia*-measured proper motion and parallax covariances in its assessment of whether two stars share the same 3D velocity. Thus, it is more robust than other methods that ignore covariances and assess whether two stars are comoving based on differences in their astrometric motions alone.

In this study, we used the Oh et al. (2017) probabilistic model to produce a catalog of *Gaia* EDR3 stellar companions to the current list of TESS Objects of Interest (TOIs). As the natural successor to the *Kepler* mission, TESS is currently carrying out high precision time series photometric observations for bright stars across ~85% of the sky, and has already discovered thousands of TOIs and dozens of confirmed planets (Ricker et al., 2015). Follow-up observations that confirm TOI planet candidates are ongoing, making the TOI catalog the source for the fastest growing sample of newly identified planet hosts.

After assembling a list of TOI hosts with comoving stellar companions, we investigated the degree of alignment between the planetary and companion orbits. We used precise astrometric measurements from *Gaia* EDR3 to compute the angle between the relative 2D position and velocity vectors of the two comoving stars in the sky plane. We then modeled the observed distribution of angles to constrain the mutual inclinations between the planet and companion orbits, a previously unexplored geometric property of planetary systems.

This paper is structured as follows. We provide details on the TOI catalog and our sample selection involving the Oh et al. (2017) probabilistic model in Section 3.2. In Section 3.3 we present our catalog of high-confidence TOI stellar companions obtained from applying the model to *Gaia* EDR3. We highlight TESS systems with stellar companions that host confirmed planets in Section 3.4, and investigate possible alignment between stellar companion and planetary orbits in Section 3.5. We discuss these results further in Section 3.6.

3.2 TOI Sample and Companion Selection

We constructed our TOI host sample from the 2241 TOIs reported in Guerrero (2020) and available on the ExoFOP-TESS[1] database. The TOIs correspond to 2140 unique stellar hosts spanning TESS Sectors 1–26 whose properties are drawn from the TICv7 (Sectors 1–13) and TICv8 (Sector 14 and onward), the two most up-to-date versions of the TESS Input Catalog (TIC) (Stassun et al., 2019). We searched for stellar companions to these 2140 TOI hosts within the *Gaia* EDR3 catalog, the latest data release resulting from the *Gaia* mission. *Gaia* EDR3 contains ~1 million new sources compared to *Gaia* DR2 and features improvements in parallax and proper motion precision of 30% and a factor of two, respectively, as well as decreased systematic uncertainties for parallaxes and proper motions by 30-40% and a factor of ~2.5, respectively. The photometric precision and homogeneity across color, magnitude, and celestial position are also improved (Gaia Collaboration, Brown, Vallenari, Prusti, de Bruijne, Babusiaux, Biermann, et al., 2021).

We employed a search radius of 10 arcmin around each TOI host to search for stellar companions. We then applied initial cuts of global parallax signal-to-noise $\bar{\omega}/\sigma_{\bar{\omega}} > 4$, distance agreement $2|r_1 - r_2|/(r_1 + r_2)$ to within 200%, and (point estimate)

[1]https://exofop.ipac.caltech.edu/tess/

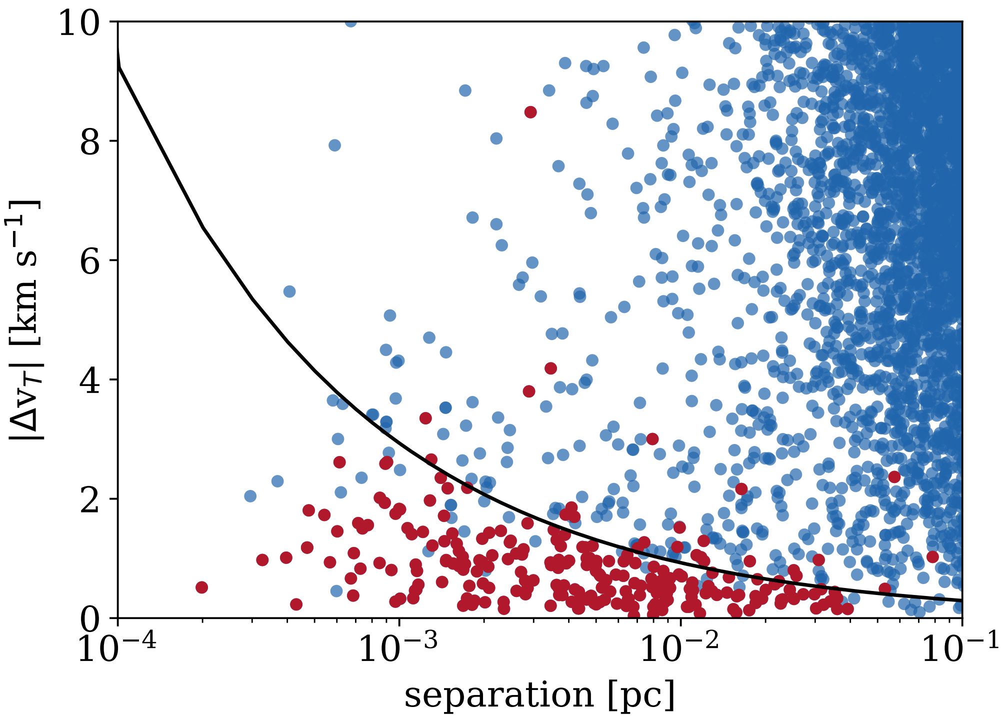


Figure 3.1: The tangential velocity differences ($\Delta|\boldsymbol{v}_T|$) and 2D separations for neighboring stars relative to their TOI hosts (blue points). The black line shows the Keplerian velocity as a function of semimajor axis for a binary pair with a total mass of $2M_\odot$ assuming circular orbits. The bound companions identified by our statistical analysis are shown in red.

tangential velocity differences less than $|\Delta \boldsymbol{v}_T| < 150$ km s^{-1}. These preliminary cuts are generous enough to allow for recovery of all stellar companions reported in Mugrauer (2019); their purpose is to trim down the list of potential TOI host companions without removing any true companions before applying the probabilistic model. We also note that the observational uncertainties on parameters involved in these cuts are small enough to be disregarded.

The tangential velocity calculation incorporates a point estimate of the distance derived from a correction to the Lutz-Kelker bias (Lutz and Kelker, 1973):

$$\hat{r} = 1000 \left[\frac{\bar{\omega}}{2} \left(1 + \sqrt{1 - \frac{16}{[\mathrm{S/N}]_{\bar{\omega}}^2}} \right) \right]^{-1} \text{ pc} \tag{3.1}$$

where $\bar{\omega}$ is the parallax in mas. The tangential velocity between two stars is estimated as

$$|\Delta \boldsymbol{v}_T| = |\hat{r}_1 \mu_1 - \hat{r}_2 \mu_2|, \tag{3.2}$$

where the proper motion vector is $\mu = (\mu_\alpha^{*}$[2], $\mu_\delta)$. We were left with ~1,200,000 possible stellar companions following these cuts, and plot their tangential velocity differences and separations relative to their corresponding TOI hosts in Figure 3.1. We found that there is a population of stellar pairs with small separations (<1 pc) and tangential velocity differences (<2 km s^{-1}) that are likely gravitationally bound, also noted by Oh et al. (2017) for the *Gaia* DR1 sample.

We employed the probabilistic model presented in Oh et al. (2017) that incorporates reported uncertainties in the *Gaia* astrometric data to yield a likelihood $\mathcal{L}_1$ that a given pair of stars is comoving based on their proper motions, distances, and 3D velocities. Similarly, the likelihood that the two stars are not comoving $\mathcal{L}_2$ can be computed and compared with $\mathcal{L}_1$ to identify truly bound pairs. For more details on the model, see Oh et al. (2017). While Oh et al. (2017) established a log-likelihood ratio value of $\ln(\mathcal{L}_1/\mathcal{L}_2) > 6$ as the threshold for high-confidence co-moving pairs, less than half of the Mugrauer (2019) stellar companions meet this threshold, with some members of the sample yielding log-likelihood ratios as low as $\ln(\mathcal{L}_1/\mathcal{L}_2) = -30$. This led us to suspect that the Oh et al. (2017) log-likelihood ratio cut was too stringent for our sample. We inspected the Mugrauer (2019) sample within the *Gaia* EDR3 database, and found that many stellar companions have *Gaia* EDR3 uncertainties on astrometric data that are less than 1% of the associated proper motion or parallax value. Such underestimated uncertainties can artificially lower the probabilistic model confidence that a pair of stars is bound. To address this, we included a "jitter" term in the proper motion and parallax uncertainties constituting 5% of the absolute proper motion value, and similarly 5% of the parallax value. This jitter term accounts for unknown systematic effects that may arise from factors such as the *Gaia* solution single-star model that fails to describe comoving binary systems. For example, TOI-837A and B have a spuriously large 3D separation of 6.6 ± 2.1 pc derived from *Gaia* quantities because the incorrect single-star model was assumed (Bouma et al., 2020), as further discussed in Section 3.4.

We ran the probabilistic model on the ~1,200,000 possible TOI host companions and recovered ~60,000 with $\ln(\mathcal{L}_1/\mathcal{L}_2) > 0$. The purpose of this generous threshold is to retain as many true comoving systems as possible. Most of these pairs have huge projected separations on the order of 10^4–10^5 AU, making it unlikely that they

[2] μ_α^* is the proper motion component in the RA direction, $\mu_\alpha^* = \mu_\alpha \cos\delta$.

are truly bound. To remove such pairs, we applied cuts of projected separations <10,000 AU and $|\Delta v_T| < 20$ km s^{-1} as motivated by the Mugrauer (2019) sample of companions whose separations and $|\Delta v_T|$ fall within 9100 AU and 20 km s^{-1}, respectively. We did not search separations beyond 10,000 AU because such wide companions are rare and not usually predicted to remain bound throughout dynamical evolution within clusters (Parker et al., 2009; Raghavan et al., 2010). We were left with ~1500 potential companions following these cuts. As mentioned earlier, the *Gaia* solution single-star model may introduce systematic errors in astrometric quantities that evade simple numerical cuts. To identify systems with such systematic errors, we visually inspected the potentially comoving systems individually, eliminating those that may still be unbound as evinced by large differences between their astrometric motions ($\Delta\mu_\alpha$ and $\Delta\mu_\delta \gtrsim 2-3$ mas yr^{-1}) and distances ($\Delta r \gtrsim 50$ pc). Because our analysis does not involve a completeness study, these non-systematic cuts should not affect our results.

Finally, we identified systems that likely harbor unresolved companions using their *Gaia* EDR3 re-normalized unit weight error (RUWE) values. Sources with RUWE > 1.4 are generally assumed to not have well-behaved *Gaia* astrometric solutions assuming they are single sources (Lindegren, 2018). However because our sample is composed of systems with detected stellar companions, their RUWE values may be high because of companion contamination, making a single cut at RUWE < 1.4 inappropriate. Belokurov et al. (2020) analyzed a set of known spectroscopic binaries with RUWE values, and found that RUWE rises steeply for binaries at separations <1.5″, marking this as the maximum angular separation at which companions can contaminate each other. Belokurov et al. (2020) also found that systems with tangential velocity differences exceeding their escape velocities have large RUWE, indicating an additional, closely separated companion. Thus, to identify which of our systems likely harbor unresolved companions, we made cuts on systems with separations >1.5″ that have TOI host or companion RUWE > 1.4, and tangential velocity differences greater that their escape velocities. We subsequently removed the 9 systems that met this criteria from our sample. We were ultimately left with 238 high-confidence companions to 234 of the original 2140 TOI hosts, yielding a raw (without completeness corrections) detection rate of ~10.9%. These companions identified through our statistical analysis are plotted in red in Figure 3.1.

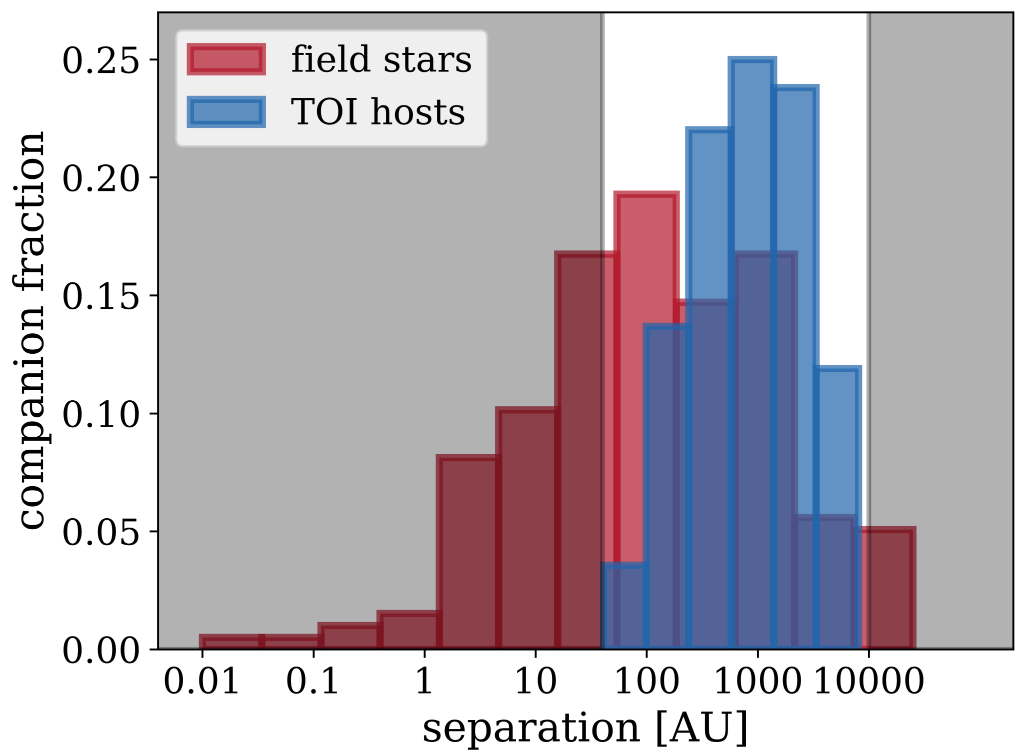


Figure 3.2: The distribution of projected separations for the 170 non-false positive TOI hosts and their 172 detected stellar companions (blue). The red histogram corresponds to the sample of companions to solar-like field stars reported in Raghavan et al. (2010). The separations of TOI hosts and companions span 40 to 10,000 AU (our search limit, delineated by gray shading), though most companions are found within 300–4000 AU. The lack of TOI host companions at separations <40 AU is due to the *Gaia* resolution limit of 0.7″, though the number of missed companions may be negligible (see Section 3.6).

3.3 Stellar Companions

The majority of the 234 TOI hosts with detected companions exhibit only one companion, comprising 230 systems with two stars. There are also 4 triple systems. Of the 234 TOI hosts with companions, 170 have non-‘FP’ (false positive), ‘FA’ (false alarm), or ‘APC’ (ambiguous planetary candidate) TESS and TESS Follow-up Observing Program Working Group (TFOPWG) dispositions whose properties are given in Table 3.1. The 170 TOI hosts currently marked as non-false positives together harbor 172 stellar companions whose properties are given in Table 3.2. Amongst the non-false positive TOI systems, there are 168 binary systems and 2 triple systems. We noted that the TIC 37770169 triple system does not appear hierarchical based on relative separations. This system may belong to a larger group

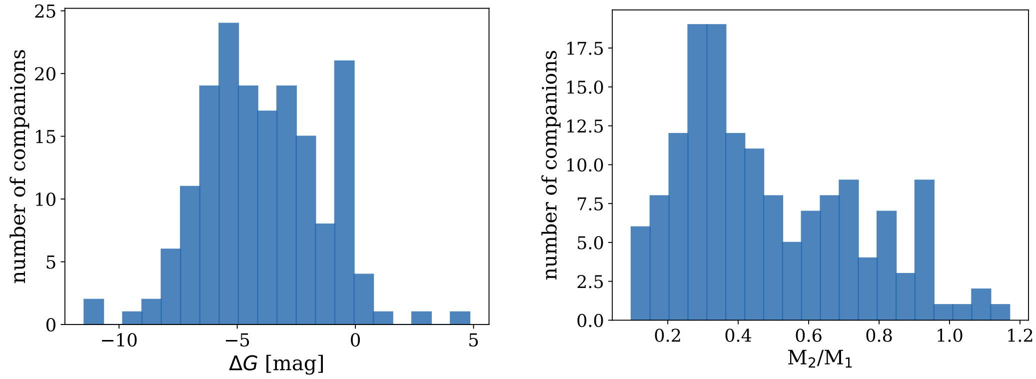


Figure 3.3: The distribution of ΔG (left) and mass ratios (right) for all 172 detected companions compared to their non-false positive TOI hosts. The companions are comparatively fainter than the TOI hosts, with ΔG ranging from (-11.5)–$(+4.8)$ and peaking at ~-5.0. The mass ratio distribution of the companions relative to the TOI hosts spans ~0.10–1.17 and peaks at ~0.3, identifying approximately half of the companions as low-mass main sequence stars.

of co-moving stars, but failed to match with any young associations according to the Bayesian classifier BANYAN Σ (Gagné et al., 2018).

The projected separations between the TOI hosts and their companions span 40 to 10,000 AU (our search limit), though most companions are found within 300-4000 AU (Figure 3.2). The G-band magnitudes for the detected companions compared to the TOI hosts are usually fainter; the distribution of ΔG ranges from (-11.5)–$(+4.8)$ and peaks at approximately -5.0 (Figure 3.3, left panel). 42 of the companions to non-false positive TOI hosts are sufficiently bright for radial velocity (RV) follow-up ($G < 13$ mag). In addition, 10 satisfy criteria required for a high level of chemical homogeneity with their TOI hosts ($\Delta T_{\rm eff} < 200$ K and $\Delta G < 0.3$ mag) (Andrews et al., 2019), making them stellar twins potentially useful for studies involving differential stellar abundances. We computed TOI host and stellar companion masses for the full sample with the `isoclassify` code that performs stellar classification with isochrone fitting by making use of astrometric and magnitude information (Huber, 2017). `isoclassify` failed to predict masses for 35 of the 238 stellar companions (~15% fail rate), likely because of close companion separations that result in blending. The predicted masses for the remaining 203 stellar companions are in the range ~0.11–1.51 $M_\odot$, with ~55% exhibiting masses below 0.5 $M_\odot$. The TOI host-companion mass ratios span ~0.1–1.17 and peak at ~0.3, indicating that the stellar companions are predominantly low-mass main sequence stars (Figure 3.3, right panel).

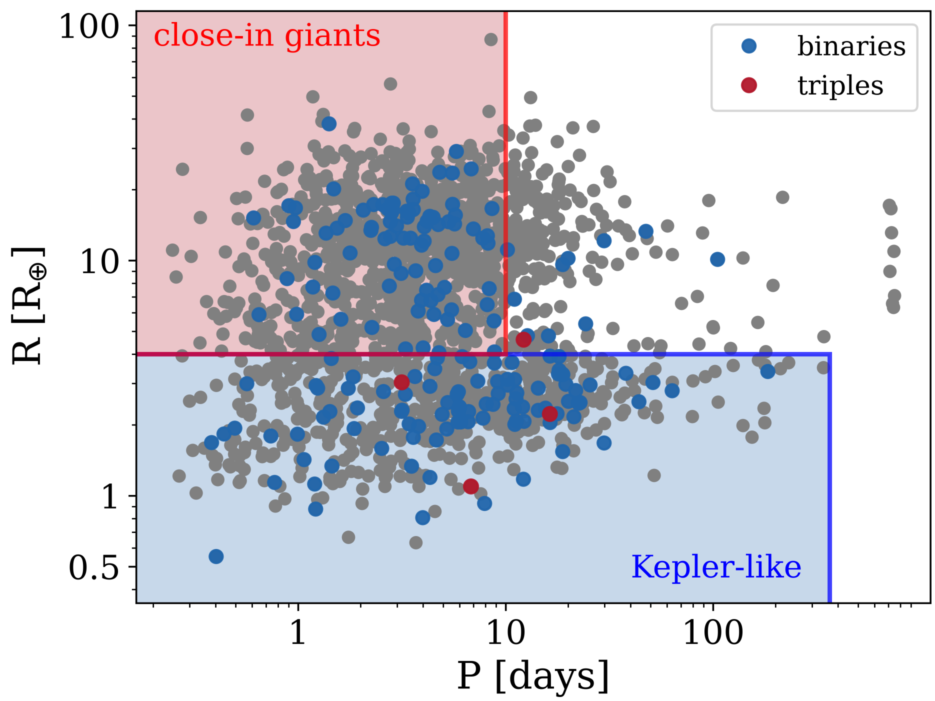


Figure 3.4: The total sample of 2140 TOIs plotted in period-radius space (gray). Non-false positive TOIs with one detected stellar companion (binary systems) and two detected companions (triple systems) are overplotted in blue and red, respectively. There does not appear to be a preferred location for TOIs with stellar companions in period-radius space. 6 TOIs fall outside this period-radius space and are not shown for easier visualization. We also delineate the regions of parameter space that harbor *Kepler*-like (blue) and close-in giant systems (red).

To construct a planet-less comparison sample, we drew from the TICv8 Candidate Target List (CTL). A subset of the TICv8 CTL targets are selected for TESS observations based on their ranked priorities that incorporate the probability of detecting small planet transits considering the host star radius, the total expected photometric precision, and the number of TESS Sectors that may include the target. For more details on the TICv8 CTL priority calculation, see Stassun et al. 2019. We selected TICv8 CTL targets with priorities above 0.005 (yielding <200,000 stars) and $T_{\rm eff}$, mass (M_*), radius (R_*), distance (r), and G-band magnitude within the 1-σ bounds of the 2140 TOI system hosts. The TOI host stellar parameters were computed with `isoclassify` which failed on 56 of the 2140 TOI hosts, again likely because of close companion separations that result in blending, leaving 2084 TOI hosts for this analysis. A search in the TICv8 CTL using the 1-σ bounds of the TOI systems yielded ~19,900 targets. This sample was further trimmed down with the following

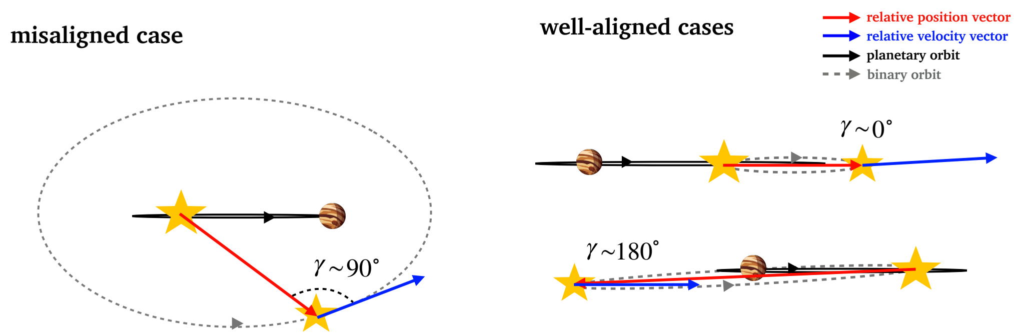


Figure 3.5: A schematic illustrating the "Linear Motion Parameter" γ, i.e., the angle between the relative 2D position and velocity vectors of the TOI and its comoving stellar companion in the sky plane. Because the transiting planet must follow a nearly edge-on orbit, γ will favor different values depending on the mutual inclination between the orbits of the planet and companion. In the well-aligned case of small mutual inclination, the companion will likely also have an edge-on orbit, with γ near 0° or 180°. In the misaligned case of large mutual inclination, the binary orbit will tend to be face-on with γ near 90°.

similarity metric:

$$D^2 = \left(\frac{\Delta T_{\rm eff}}{\sigma_{T_{\rm eff}}}\right)^2 + \left(\frac{\Delta M_*}{\sigma_{M_*}}\right)^2 + \left(\frac{\Delta R_*}{\sigma_{R_*}}\right)^2 + \left(\frac{\Delta r}{\sigma_r}\right)^2 + \left(\frac{\Delta \rm G}{\sigma_{\rm G}}\right)^2 \tag{3.3}$$

which incorporates the relative $T_{\rm eff}$, M_*, R_*, r, and G-band magnitude between the two stars, and their associated errors added in quadrature. We constructed the comparison sample by selecting the five CTL targets with the smallest comparison metric value relative to each TOI host in the 2084 TOI host sample. After removing duplicates, we were left with 4146 TICv8 CTL stars. We then recovered 703 stellar companions to 679 of the 4146 TICv8 CTL targets searched within *Gaia* EDR3 with the same procedure we used to identify stellar companions to TOI systems. This sample of companions to planet-less CTL stars appears to be a good comparison sample for the TOI host companions by virtue of exhibiting similar distributions in parameters such as separation and brightness difference.

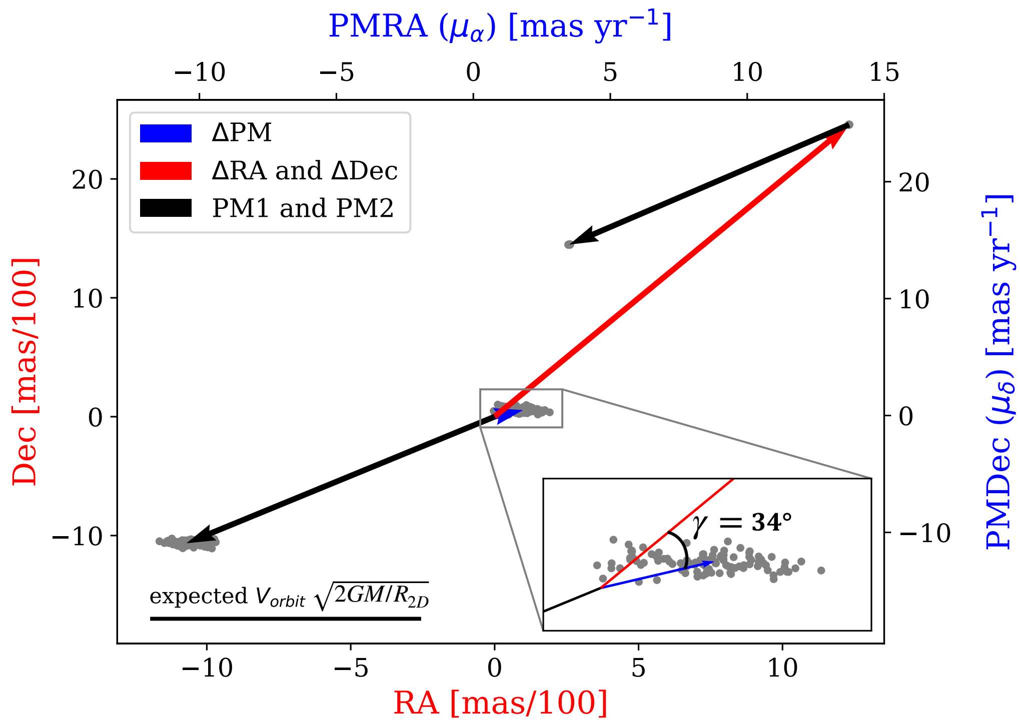


Figure 3.6: The differential position (red), differential velocity (blue), and individual velocity vectors (black) in the sky plane for one of the 155 binary TOI systems used in our alignment analysis (TOI-1473), derived from the equatorial coordinates (RA and Dec) and proper motions in the RA (μ_α) and Dec (μ_δ) directions. The velocity vectors are in units of mas yr^{-1} while the position vector is in units of mas/100 for easier visualization. The position and velocity vector scatter is derived from sampling the covariance matrix for 100 iterations, and is represented by the clouds of gray dots. The zoomed inset displays how γ is measured. The expected orbital velocity (mas yr^{-1}) assuming circular orbits is shown for comparison and represented by the length of the black line in the lower left corner.

3.4 TOI and Confirmed Planets

The 234 TOI hosts with detected stellar companions host a total number of 254 TOIs whose properties are reported in Table 3.3. We plot these 254 TOIs with the total sample of 2140 TOIs in Figure 3.4 using the planetary orbital periods and radii.

To determine how many of our TOI hosts with stellar companions harbor confirmed planets, we selected those with TESS and TFOPWG dispositions of ‘CP’ (confirmed planet) or ‘KP’ (known planet), and cross-matched our catalog with the NASA Exoplanet Archive[3] (Akeson et al., 2013) and the living catalog of confirmed planets

[3] https://exoplanetarchive.ipac.caltech.edu/

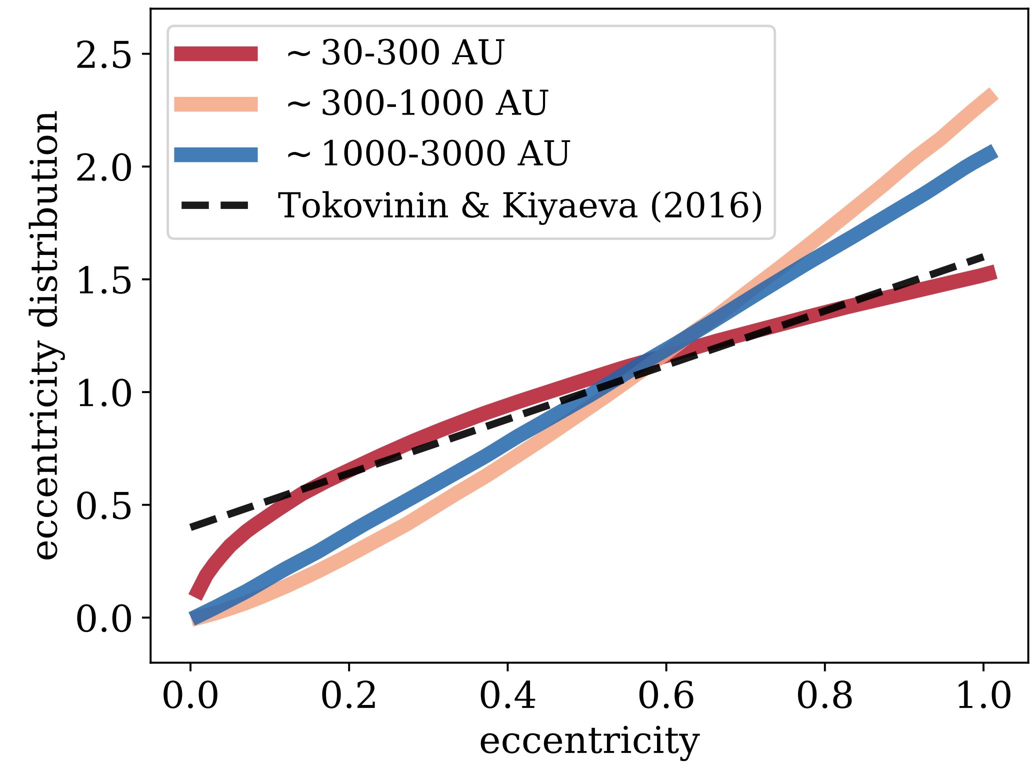


Figure 3.7: The eccentricity distribution model derived by Tokovinin and Kiyaeva (2016) (black dashed) and the eccentricity distributions from Hwang, Ting, and Zakamska (2022) for binary separations ranging from ∼30−300 AU (red), ∼300−1000 AU (orange), and ∼1000−3000 AU (blue). The Tokovinin and Kiyaeva (2016) model is in good agreement with the Hwang, Ting, and Zakamska (2022) distributions in this binary separation regime of ∼30−3000 AU.

discovered by TESS[4]. As of January 9, 2022, this yielded 99 confirmed planets amongst the 254 TOIs, marked with an "*" next to their corresponding TIC in Table 3.3. The confirmed systems span a wide range of architectures, including one planet in the Neptune desert (Díaz et al., 2020) and a young hot Jupiter system (Bouma et al., 2020). We discuss these systems in more detail below.

We detected one previously unknown stellar companion (*Gaia* EDR3 6520880036122448000, hereafter referred to as TOI-132B) to TOI-132, a G = 11.3 G dwarf that hosts a planet in the Neptune desert (Díaz et al., 2020). TOI-132B is faint (G = 18.4) and located at a projected separation of 3231 AU (19.6″). Díaz et al. (2020) also searched for nearby companions with speckle imaging using HRCam on the 4.1-m Southern Astrophysical Research (SOAR) telescope; they noted a potential compan-

[4]https://tess.mit.edu/publications//

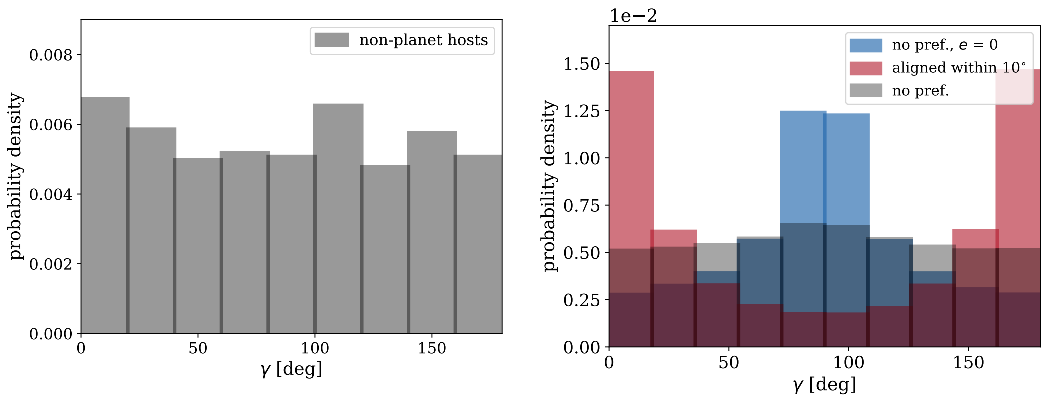


Figure 3.8: The left panel displays the normalized histogram of angles γ between the differential position and velocity vectors in the sky plane for the 512 TICv8 CTL targets in our comparison sample that pass our $\gamma_{\rm err} < 5$ degrees cut (gray). The right panel displays the theoretical γ distributions corresponding to an isotropic distribution of α with circular orbits (blue) or eccentric orbits (gray). The red distribution corresponds to the case of mutual inclination α between the transiting planet and companion agreeing within 10 degrees. Note that the planet-less comparison sample shows a rather uniform distribution of γ as one would expect for an isotropic distribution that also accounts for orbital eccentricities.

ion at 0.079″ but ultimately ruled it out based on visual inspection of the individual spectra.

We also detected one previously unknown stellar companion (*Gaia* EDR3 5251470948231619200, hereafter referred to as TOI-837B) at a close projected separation of 330 AU (2.3″) to TOI-837, a young hot Jupiter host (P = 0.83 days, $R_P = 0.77^{+0.09}_{-0.07}\ R_J$) (Bouma et al., 2020). TOI-837 is a member of the 35^{+11}_{-5} Myr southern open cluster IC 2602, making its hot Jupiter one of the youngest transiting planets currently known. Bouma et al. (2020) identified TOI-837B with *Gaia* DR2 imaging and noted it as another IC 2602 member, but ruled it out as a bound stellar companion based on its large 3D separation of 6.6 ± 2.1 pc (Bouma et al. 2020, *private correspondence*). However, the 3D separation calculation is quite sensitive to the *Gaia* distance estimate, which assumes a single-star model that likely introduced systematic errors in the parallax measurement, as discussed in Section 3.2. Our probabilistic analysis indicates that these two stars are bound with $\ln(\mathcal{L}_1/\mathcal{L}_2) \approx 4.09$.

3.5 Planet-Companion Orbit Alignment

We sought to constrain the relative orientations between planetary and stellar companion orbits in our sample, i.e., the mutual inclination between the two orbital axes,

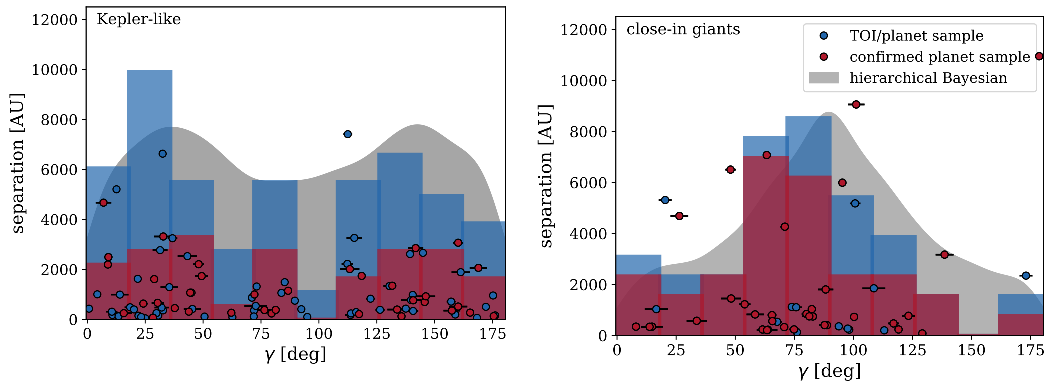


Figure 3.9: The right panel displays the scatter plot of binary separations and angles γ between the 2D differential position and velocity vectors for the TOI (blue) and confirmed planet (red) systems that harbor close-in giant planets (P < 10 days and $R_P > 4\ R_\oplus$). The left panel displays the binary separations and γ for the *Kepler*-like ($R_p \leq 4\ R_\oplus$ and a < 1 AU) TOI and confirmed planet systems. The γ distribution histograms are provided in the background of all plots and scaled for easier visualization. We also show the best-fit theoretical γ distributions derived from HBM (gray).

which we refer to as α. Using the *Gaia* EDR3 equatorial coordinates (RA and Dec) and proper motions, we were able to precisely measure the 2D relative position and velocity vectors in the sky plane between the planet host and its comoving stellar companion. The angle between these vectors, which we refer to as γ, encodes information on the true planet-companion mutual inclination α. Because all planets in our sample are transiting and thus have orbital inclinations close to 90°, it follows that if the planet and companion orbits are well-aligned (low α values), the orbital axis of the companion will also likely reside in the sky plane. In other words, if the companion follows an edge-on orbit, the angle between the 2D relative position and velocity vectors of the planet host and companion γ will be near 0° or 180°. In contrast, if the companion and transiting planet orbits are misaligned (large α values), the binary orbit will more likely reside in orientations approaching face-on with γ near 90° (Figure 3.5). Thus, the γ distribution can be used to deduce the underlying distribution of α. Figure 3.6 illustrates application of this method to TOI-1473. We note that the quantity γ or "Linear Motion Parameter" was also used by Tokovinin and Kiyaeva (2016) to constrain the eccentricity distribution of wide binaries (>30 AU). They modeled the eccentricity probability density distribution as $p(e) \approx 1.2e + 0.4$ with $\langle e \rangle = 0.59 \pm 0.02$, which we subsequently adopted for our investigation of mutual inclinations between planetary and companion orbits. Similarly, Hwang,

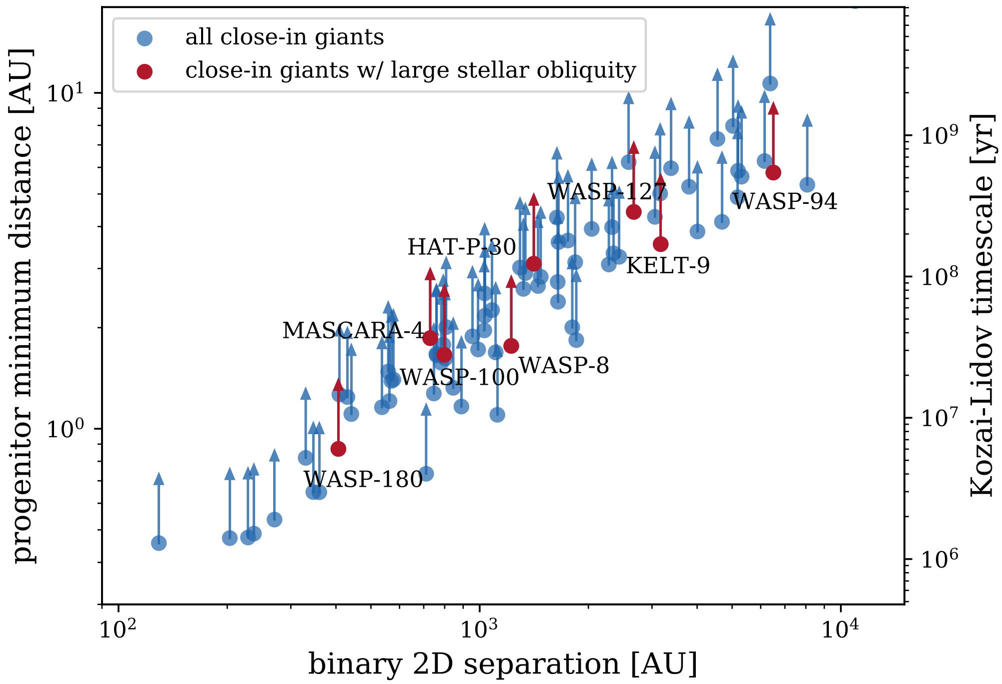


Figure 3.10: The minimum orbital distances of the close-in giant progenitors, with arrows representing the upper limit, versus the 2D projected separation between the planet host and comoving companion in a Kozai-Lidov formation scenario. The progenitors must initially have large enough orbital periods to ensure that Kozai-Lidov oscillations induced by the comoving companion are not suppressed by GR precession. The right y-axis denotes the Kozai-Lidov timescale for the comoving binaries at their observed 2D separations. We assumed a fixed eccentricity of 0.5 for the comoving binary, and fixed masses of 1 and 0.3 $M_\odot$ for the planet host and the comoving binary. These assumptions permit a one-to-one correspondence between the left and right axes. We highlight and label the planets that have both large stellar obliquities λ and mutual inclinations γ in red.

Ting, and Zakamska (2022) employed the 2D relative position and velocity vectors to derive eccentricity distributions of wide binaries, and found evidence for two distinct formation pathways operating at different separation regimes. The $p(e) \approx 1.2e + 0.4$ eccentricity distribution model from Tokovinin and Kiyaeva (2016) used in our alignment analysis is derived from a clean sample of 477 binaries that lack planets and have no preferential orientation. It also well-represents the eccentricity distributions derived by Hwang, Ting, and Zakamska (2022) for separations of ~100−3000 AU, which spans the bulk of separations in our sample (Figures 3.2 and 3.7). Thus, the Tokovinin and Kiyaeva (2016) model is an appropriate prior for our

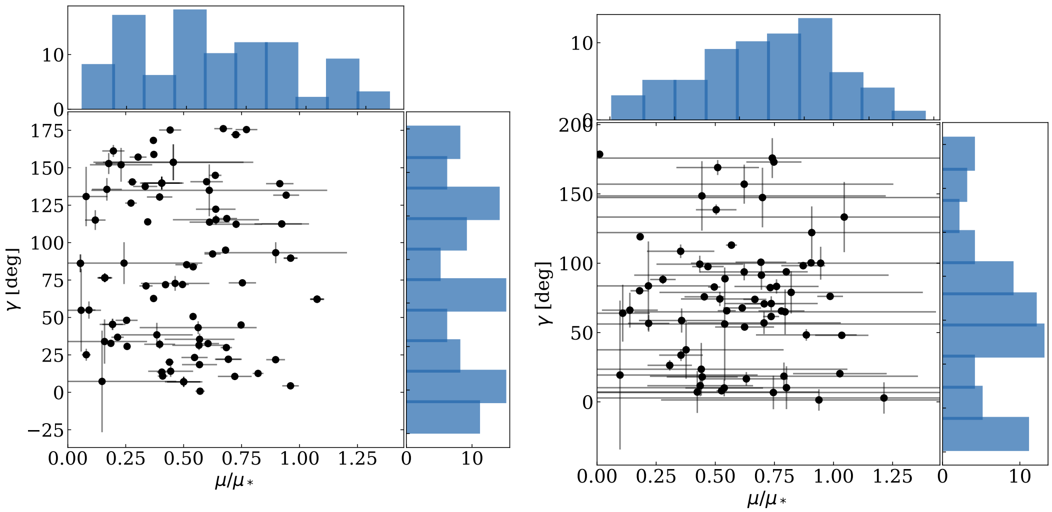


Figure 3.11: The distributions of angles γ and normalized relative motion μ' for the TOI and confirmed planet systems that harbor close-in giant planets ($P < 10$ days and $R_P > 4\ R_\oplus$) (right panel), and those with *Kepler*-like architectures ($R_p \leq 4\ R_\oplus$ and $a < 1$ AU) (left panel). The γ and μ' distributions appear uncorrelated.

entire sample of TOI systems with stellar companions.

We limited our analysis to the 168 binaries in our TOI sample as γ is only measurable in two-body systems. We verified that the 2D relative velocity between the two stars did not exceed their expected orbital velocity $2G(M_1 + M_2)/a$ considering their 2D projected separation and assuming circular orbits. Following this cut, we were left with 159 of the 168 binaries that together host 179 TOIs. To create a larger sample of planet/planet candidate-hosting binary systems, we added the Mugrauer (2019) confirmed exoplanet systems with stellar companions that pass our expected orbital velocity cuts (182 additional systems), yielding a total of 341 systems. We also constructed a comparison sample by applying our full set of cuts to the planet-less CTL systems, leaving 673 systems.

For the sample of 341 planet/planet candidate systems, as well as the planet-less CTL sample, we measured γ from the equatorial coordinates and proper motions, and estimated the uncertainty on γ by sampling from multi-dimensional Gaussians that correspond to covariance matrices reported by *Gaia* EDR3. Specifically, we took 100 random draws from the covariance matrices to generate 100 iterations of γ per system, and took the standard deviation as $\gamma_{\rm err}$. We then made a final cut to remove systems with $\gamma_{\rm err} > 5$ degrees, leaving 155 planet/planet candidate systems, and 512 planet-less CTL systems. We provide the γ distribution of the planet-less

CTL comparison sample in the left panel of Figure 3.8, which appears flat. The sample of 155 binaries with γ measurements served as the posterior sample for our hierachical Bayesian analysis outlined in the following section on hierachical Bayesian modeling. We were able to ignore the influence of transiting planets on the astrometry of comoving stars because they will not generate significant perturbations to the two-body Keplerian motion considering their masses and orbital periods.

We constructed theoretical γ distributions by generating 100,000 simulated systems with the Keplerian solver implemented in `orbitize` (Blunt et al., 2020). For each simulated system we sampled all Keplerian orbital elements uniformly in phase space and adopted the eccentricity distribution model $p(e) \approx 1.2e + 0.4$ reported by Tokovinin and Kiyaeva (2016). We assumed solar mass stars and 1000 AU separations. Supposing that the planets follow edge-on orbits ($i_p = 90°$), we can compute the orbital inclinations of the stellar companions with the mutual inclination α. The right panel of Figure 3.8 displays theoretical γ distributions generated under different assumptions. Specifically, the blue distribution corresponds to a sample of systems with circular companion orbits and no preferential alignment between the planet and companion. In this case, γ is strongly peaked at 90°. If we allow for eccentric orbits that follow $p(e) \approx 1.2e + 0.4$, γ is smeared out into a relatively uniform distribution between $0-180°$ as shown by the gray distribution. However if we assume that the planet and the companion orbits are well-aligned ($\alpha \lesssim 10°$), the distribution of γ exhibits a symmetric double peak pattern at 0° and 180° as expected from the schematic shown in Figure 3.5.

Hierarchical Bayesian Modeling

We employed a hierarchical Bayesian model (HBM) to translate the observed γ distribution of the TOI systems to constraints on the true mutual inclination α. We divided the TOI sample into two architecture sub-samples, namely *Kepler*-like systems featuring sub-Neptunes/super-Earths ($R_P \leq 4\ R_\oplus$ and a < 1 AU) versus single, close-in gas giants ($P < 10$ days and $R_P > 4\ R_\oplus$) (Figure 3.4). These architectures were chosen based on numerous observational and theoretical lines of evidence that suggest $Kepler$-like and close-in giant systems may have distinct formation pathways, e.g., giant planets strongly favor metal-rich environments (Fischer and Valenti, 2005) while *Kepler*-like systems form readily in lower metallicity environments (Petigura et al., 2018), and giant planets are often lonely and misaligned while *Kepler*-like planets frequently reside in multi-planet systems with low mutual inclinations (e.g., Winn and Fabrycky 2015 and references therein).

We employed an HBM (Hogg, Myers, and Bovy, 2010; Foreman-Mackey, Hogg, and Morton, 2014) to model the distribution of mutual inclinations between the planetary and stellar companion orbits. We considered two possible hypotheses for the α distributions:

1. The planetary and the stellar companion orbits are uncorrelated (no preferred α angle), and the resultant distribution of γ is approximately uniform as exemplified by the gray histogram in the right panel of Figure 3.8.

2. A certain fraction (f) of the planetary systems have a preferred orientation such that α can be approximated by a von Mises distribution with mean α_0 and κ parameter that encodes the width of the distribution, while the remaining 1-f of the systems follow an isotropic α distribution. The resultant distribution of γ will significantly deviate from uniformity.

We approximated the likelihood function using the γ samples computed from the covariance matrix of each system:

$$p(data|x) \propto \prod_{k=1}^{K} \frac{1}{N} \sum_{n=1}^{N} \frac{p(\gamma_{n,k}|x)}{p_0(\gamma)}, \tag{3.4}$$

where $data$ represents the observed distribution of γ, x is the set of hyperparameters describing the distribution of mutual inclinations α, K is the total number of observed systems, and N is the number of covariance samples. We converted the distribution of α described by the hyperparameters x to a distribution of γ by marginalizing over various Keplerian orbital elements. Our total set of nuisance and hyperparameters include eccentricity with the Tokovinin and Kiyaeva (2016) model taken as the prior; time of observation and argument of periapse modeled with uniform priors; stellar companion orbital inclination with a prior set by α and ϕ (an arbitrary azimuthal angle marginalized uniformly); and planetary orbital inclination, orbital period, and longitude of ascending node, all held fixed. After marginalization, we numerically evaluated $p(\gamma|x)$ and $p_0(\gamma)$, and sampled from the hyperparameter posterior distribution and Bayesian evidence Z simultaneously using the nested sampling code `MultiNest` (Feroz, Hobson, and Bridges, 2009). We then computed the Bayesian evidence for model comparison. We sampled the various parameters uniformly in phase space except for eccentricity, which we derived from the distribution $p(e) \approx 1.2e + 0.4$ (Tokovinin and Kiyaeva, 2016).

We first tested the planet-less CTL sample, and found that it favors an isotropic distribution of γ consistent with the first hypothesis. Testing the second hypothesis yielded a fraction of non-isotropic components f that converged towards 0 with a 95% upper limit of $f < 18\%$, indicating that an isotropic distribution is favored again. Overall, the planet-less CTL sample favors the first hypothesis with a Bayes factor of $\Delta log(Z) = 3.1$. Considering the TOI sample, we found that $Kepler$-like systems slightly favor the second hypothesis with $\Delta log(Z) = 1.1$ over the isotropic model. Specifically, for $73^{+14}_{-20}\%$ of the $Kepler$-like systems, the planet and companion orbits appear to favor alignment with an α of typically $35 \pm 24°$. Note that we combined the α_0 and κ posteriors together because both parameters describe the distribution of α. The close-in giants also favor the second hypothesis with $\Delta log(Z) = 2.0$, but in contrast appear to favor a perpendicular geometry, with $65^{+20}_{-35}\%$ of close-in giants exhibiting an α of typically $89 \pm 21°$. The best-fit γ distributions for the *Kepler*-like and close-in giant systems are provided in Figure 3.9. These results are weakly significant according to Jeffreys (1998) and Kass and Raftery (1995), which assert that $2.5 < log(Z) < 5$ indicates strong significance, and $1 < log(Z) < 2$ indicates positive significance. The low statistical significance of our results may be partially due to the small sizes of our $Kepler$-like (108 systems) and close-in giant (47 systems) samples. Moreover, γ only provides an indirect measurement of α. We further discuss these results and consider how the *Gaia* final data release may improve constraints in Section 3.6.

3.6 Discussion

Comoving Stars for Further Characterization

We present a new catalog of *Gaia* EDR3 stellar companions to the 2140 unique TOI hosts from TESS Sectors 1–26. We note that the *Gaia* resolution limit of 0.7″ allows for companion detection if projected separations are >40 AU at distances less than ~60 pc. This may have contributed to lower companion completeness in our catalog; nearly all our TOI hosts reside at farther distances, thus explaining our lack of detected stellar companions at these small separations (Figure 3.2). More specifically, 84% of our companions reside at separations of 300–4000 AU while only ~20.5% of field star companions detected by Raghavan et al. (2010) fall in this range. However, stellar companion surveys of planet host stars yield few companions at separations interior to ~100 AU because planet formation is suppressed by dynamical effects from close companions (Kraus et al., 2016; Moe and Kratter, 2021). This implies that the actual number of missed stellar companions to TOI hosts due to the *Gaia* resolution limit is likely negligible.

We found a total of 238 comoving stellar companions to 234 TOI hosts, yielding a raw companion detection rate of ~10.9% with respect to the total number of 2140 searched hosts. The 234 systems include 230 binaries (~10.7%) and 4 triple systems (~0.19%). These fractions are lower than those of solar-type field stars which exhibit binary and triple system fractions of 33 ± 2% and 8 ± 1%, respectively (Raghavan et al., 2010). While our lower binary and triple detection rates likely do not stem from the *Gaia* resolution limit, they may be affected by *Gaia* pipeline incompleteness, shortcomings in our probabilistic approach, or true astrophysical differences between the companion fractions of planet host stars versus field stars. The last possibility, while intriguing, cannot be entertained until the first two possibilities are ruled out.

If we consider only non-false positive systems, there are 172 companions to 170 TOI hosts. Amongst these systems are 10 stellar twin binaries potentially useful for future studies involving differential stellar abundances ($\Delta T_{\rm eff} < 200$ K and $\Delta G < 0.3$ mag), and 42 systems amenable to RV follow-up for planet detection ($G < 13$ mag). Additionally, 11 of the 172 companions exhibit masses below 0.15 $M_\odot$ and distances below 100 pc, with 3 within 25 pc. This demonstrates that the probabilistic framework used in this study can identify nearby low mass stellar companions. Such objects are valuable for surveys of stars within the Solar neighborhood and their hosted planets such as the RECONS (REsearch Consortium On Nearby Stars) survey within 25 pc (e.g., Henry, Kirkpatrick, and Simons 1994) and the TRAPPIST survey of planets around nearby ultra-cool dwarfs (Gillon et al., 2013).

Thick Disk Membership

Many of the TOI hosts with stellar companions have large tangential velocities that suggest Galactic thick disk membership. To investigate this, we computed their Galactic space motion velocities U, V, and W with the procedure detailed in Johnson and Soderblom (1987) assuming the local standard of rest from Coşkunoğlu et al. (2012). Using the methodology of Reddy, Lambert, and Allende Prieto (2006), we found that 3 TOI hosts with companions (TIC 166833457, TIC 175532955, TIC 23434737) have a >50% probability of belonging to the thick disk. These stars also have thick-to-thin disk probability ratios computed from the probabilistic framework of Bensby, Feltzing, and Lundström (2004) and Bensby, Feltzing, and Oey (2014) of TD/D = 4, 33 and 114, respectively (Carrillo et al., 2020). TIC 166833457 has already been confirmed as a thick disk member via extensive chemokinetic follow-up (Mancini et al., 2016; Southworth et al., 2020), and its hot Jupiter candidate confirmed as WASP-98B. Though TIC 23434737 and TIC 175532955 are not confirmed to reside in the thick disk, TIC 23434737 has a TICv8 metallicity of [Fe/H] = −0.39 ± 0.05, consistent with the thick disk population. Additionally, these stars host warm Neptune (TOI-1203) and hot super-Earth planet candidates (TOI-929), respectively, making potential thick disk membership particularly interesting as it would provide evidence that small, rocky planets are able to form in metal-poor environments and avoid being tidally destroyed around old stars (Buchhave et al., 2012; Hamer and Schlaufman, 2020).

Kozai-Lidov Migration

One possible formation scenario for hot Jupiters involves a stellar companion inducing Kozai-Lidov oscillations between the hot Jupiter progenitor and its host star, leading to high-eccentricity tidal migration of the planet to its current close-in location (e.g. Fabrycky and Tremaine, 2007). We considered if the companions to our close-in giants could have instigated such Kozai-Lidov migration. Ngo et al. (2016) noted that the occurrence rate of stellar companions to hot Jupiter hosts at separations of 50–2000 AU (47% ± 7%) is a factor of 2.9 higher than the rate for field stars in the same range. However, Ngo et al. (2016) also suggested that most of these companions are too far away to have instigated Kozai-Lidov migration. We performed a similar set of calculations for the close-in giants in our sample. For example, the timescale for Kozai-Lidov oscillations in the young hot Jupiter system TOI-837 is

$$\tau_{KL} = \frac{2P_b^2}{3\pi P_p^2}\frac{M_1 + M_2}{M_2}(1 - e_b^2)^{3/2} \tag{3.5}$$

(Kozai 1962; Kiseleva, Eggleton, and Mikkola 1998), where P_p is the original orbital period of the hot Jupiter progenitor around its host, M_1 and M_2 are the masses of the TOI host and stellar companion, respectively, and P_b and e_b are the period and eccentricity of the stellar companion. We calculated τ_{KL} using `isoclassify`-predicted masses for TOI-837A and TOI-837B. We then compared τ_{KL} to the timescale of general relativistic (GR) precession. This is relevant because Kozai-Lidov oscillations require a slow changing argument of perihelion that will not occur if GR precession is sufficiently fast. We computed the expected GR

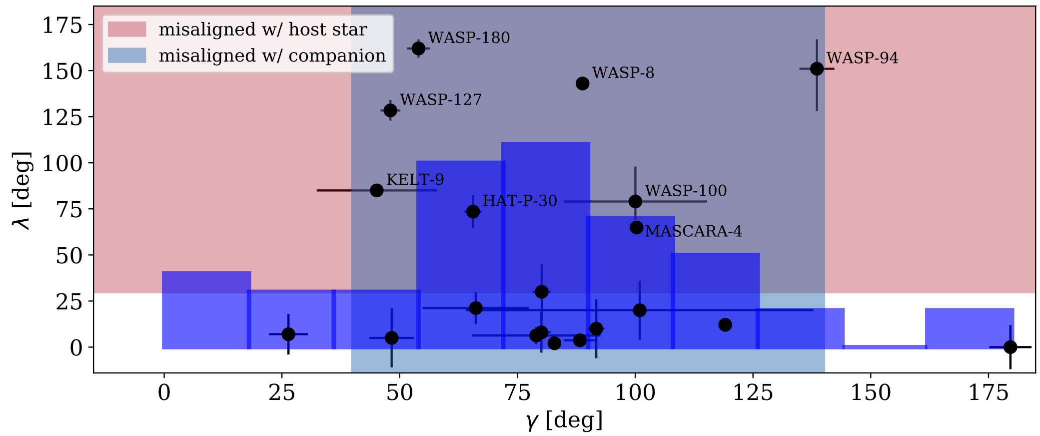


Figure 3.12: The sky-projected stellar obliquity λ and γ angles for the subset of 20 close-in giant planet systems in our sample with published λ values. We overplot the histogram of all close-in giants in our sample and scale it for easier visualization (blue). Close-in giants that are misaligned with respect to their host star (red region, $\lambda > 30°$) are often also misaligned with respect to their comoving companion (blue, $40 < \gamma < 140°$).

precession rates for TOI-837 as follows:

$$\dot{\omega}_{GR} = \frac{GM_*}{a_b c^2} \frac{3n_b}{G_b^2}, \tag{3.6}$$

where $n_b = 2\pi/P_b$, $G_b = \sqrt{1 - e_b^2}$, and c is the speed of light. The condition for Kozai-Lidov oscillations to be suppressed by relativistic precession is $\tau_{KL}\dot{\omega}_{GR} > 3$ (Fabrycky and Tremaine, 2007), which is well-satisfied by TOI-837b for a typical choice of unknown system parameters.

However, if we allow the progenitors of close-in giant planets to form along more distant orbits, the rate of GR precession can be suppressed by several orders of magnitude. Using the 2D projected separation and `isoclassify` stellar masses, we calculated the minimum orbital distances for the progenitor planets that ensure Kozai-Lidov oscillations are not quenched by GR precession. We found that the progenitors of close-in giants drawn from our sample of comoving binaries must have formed at minimum distances of ~0.5-10 AU for Kozai-Lidov oscillations to proceed (Figure 3.10). Such orbital distances are consistent with the progenitors starting as cold to warm Jupiters (e.g., Fulton et al., 2021; Dawson and Albrecht, 2021).

Planet-Companion Orbital Alignment

Stellar obliquity, or the angle between the planetary orbit and rotation axis of the host star, has been the subject of numerous exoplanet studies throughout the past decade. The diversity of the ~200 reported stellar obliquity measurements have revealed intriguing trends with respect to planet and stellar host properties (Winn, Fabrycky, et al., 2010; Mazeh et al., 2015; Louden et al., 2021) that have far-reaching implications for planet formation, migration, orbit tilting, and tidal realignment (Winn and Fabrycky, 2015). In this work, we placed constraints on a similar yet distinct geometric property of planetary systems, i.e., the alignment between a planetary orbit and that of a comoving companion star. We measured the angle γ between the 2D relative position and velocity vectors of the planet host and companion in the sky plane. This angle γ cannot be translated to the true mutual inclination α on a system-by-system level due to its dependence on other orbital elements, particularly the orbital eccentricity. However on a population level, it is possible to marginalize over various Keplerian orbital elements and subsequently deduce the underlying α distribution from the observed distribution of γ.

The 2D relative velocity vector magnitude can potentially provide additional information on the true mutual inclination α. To investigate this, we computed the normalized relative motion $\mu^{'} = \mu/\mu^{*}$ for each system, where μ is the proper motion magnitude and μ^{*} is the relative orbital motion (Tokovinin and Kiyaeva, 2016). In Figure 3.11, we show the measured 2D distribution of $\mu^{'}$ and γ. $\mu^{'}$ is subject to more measurement uncertainty (e.g., the total mass of the binary system). Moreover, our sample size is too small to warrant 2D analysis. We chose to focus on the more informative γ distribution.

Before assessing the observed γ distributions, we consider a related question: do we expect any correlation between the spin axis of a star and the orbit of its comoving companion? Considering a simple core accretion model, one might assume that planets form within a protoplanetary disk that is well-aligned with the spin axis of the host star. If there is a distant comoving star in the system, would the companion orbit also be coplanar with the inner planetary system? Hale (1994) argued that the rotation axis of a star and the orbit of a close-in companion within <30 AU are aligned based on a comparison between $v\sin i$ measurements and stellar rotation periods. However the situation is less clear for more distant binaries which make up the majority of our current sample. Moreover, recent analysis by Justesen and Albrecht (2020) of a larger, more well-constrained sample from TESS proved insufficient for deriving spin-binary orientations, even for close-in binaries.

If the companion orbit is uncorrelated with the planetary orbit and host star spin axis, we would expect an isotropic distribution of γ (Fig. 3.8, gray distribution). However, our results suggest that both *Kepler*-like and close-in giant systems may exhibit non-isotropic γ distributions, though with low statistical significance ($\Delta log(Z)$ = 1.1 and 2.0, respectively). Specifically, 73^{+14}_{-20}% of $Kepler$-like systems appear to favor alignment between the transiting planet and companion orbits, exhibiting a typical mutual inclination α of approximately $35 \pm 24°$. On the other hand, 65^{+20}_{-35}% of close-in giants appear to favor perpendicular orientations between the transiting planet and companion, exhibiting mutual inclinations α that cluster around $89 \pm 21°$. As a comparison nonparametric method, we applied Kuiper's test to our architecture sub-samples, which is well-suited to quantifying deviations from uniformity for angular data (e.g., Fisher 1993). We derived test statistics of 0.73 and 0.58 for the close-in giant and *Kepler*-like systems, respectively. This indicates borderline significant deviation from uniformity for the close-in giants and less significant deviation for *Kepler*-like systems, in agreement with our Bayes factor results.

There are 20 systems in our sample with reported stellar obliquity λ measurements derived using the Rossiter-McLaughlin effect (Rossiter, 1924; McLaughlin, 1924). Comparison of γ and these λ values reveals another interesting trend; the planets that are misaligned with their host stars according to obliquity measurements (9 systems with $\lambda \gtrsim 30°$) also display γ angles near 90° (Figure 3.12). These perpendicular γ values indicate relatively face-on orbits for the comoving binaries, which in turn betray misaligned orbits between the planets and comoving companions. We performed a simple binomial probability calculation: if we assume a uniform distribution for γ between $0-180°$ as expected from an isotropic planet-companion orientation, the probability that all of the 9 high-λ systems will fall within the observed range of $40 < \gamma < 140°$ is about $\sim$1%. We note that is an a posteriori result; we devised this statistical test based on observations of the data. The true probability of finding all 9 systems within the observed range is likely several times higher than 1%. Still, the correlation between λ and γ is obvious, and suggests that close-in giants that are misaligned with their host stars are likely also misaligned with comoving companions. This result, if confirmed by future studies, may have implications for which proposed mechanisms are most effective at tilting planetary orbits (e.g., Fabrycky and Tremaine 2007; Dawson and Johnson 2018).

However, we emphasize the low statistical significance of the planet-companion alignment. The *Kepler*-like and close-in giant systems favor non-isotropic γ distributions with only small Bayes factor values of $\Delta log(Z)$ = 1.1 and 2.0, respectively. This low significance can be partly attributed to the small size of our TOI sample resulting from our various data quality cuts. Moreover, *Gaia* EDR3 equatorial coordinates, proper motions, and parallaxes were all solved assuming a single-star model, which is bound to introduce systematic errors of varying extent into our sample of comoving binaries as discussed earlier in Section 3.2. Finally, γ is an indirect proxy for the true mutual inclination α between the planet and companion orbits; it cannot be translated to the true planet-companion mutual inclination without knowledge of other Keplerian orbital elements.

Nevertheless, our alignment trend findings echo recent results that point to an excess of perpendicular planetary systems. In particular, Albrecht et al. (2021) found a significant preference for misaligned geometries amongst 57 close-in giant systems as evinced by their true, 3D obliquity measurements. If this excess of polar orbits is corroborated by future studies, it could illuminate which obliquity excitation mechanisms predominantly shape planet architectures. Future alignment analyses will also be aided by upcoming *Gaia* data releases; individual astrometric measurements will be provided in the full, final data release of the nominal mission that will make it possible to directly constrain orbital inclinations for a subset of our sample, thus enabling a more direct and definitive investigation of planet-companion alignment.

References

Akeson, R. L. et al. (Aug. 2013). “The NASA Exoplanet Archive: Data and Tools for Exoplanet Research”. In: *Publications of the Astronomical Society of the Pacific* 125.930, p. 989. DOI: 10.1086/672273. arXiv: 1307.2944 [astro-ph.IM].

Albrecht, Simon H. et al. (July 2021). “A Preponderance of Perpendicular Planets”. In: *The Astrophysical Journal Letters* 916.1, L1, p. L1. DOI: 10.3847/2041-8213/ac0f03. arXiv: 2105.09327 [astro-ph.EP].

Andrews, Jeff J. et al. (Jan. 2019). “Using APOGEE Wide Binaries to Test Chemical Tagging with Dwarf Stars”. In: *The Astrophysical Journal* 871.1, 42, p. 42. DOI: 10.3847/1538-4357/aaf502. arXiv: 1811.12032 [astro-ph.SR].

Belokurov, Vasily et al. (Aug. 2020). “Unresolved stellar companions with Gaia DR2 astrometry”. In: *Monthly Notices of the Royal Astronomical Society* 496.2, pp. 1922–1940. DOI: 10.1093/mnras/staa1522. arXiv: 2003.05467 [astro-ph.SR].

Bensby, T., S. Feltzing, and I. Lundström (Feb. 2004). "Oxygen trends in the Galactic thin and thick disks". In: *Astronomy and Astrophysics* 415, pp. 155–170. DOI: 10.1051/0004-6361:20031655. arXiv: astro-ph/0310741 [astro-ph].

Bensby, T., S. Feltzing, and M. S. Oey (Feb. 2014). "Exploring the Milky Way stellar disk. A detailed elemental abundance study of 714 F and G dwarf stars in the solar neighbourhood". In: *Astronomy and Astrophysics* 562, A71, A71. DOI: 10.1051/0004-6361/201322631. arXiv: 1309.2631 [astro-ph.GA].

Blunt, Sarah et al. (Mar. 2020). "orbitize!: A Comprehensive Orbit-fitting Software Package for the High-contrast Imaging Community". In: *The Astronomical Journal* 159.3, 89, p. 89. DOI: 10.3847/1538-3881/ab6663. arXiv: 1910.01756 [astro-ph.EP].

Bouma, L. G. et al. (Nov. 2020). "Cluster Difference Imaging Photometric Survey. II. TOI 837: A Young Validated Planet in IC 2602". In: *The Astronomical Journal* 160.5, 239, p. 239. DOI: 10.3847/1538-3881/abb9ab. arXiv: 2009.07845 [astro-ph.EP].

Buchhave, Lars A. et al. (June 2012). "An abundance of small exoplanets around stars with a wide range of metallicities". In: *Nature* 486.7403, pp. 375–377. DOI: 10.1038/nature11121.

Carrillo, Andreia et al. (Jan. 2020). "Know thy star, know thy planet: chemo-kinematically characterizing TESS targets". In: *Monthly Notices of the Royal Astronomical Society* 491.3, pp. 4365–4381. DOI: 10.1093/mnras/stz3255. arXiv: 1911.07825 [astro-ph.SR].

Coşkunoğlu, B. et al. (Feb. 2012). "Local stellar kinematics from RAVE data - II. Radial metallicity gradient". In: *Montly Notices of the Royal Astronomical Society* 419.4, pp. 2844–2854. DOI: 10.1111/j.1365-2966.2011.19925.x. arXiv: 1109.6519 [astro-ph.GA].

Dawson, Rebekah I. and Simon H. Albrecht (Aug. 2021). "From Warm Planets to Perpendicular Hot Planets". In: *arXiv e-prints*, arXiv:2108.09325, arXiv:2108.09325. arXiv: 2108.09325 [astro-ph.EP].

Dawson, Rebekah I. and John Asher Johnson (Sept. 2018). "Origins of Hot Jupiters". In: *Astronomy and Astrophysics Review* 56, pp. 175–221. DOI: 10.1146/annurev-astro-081817-051853. arXiv: 1801.06117 [astro-ph.EP].

Díaz, Matías R. et al. (Mar. 2020). "TOI-132 b: A short-period planet in the Neptune desert transiting a V = 11.3 G-type star?". In: *Monthly Notices of the Royal Astronomical Society* 493.1, pp. 973–985. DOI: 10.1093/mnras/staa277. arXiv: 1911.02012 [astro-ph.EP].

Dressing, Courtney D. et al. (Nov. 2014). "Adaptive Optics Images. III. 87 Kepler Objects of Interest". In: *The Astronomical Journal* 148.5, 78, p. 78. DOI: 10.1088/0004-6256/148/5/78. arXiv: 1407.1848 [astro-ph.EP].

Fabrycky, Daniel and Scott Tremaine (Nov. 2007). "Shrinking Binary and Planetary Orbits by Kozai Cycles with Tidal Friction". In: *The Astrophysical Journal* 669.2, pp. 1298–1315. DOI: 10.1086/521702. arXiv: 0705.4285 [astro-ph].

Feroz, F., M. P. Hobson, and M. Bridges (Oct. 2009). "MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics". In: *Monthly Notices of the Royal Astronomical Society* 398, pp. 1601–1614. DOI: 10.1111/j.1365-2966.2009.14548.x. arXiv: 0809.3437.

Fischer, Debra A. and Jeff Valenti (Apr. 2005). "The Planet-Metallicity Correlation". In: *The Astrophysical Journal* 622.2, pp. 1102–1117. DOI: 10.1086/428383.

Fisher, N. I. (1993). *Statistical Analysis of Circular Data*. Cambridge University Press. DOI: 10.1017/CBO9780511564345.

Foreman-Mackey, D., D. W. Hogg, and T. D. Morton (Nov. 2014). "Exoplanet Population Inference and the Abundance of Earth Analogs from Noisy, Incomplete Catalogs". In: *The Astrophysical Journal* 795, 64, p. 64. DOI: 10.1088/0004-637X/795/1/64. arXiv: 1406.3020 [astro-ph.EP].

Fulton, Benjamin J. et al. (July 2021). "California Legacy Survey. II. Occurrence of Giant Planets beyond the Ice Line". In: *The Astrophysical Journal Supplement Series* 255.1, 14, p. 14. DOI: 10.3847/1538-4365/abfcc1. arXiv: 2105.11584 [astro-ph.EP].

Gagné, Jonathan et al. (Mar. 2018). "BANYAN. XI. The BANYAN Σ Multivariate Bayesian Algorithm to Identify Members of Young Associations with 150 pc". In: *The Astrophysical Journal* 856.1, 23, p. 23. DOI: 10.3847/1538-4357/aaae09. arXiv: 1801.09051 [astro-ph.SR].

Gaia Collaboration, A. G. A. Brown, A. Vallenari, T. Prusti, J. H. J. de Bruijne, C. Babusiaux, C. A. L. Bailer-Jones, et al. (Aug. 2018). "Gaia Data Release 2. Summary of the contents and survey properties". In: *Astronomy and Astrophysics* 616, A1, A1. DOI: 10.1051/0004-6361/201833051. arXiv: 1804.09365 [astro-ph.GA].

Gaia Collaboration, A. G. A. Brown, A. Vallenari, T. Prusti, J. H. J. de Bruijne, C. Babusiaux, M. Biermann, et al. (May 2021). "Gaia Early Data Release 3. Summary of the contents and survey properties". In: *Astronomy and Astrophysics* 649, A1, A1. DOI: 10.1051/0004-6361/202039657. arXiv: 2012.01533 [astro-ph.GA].

Gillon, M. et al. (Apr. 2013). "TRAPPIST-UCDTS: A prototype search for habitable planets transiting ultra-cool stars". In: *European Physical Journal Web of Conferences*. Vol. 47. European Physical Journal Web of Conferences, 03001, p. 03001. DOI: 10.1051/epjconf/20134703001.

Guerrero, N. (Jan. 2020). "The TESS Objects of Interest Catalog from Year 1 of the TESS Mission". In: *American Astronomical Society Meeting Abstracts #235*. Vol. 235. American Astronomical Society Meeting Abstracts, 327.03, p. 327.03.

Hale, A. (Jan. 1994). "Orbital CoPlanetary in Solar-Type Binary Systems: Implications for Planetary System Formation and Detection". In: *The Astronomical Journal* 107, p. 306. DOI: 10.1086/116855.

Hamer, Jacob H. and Kevin C. Schlaufman (Sept. 2020). "Ultra-short-period Planets Are Stable against Tidal Inspiral". In: *The Astronomical Journal* 160.3, 138, p. 138. DOI: 10.3847/1538-3881/aba74f. arXiv: 2007.10944 [astro-ph.EP].

Henry, Todd J., J. Davy Kirkpatrick, and Douglas A. Simons (Oct. 1994). "The Solar Neighborhood. I. Standard Spectral Types (K5-M8) for Northern Dwarfs Within Eight Parsecs". In: *The Astronomical Journal* 108, p. 1437. DOI: 10.1086/117167.

Hogg, David W., Adam D. Myers, and Jo Bovy (Dec. 2010). "Inferring the Eccentricity Distribution". In: *The Astrophysical Journal* 725.2, pp. 2166–2175. DOI: 10.1088/0004-637X/725/2/2166. arXiv: 1008.4146 [astro-ph.SR].

Huber, Daniel (May 2017). *isoclassify: v1.2*. Version v1.2. DOI: 10.5281/zenodo.573372. URL: https://doi.org/10.5281/zenodo.573372.

Hwang, Hsiang-Chih, Yuan-Sen Ting, and Nadia L. Zakamska (May 2022). "The eccentricity distribution of wide binaries and their individual measurements". In: *Monthly Notices of the Royal Astronomical Society* 512.3, pp. 3383–3399. DOI: 10.1093/mnras/stac675. arXiv: 2111.01789 [astro-ph.SR].

Jeffreys, H. (1998). *The Theory of Probability*. Oxford Classic Texts in the Physical Sciences. OUP Oxford. ISBN: 9780191589676. URL: https://books.google.com/books?id=vh9Act9rtzQC.

Johnson, Dean R. H. and David R. Soderblom (Apr. 1987). "Calculating Galactic Space Velocities and Their Uncertainties, with an Application to the Ursa Major Group". In: *The Astronomical Journal* 93, p. 864. DOI: 10.1086/114370.

Justesen, A. B. and S. Albrecht (Oct. 2020). "The spin-orbit alignment of visual binaries". In: *Astronomy and Astrophysics* 642, A212, A212. DOI: 10.1051/0004-6361/202039138. arXiv: 2008.12068 [astro-ph.SR].

Kass, Robert E. and Adrian E. Raftery (1995). "Bayes Factors". In: *Journal of the American Statistical Association* 90.430, pp. 773–795. DOI: 10.1080/01621459.1995.10476572. eprint: http://www.tandfonline.com/doi/pdf/10.1080/01621459.1995.10476572. URL: /brokenurl#%20http://www.tandfonline.com/doi/abs/10.1080/01621459.1995.10476572.

Kiseleva, L. G., P. P. Eggleton, and S. Mikkola (Oct. 1998). "Tidal friction in triple stars". In: *Monthly Notices of the Royal Astronomical Society* 300.1, pp. 292–302. DOI: 10.1046/j.1365-8711.1998.01903.x.

Kozai, Yoshihide (Nov. 1962). "Secular perturbations of asteroids with high inclination and eccentricity". In: *The Astronomical Journal* 67, pp. 591–598. DOI: 10.1086/108790.

Kraus, Adam L. et al. (July 2016). "The Impact of Stellar Multiplicity on Planetary Systems. I. The Ruinous Influence of Close Binary Companions". In: *The Astronomical Journal* 152.1, 8, p. 8. DOI: 10.3847/0004-6256/152/1/8. arXiv: 1604.05744 [astro-ph.EP].

Lindegren, L. (2018). In: *Gaia Technical Note GAIA-C3-TN-LU-LL-124-01. Obtained from https://www.cosmos.esa.int/web/gaia/public-dpacdocuments*.

Louden, Emma M. et al. (Feb. 2021). "Hot Stars with Kepler Planets Have High Obliquities". In: *The Astronomical Journal* 161.2, 68, p. 68. DOI: 10.3847/1538-3881/abcebd. arXiv: 2012.00776 [astro-ph.EP].

Lutz, Thomas E. and Douglas H. Kelker (Oct. 1973). "On the Use of Trigonometric Parallaxes for the Calibration of Luminosity Systems: Theory". In: *Publications of the Astronomical Society of the Pacific* 85, p. 573. DOI: 10.1086/129506. URL: https://doi.org/10.1086%2F129506.

Mancini, L. et al. (Sept. 2016). "An optical transmission spectrum of the transiting hot Jupiter in the metal-poor WASP-98 planetary system". In: *Monthly Notices of the Royal Astronomical Society* 461.1, pp. 1053–1061. DOI: 10.1093/mnras/stw1386. arXiv: 1606.00432 [astro-ph.EP].

Matson, Rachel et al. (Jan. 2018). "Stellar Companions of Exoplanet Host Stars in K2". In: *American Astronomical Society Meeting Abstracts #231*. Vol. 231. American Astronomical Society Meeting Abstracts, 109.02, p. 109.02.

Mazeh, Tsevi et al. (Mar. 2015). "Photometric Amplitude Distribution of Stellar Rotation of KOIs—Indication for Spin-Orbit Alignment of Cool Stars and High Obliquity for Hot Stars". In: *The Astrophysical Journal* 801.1, 3, p. 3. DOI: 10.1088/0004-637X/801/1/3. arXiv: 1501.01288 [astro-ph.EP].

McLaughlin, D. B. (July 1924). "Some results of a spectrographic study of the Algol system." In: *The Astrophysical Journal* 60, pp. 22–31. DOI: 10.1086/142826.

Michel, K. -U. and M. Mugrauer (Mar. 2021). "Search for (sub)stellar companions of exoplanet hosts by exploring the second ESA-Gaia data release". In: *Frontiers in Astronomy and Space Sciences* 8, 14, p. 14. DOI: 10.3389/fspas.2021.624907. arXiv: 2102.04385 [astro-ph.EP].

Moe, Maxwell and Kaitlin M Kratter (Aug. 2021). "Impact of binary stars on planet statistics - I. Planet occurrence rates and trends with stellar mass". In: *Monthly Notices of the Royal Astronomical Society* 507.3, pp. 3593–3611. ISSN: 0035-8711. DOI: 10.1093/mnras/stab2328. eprint: https://academic.oup.com/mnras/article-pdf/507/3/3593/40327728/stab2328.pdf. URL: https://doi.org/10.1093/mnras/stab2328.

Mugrauer, M (Nov. 2019). "Search for stellar companions of exoplanet host stars by exploring the second ESA-Gaia data release". In: *Monthly Notices of the Royal Astronomical Society* 490.4, pp. 5088–5102. ISSN: 0035-8711. DOI: 10.1093/mnras/stz2673. eprint: https://academic.oup.com/mnras/article-pdf/490/4/5088/30656870/stz2673.pdf. URL: https://doi.org/10.1093/mnras/stz2673.

Mugrauer, M. and K. -U. Michel (Dec. 2020). "Gaia Search for stellar Companions of TESS Objects of Interest". In: *Astronomische Nachrichten*, arXiv:2009.12234, arXiv:2009.12234. arXiv: 2009.12234 [astro-ph.SR].

Mugrauer, Markus and Kai-Uwe Michel (July 2021). "Gaia search for stellar companions of TESS Objects of Interest II". In: *Astronomische Nachrichten* 342.6, pp. 840–864. DOI: 10.1002/asna.202113972. arXiv: 2105.04625 [astro-ph.SR].

Ngo, Henry et al. (Aug. 2016). "Friends of Hot Jupiters. IV. Stellar Companions Beyond 50 au Might Facilitate Giant Planet Formation, but Most are Unlikely to Cause Kozai-Lidov Migration". In: *The Astrophysical Journal* 827.1, 8, p. 8. DOI: 10.3847/0004-637X/827/1/8. arXiv: 1606.07102 [astro-ph.EP].

Oh, Semyeong et al. (June 2017). "Comoving Stars in Gaia DR1: An Abundance of Very Wide Separation Comoving Pairs". In: *The Astronomical Journal* 153.6, 257, p. 257. DOI: 10.3847/1538-3881/aa6ffd. arXiv: 1612.02440 [astro-ph.SR].

Parker, Richard J. et al. (Aug. 2009). "Do binaries in clusters form in the same way as in the field?" In: *Monthly Notices of the Royal Astronomical Society* 397.3, pp. 1577–1586. DOI: 10.1111/j.1365-2966.2009.15032.x. arXiv: 0905.2140 [astro-ph.GA].

Petigura, Erik A. et al. (Feb. 2018). "The California-Kepler Survey. IV. Metal-rich Stars Host a Greater Diversity of Planets". In: *The Astronomical Journal* 155.2, 89, p. 89. DOI: 10.3847/1538-3881/aaa54c. arXiv: 1712.04042 [astro-ph.EP].

Raghavan, Deepak et al. (Sept. 2010). "A Survey of Stellar Families: Multiplicity of Solar-type Stars". In: *The Astrophysical Journal Supplement Series* 190.1, pp. 1–42. DOI: 10.1088/0067-0049/190/1/1. arXiv: 1007.0414 [astro-ph.SR].

Reddy, Bacham E., David L. Lambert, and Carlos Allende Prieto (Apr. 2006). "Elemental abundance survey of the Galactic thick disc". In: *Monthly Notices of the Royal Astronomical Society* 367.4, pp. 1329–1366. DOI: 10.1111/j.1365-2966.2006.10148.x. arXiv: astro-ph/0512505 [astro-ph].

Ricker, George R. et al. (Jan. 2015). "Transiting Exoplanet Survey Satellite (TESS)". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 1, 014003, p. 014003. DOI: 10.1117/1.JATIS.1.1.014003.

Rossiter, R. A. (July 1924). "On the detection of an effect of rotation during eclipse in the velocity of the brigher component of beta Lyrae, and on the constancy of velocity of this system." In: *The Astrophysical Journal* 60, pp. 15–21. DOI: 10.1086/142825.

Southworth, John et al. (Oct. 2020). “A white dwarf bound to the transiting planetary system WASP-98”. In: *Monthly Notices of the Royal Astronomical Society* 497.4, pp. 4416–4422. DOI: 10.1093/mnras/staa2328. arXiv: 2008.02533 [astro-ph.SR].

Stassun, Keivan G. et al. (Oct. 2019). “The Revised TESS Input Catalog and Candidate Target List”. In: *The Astronomical Journal* 158.4, 138, p. 138. DOI: 10.3847/1538-3881/ab3467. arXiv: 1905.10694 [astro-ph.SR].

Tokovinin, A. and O. Kiyaeva (Feb. 2016). “Eccentricity distribution of wide binaries”. In: *Monthly Notices of the Royal Astronomical Society* 456.2, pp. 2070–2079. DOI: 10.1093/mnras/stv2825. arXiv: 1512.00278 [astro-ph.SR].

Wang, Ji et al. (June 2015). “Influence of Stellar Multiplicity On Planet Formation. III. Adaptive Optics Imaging of Kepler Stars With Gas Giant Planets”. In: *The Astrophysical Journal* 806.2, 248, p. 248. DOI: 10.1088/0004-637X/806/2/248. arXiv: 1505.05363 [astro-ph.EP].

Winn, Joshua N., Daniel Fabrycky, et al. (Aug. 2010). “Hot Stars with Hot Jupiters Have High Obliquities”. In: *The Astrophysical Journal Letters* 718.2, pp. L145–L149. DOI: 10.1088/2041-8205/718/2/L145. arXiv: 1006.4161 [astro-ph.EP].

Winn, Joshua N. and Daniel C. Fabrycky (Aug. 2015). “The Occurrence and Architecture of Exoplanetary Systems”. In: *Astronomy and Astrophysics Review* 53, pp. 409–447. DOI: 10.1146/annurev-astro-082214-122246. arXiv: 1410.4199 [astro-ph.EP].

Ziegler, Carl, Andrei Tokovinin, César Briceño, et al. (Jan. 2020). “SOAR TESS Survey. I. Sculpting of TESS Planetary Systems by Stellar Companions”. In: *The Astronomical Journal* 159.1, 19, p. 19. DOI: 10.3847/1538-3881/ab55e9. arXiv: 1908.10871 [astro-ph.EP].

Ziegler, Carl, Andrei Tokovinin, Madelyn Latiolais, et al. (Oct. 2021). “SOAR TESS Survey. II. The Impact of Stellar Companions on Planetary Populations”. In: *The Astronomical Journal* 162.5, p. 192. DOI: 10.3847/1538-3881/ac17f6. URL: https://dx.doi.org/10.3847/1538-3881/ac17f6.

Table 3.1: Properties of TOI Hosts with Stellar Companions

TIC	RA hh:mm:ss	Dec deg:mm:ss	μ_α mas yr^{-1}	μ_δ mas yr^{-1}	parallax ($\bar{\omega}$) mas	G mag
332064670	08:52:35.22	28:19:47.22	−485.68 ± 0.04	−233.52 ± 0.04	79.45 ± 0.04	5.7
178155732	02:51:56.40	−30:48:50.57	123.44 ± 0.02	106.00 ± 0.04	31.54 ± 0.04	6.3
21832928	17:07:55.63	32:06:19.07	−161.88 ± 0.02	−42.11 ± 0.02	27.27 ± 0.02	7.1
16740101	20:31:26.38	39:56:20.11	16.72 ± 0.03	20.96 ± 0.03	4.83 ± 0.02	7.6
396356111	00:10:23.89	58:29:21.98	−98.73 ± 0.02	−9.84 ± 0.02	10.78 ± 0.02	7.7
263003176	00:50:11.27	−83:44:37.55	139.46 ± 0.03	30.39 ± 0.02	17.42 ± 0.02	7.8
243187830	01:07:37.99	22:57:10.08	103.17 ± 0.04	−490.28 ± 0.02	48.68 ± 0.03	8.0
371443216	09:50:19.22	−66:06:50.14	5.91 ± 0.02	−15.04 ± 0.02	5.94 ± 0.02	8.2
202426247	15:05:49.61	64:02:51.71	−122.14 ± 0.02	110.57 ± 0.02	21.86 ± 0.02	8.2
410214986	23:39:39.72	−69:11:45.79	79.53 ± 0.01	−67.55 ± 0.02	22.64 ± 0.02	8.3

This table lists the TESS Input Catalog (TIC) ID and *Gaia* EDR3-derived RA, Dec, proper motion in the RA (μ_α) and Dec (μ_δ) directions, parallax, and G-band magnitudes for the 170 non-false positive TOI hosts with detected stellar companions. The TOI hosts are sorted by their G-band magnitudes.

Table 3.2: Properties of Stellar Companions to TOI Hosts

TIC	sep	$\Delta\mu_\alpha$	$\Delta\mu_\delta$	Δparallax ($\bar{\omega}$)	ΔG	M_*	$\ln(\mathcal{L}_1/\mathcal{L}_2)$
	AU	mas yr^{-1}	mas yr^{-1}	mas	mag	$M_\odot$	
175532955	1828	0.25 ± 0.02	0.09 ± 0.02	0.05 ± 0.02	−2.3	$0.457^{+0.030}_{-0.030}$	12.96
1551345500	207	1.65 ± 0.03	3.67 ± 0.04	0.06 ± 0.02	1.0	$0.604^{+0.029}_{-0.029}$	9.43
306362738	441	0.48 ± 0.08	0.72 ± 0.11	0.10 ± 0.10	−5.6	$0.296^{+0.033}_{-0.033}$	8.34
73649615	955	0.80 ± 0.03	0.30 ± 0.03	0.01 ± 0.03	−1.3	$0.265^{+0.024}_{-0.024}$	7.31
23434737	763	0.01 ± 0.43	0.51 ± 0.38	0.11 ± 0.48	−11.5	$0.112^{+0.004}_{-0.004}$	7.29
281781375	978	0.74 ± 0.02	0.18 ± 0.02	0.00 ± 0.02	−1.2	$0.782^{+0.038}_{-0.038}$	6.97
236445129	16184	0.47 ± 0.01	0.08 ± 0.02	0.00 ± 0.02	−0.4	$1.191^{+0.172}_{-0.172}$	6.84
452866790	1346	3.78 ± 0.07	0.19 ± 0.05	0.16 ± 0.07	−3.5	$0.111^{+0.004}_{-0.004}$	6.78
322307342	11838	0.28 ± 0.07	0.16 ± 0.07	0.06 ± 0.07	−5.2	–	6.49
280095254	114	1.01 ± 0.14	0.65 ± 0.21	0.35 ± 0.09	−1.5	$0.834^{+0.041}_{-0.042}$	6.45

This table lists the TIC ID and differences in *Gaia* EDR3-derived RA, Dec, proper motion in the RA (μ_α) and Dec (μ_δ) directions, parallax, and G-band magnitude for the 172 stellar companions to non-false positive TOI hosts, as well as their masses derived from `isoclassify` (Huber, 2017). The companions are sorted by their log-likelihood ratios $\ln(\mathcal{L}_1/\mathcal{L}_2)$.

Table 3.3: Properties of TOIs with Stellar Companions

TIC	TOI in System	Num. Companions in System	TESS Dis.	TFOPWG Dis.	Source	Period days	R_P $R_\oplus$	T_{eq} K	$\ln(\mathcal{L}_1/\mathcal{L}_2)$
175532955	929.010	1	PC	PC	qlp	5.830 ± 0.000	2.710 ± 0.240	774.000	12.960
1551345500	1764.010	1	PC	PC	spoc	47.390 ± 0.000	13.320 ± 1.170	321.120	9.430
306362738*	479.010	1	KP	KP	spoc	2.780 ± 0.000	12.680 ± 0.610	1270.340	8.340
73649615	756.010	1	PC	PC	spoc	1.240 ± 0.000	2.870 ± 0.390	827.940	7.310
23434737*	1203.010	1	CP	CP	spoc	25.520 ± 0.004	2.960 ± 0.270	662.890	7.290
281781375	204.010	1	PC	PC	spoc	43.830 ± 0.006	2.510 ± 1.930	507.340	6.970
236445129*	1282.010	1	KP	KP	spoc	0.970 ± 0.000	16.740 ± 0.770	2332.220	6.840
452866790*	488.010	1	CP	CP	spoc	1.200 ± 0.000	1.120 ± 1.520	704.630	6.780
322307342*	117.010	1	KP	KP	spoc	3.590 ± 0.000	16.470 ± 0.920	1658.390	6.490
280095254	235.010	1	PC	PC	spoc	10.090 ± 0.000	–	795.970	6.450

This table lists the properties of all 254 TOIs with detected stellar companions. These include the TIC ID, TOI ID, number of detected *Gaia* EDR3 companions in the system, TESS and TFOPWG dispositions denoting candidate status (e.g., FP = false positive, APC = ambiguous planetary candidate, PC = planetary candidate, CP = confirmed planet, KP = known planet), TESS processing pipeline source (i.e., the NASA/Ames Science Processing Operations Center – SPOC or the TESS Science Office Quick-Look Pipeline – QLP), orbital period, radius, and equilibrium temperature. The TOIs are sorted by their detected companion log-likelihood ratios $\ln(\mathcal{L}_1/\mathcal{L}_2)$. Confirmed planet hosts are marked with an "*″ next to their corresponding TIC.

Chapter 4

PLANET ENGULFMENT DETECTIONS ARE RARE ACCORDING TO OBSERVATIONS AND STELLAR MODELING

Behmard, Aida et al. (May 2023). "Planet engulfment detections are rare according to observations and stellar modelling". In: *Monthly Notices of the Royal Astronomical Society* 521.2, pp. 2969–2987. DOI: `10.1093/mnras/stad745`. arXiv: `2210.12121 [astro-ph.EP]`.

Abstract

Dynamical evolution within planetary systems can cause planets to be engulfed by their host stars. Following engulfment, the stellar photosphere abundance pattern will reflect accretion of rocky material from planets. Multi-star systems are excellent environments to search for such abundance trends because stellar companions form from the same natal gas cloud and are thus expected to share primordial chemical compositions to within 0.03−0.05 dex. Abundance measurements have occasionally yielded rocky enhancements, but few observations targeted known planetary systems. To address this gap, we carried out a Keck-HIRES survey of 36 multi-star systems where at least one star is a known planet host. We found that only HAT-P-4 exhibits an abundance pattern suggestive of engulfment, but is more likely primordial based on its large projected separation (30,000 ± 140 AU) that exceeds typical turbulence scales in molecular clouds. To understand the lack of engulfment detections among our systems, we quantified the strength and duration of refractory enrichments in stellar photospheres using `MESA` stellar models. We found that observable signatures from 10 $M_\oplus$ engulfment events last for ~90 Myr in 1 $M_\odot$ stars. Signatures are largest and longest lived for 1.1−1.2 $M_\odot$ stars, but are no longer observable ~2 Gyr post-engulfment. This indicates that engulfment will rarely be detected in systems that are several Gyr old.

4.1 Introduction

Gravitationally bound stars form from the approximately homogeneous material of their shared natal gas cloud; it follows that differences in their elemental abundances are expected to fall within the small range of chemical dispersion observed in stellar clusters and associations (e.g., De Silva, Freeman, Asplund, et al. 2007;

De Silva, Freeman, and Bland-Hawthorn 2009; Bland-Hawthorn, Krumholz, and Freeman 2010). However, several studies have found abundance differences >0.05 dex[1] between stars in binary systems (e.g., Ramírez, Meléndez, et al. 2011; Mack, Schuler, et al. 2014; Tucci Maia, Meléndez, and Ramírez 2014; Teske, Ghezzi, et al. 2015; Jofré et al. 2021), with extreme cases exhibiting differences up to ∼0.2 dex (Oh, Price-Whelan, Brewer, et al., 2018).

There are various proposed mechanisms for these abundance differences related to planet formation. For example, observed refractory element depletion can be attributed to missing solid material locked up in rocky planets. Meléndez et al. (2009) put forward this scenario to explain the Sun's observed depletion pattern, but noted that it only makes sense if the combined mass of the Solar System terrestrial planets is removed from just the solar convective zone. It is possible that dust-depleted gas was accreted onto the Sun 10−25 Myr after Solar System formation, once the solar convective zone began shrinking to its current mass fraction (∼2%, Hughes, Rosner, and Weiss 2007). However, only 1% of stars with ages ≥13 Myr show signs of accretion (White and Hillenbrand, 2005; Currie et al., 2007), indicating that late-stage accretion after the protoplanetary disk has dissipated (typical lifetimes 1−3 Myr, Li and Xiao 2016) is rare. Thus, we do not expect that sequestration of refractory material in planets will produce strong depletion signals. Alternatively, Booth and Owen (2020) suggested that depletion trends may emerge from gaps in protoplanetary disks created by forming giant planets. These gaps could create pressure traps that prevent accretion of refractory material onto the host star.

Abundance differences can also be produced from refractory enrichment. A particularly promising scenario for producing strong enrichment signals is planet engulfment, which could deposit large amounts of rocky planetary material within the convective regions of engulfing stars. Spectral analysis of polluted white dwarfs provide strong evidence for accretion of planetary remnants (e.g., Zuckerman, Melis, et al. 2010; Koester, Gänsicke, and Farihi 2014; Farihi 2016), with some white dwarfs exhibiting surface abundance patterns that closely match bulk Earth composition material (e.g., Zuckerman, Koester, et al. 2007; Klein et al. 2010). There is also evidence for planet engulfment in solar-like stars. For example, Oh, Price-Whelan, Brewer, et al. (2018) recently reported a strong (∼0.2 dex) potential signature of

[1] In this work, we adopt the standard "bracket" chemical abundance notation [X/H] = $A(\mathrm{X})$ - $A(\mathrm{X})_{\odot}$, where $A(\mathrm{X}) = \log(n_{\mathrm{X}}/n_H) + 12$ and n_{X} is the number density of species X in the star's photosphere.

planet engulfment in the HD 240429-30 (Kronos-Krios) system. We investigate abundance differences between stellar companions through the lens of planet engulfment here. Throughout this study, we use the term engulfment to reference planetary material ingestion events that occur prior to post-main sequence host star expansion, such as refractory material accretion due to dynamical scattering during the early stages of system evolution.

There are ten binary systems reported in the literature with one star significantly enhanced in refractories (>0.05 dex) compared to its stellar companion. Among these ten systems, seven host known planets (Ramírez, Meléndez, et al., 2011; Mack, Schuler, et al., 2014; Tucci Maia, Meléndez, and Ramírez, 2014; Teske, Ghezzi, et al., 2015; Ramírez, Khanal, Aleo, et al., 2015; Biazzo et al., 2015; Teske, Khanal, and Ramírez, 2016; Saffe, Jofré, Martioli, et al., 2017; Tucci Maia, Meléndez, Lorenzo-Oliveira, et al., 2019; Jofré et al., 2021). Depending on the study, four to seven of these planet host systems have refractory differences that trend with condensation temperature T_c (Table 3.1). We expect a T_c-dependent differential abundance pattern following planet engulfment; in the absence of engulfment, elements with higher T_c are more likely to be condensed throughout the disk and become locked in solid planetary material. Conversely, elements with lower T_c are more likely to reside in the gas phase and become depleted through accretion onto the host star. Thus, rocky planetary compositions are dictated by the radial temperature gradient in the disk, with higher abundances of refractory species in order of T_c. Additionally, a T_c-dependent differential abundance pattern is not expected to result from stellar processes alone.

There have been a few differential abundance studies for larger samples. For example, Hawkins et al. (2020) reported abundances for 25 comoving, wide binaries and found that while 80% (20 pairs) are homogeneous in [Fe/H] at levels below 0.02 dex, the five remaining systems exhibit Δ[Fe/H] ~ 0.10 dex. If we assume that these refractory enhancements indicate planet engulfment, they imply an engulfment rate of 20%. However, the authors did not recover a strong T_c trend for any of the Δ[Fe/H] ~ 0.10 dex systems, suggesting that the abundance differences may stem from other processes. The absence of a strong T_c trend could also be attributed to a lack of low T_c element measurements in the Hawkins et al. (2020) sample, which makes the T_c trend difficult to discern, or abundance measurement error. More recently, Spina et al. (2021) analyzed differential abundances among 107 binary systems. While they did not assess T_c trends, they found that ~20−35% of their sample exhibits large refractory-to-volatile abundance ratios that may be indicative of engulfment. While these results are intriguing, they highlight the need for further high-precision abundance studies that consider T_c to constrain the true rate of planet engulfment.

Understanding the conditions and prevalence of planet engulfment is vital for mapping the fate of refractory material within planetary systems. There are multiple lines of evidence that solid planetary material is predominantly refractory. For example, white dwarf pollution patterns from planet debris exhibit rocky compositions (Xu et al., 2019; Putirka and Xu, 2021), and the bulk densities of several super-Earth exoplanets, e.g., the TRAPPIST-1 planets and Kepler-93b (Dressing et al., 2015), are indicative of Earth-like rock-iron ratios. Thus, the building blocks of planets are sourced from the dusty component of protoplanetary disks. However, it is not clear how much disk dust becomes locked in planets or sequestered in debris disks (e.g., Booth and Owen 2020), is engulfed by the host star following a combination of radial drift and dynamical interactions, or is blown out of the system. In other words, we have not quantified the efficiency of planet formation. Refractory enhancements in planet host stars due to engulfment can be used to back out mass measurements of polluting refractory material, which will shed light on how much mass went into planets or was trapped in the outer disk, and how that mass was redistributed in the system after the disk dissipated.

The prevalence of planet engulfment also has implications for stellar chemical evolution. Stars are born together in clusters, but disperse on timescales of $\sim$100 Myr post-intracluster gas removal (Krumholz, McKee, and Bland-Hawthorn, 2019). Galactic archaeology attempts to link stars back to their siblings through chemical tagging that can trace the chemical and kinematic evolution of the Milky Way. However, chemical tagging relies on the assumption that such stellar siblings are coeval and share the same elemental abundance patterns to within 0.03$-$0.05 dex (e.g., De Silva, Freeman, Asplund, et al. 2007; Bovy 2016; Ness et al. 2018). This assumption may not be true if planet engulfment is a common phenomenon. Indeed, it has been suggested that observations of significant chemical dispersion observed within stellar clusters and associations, such as inhomogeneities in neutron capture elements within the open cluster M67 (Liu, Asplund, Yong, et al., 2016), and abundance differences at the 0.02 dex level for 19 elements in the Hyades open cluster (Liu, Yong, Asplund, Ramírez, et al., 2016), are due to planet engulfment (Oh, Price-Whelan, Hogg, et al., 2017; Ness et al., 2018).

In addition, there are no high-precision abundance surveys that specifically targeted planet hosts. Assessing engulfment signatures in systems with existing planets is important for understanding the dynamical conditions that may give rise to planet engulfment, such as planet-planet scattering in multi-planet systems (Rasio and Ford, 1996; Weidenschilling and Marzari, 1996). To fill this gap, we carried out a survey with the Keck High Resolution Echelle Spectrometer (HIRES) of 36 confirmed planet host systems with stellar companions to investigate the role of engulfment in planetary system evolution, and shed light on which dynamical pathways may dominate. For more details on the sample, see Section 4.2. The abundance analysis and engulfment model used to derive mass measurements of engulfed material are presented in Sections 4.3 and 4.4, respectively. Our `MESA` analysis is outlined in Section 4.5. The results of our survey are presented in Section 4.6, and are compared to previously published results in Section 4.7. Implications for planet engulfment and chemical homogeneity in multi-star systems are discussed in Section 4.8. Finally, we summarize our findings in Section 4.9.

4.2 Planet Engulfment Sample

Our planet engulfment sample consists of multi-star systems where at least one star is a confirmed planet host. The sample is largely sourced from the Mugrauer (2019) catalog of 207 confirmed planet hosts with stellar companions at separations of <9100 AU, compiled from the second data release of the *Gaia* mission (*Gaia* DR2, Gaia Collaboration et al. 2018). The companions were identified through a set of astrometric conditions based on the parallax and proper motion measurements of each planet host star and potential companion. Specifically, a cut on a quantity referred to as the common proper motion (CPM) index by Mugrauer (2019), defined as $|\mu_{PH} - \mu_{comp}|/\mu_{rel}$, where μ_{PH} and μ_{comp} are the proper motions of the exoplanet host star and potential companion, and μ_{rel} is their relative tangential velocity. Mugrauer (2019) defined bound companions as those with a CPM-index > 3, and a difference in parallax $\Delta\pi < 3$ mas. When met, these conditions indicate that the pair of stars share a common proper motion and are thus gravitationally bound. For more details on the companion selection criteria, see Mugrauer (2019).

We applied a projected separation cut of $>1.5''$ to ensure that the two stars would be cleanly resolved by Keck-HIRES, as well as an effective temperature cut of $T_{\rm eff}$ = 4700–6500 K. The latter cut was applied because the spectral synthesis code used for our abundance analysis (Spectroscopy Made Easy, `SME`) does not produce reliable abundances outside of this temperature range (Valenti and Piskunov 1996; Brewer, Fischer, et al. 2016). For the companions, we used their $T_{\rm eff}$ values reported in Mugrauer (2019). These were determined from absolute G-bands magnitudes and the Baraffe et al. (2015) (sub)stellar evolution models assuming an age of 5 Gyr, which is the average age of systems in the Mugrauer (2019) sample. For the planet hosts, we used the most recently reported $T_{\rm eff}$ from the *NASA Exoplanet Archive*[2]. We foreshadow here that `SME` provides more accurate $T_{\rm eff}$ measurements (typical errors are 7−14 K due to varying SNR levels, added in quadrature with an additional 25 K stemming from instrumental and stellar sources, e.g., the spectral line spread function (SLSF) or point spread function (PSF), telluric lines, and stellar activity Brewer and Fischer 2018), so this cut was redone after collecting spectra for our targets and running them through `SME`. This eliminated a further seven systems, which is described in more detail below. However at this point, we were left with 35 systems. We augmented this sample by searching for stellar companions to planet hosts that met these criteria in the *NASA Exoplanet Archive*, which resulted in an additional two systems (HAT-P-4 and WASP-180). Eleven of the 37 planet host binaries qualify as stellar twins ($\Delta T_{\rm eff}$ < 200 K, Andrews et al. 2019), which are well suited to differential abundance analyses given their near-identical evolutionary states. All systems in our sample were verified to host confirmed planets according to the *NASA Exoplanet Archive*[2]. Finally, we removed any systems that display evidence of spectroscopic binary contamination in their spectral cross-correlation; such contamination will lead to inaccurate `SME` abundance predictions. This was the case for ψ^1 Dra, leaving 36 systems.

The final sample of 36 systems contains 28 binaries and eight triples. Though four of the eight triples are hierarchical, we determined that the spectra of individual stars in these systems are not blended with those of nearby companions using the `ReaMatch` code (Kolbl et al., 2015). We also ensured that the planet hosts and their stellar companions have similar rotational velocities by checking that their $v \sin i$ agree to within ∼10 km/s. Each of the triple systems has only one stellar companion that meets the $T_{\rm eff}$ and projected separation criteria. Thus, two stars were always analyzed per system. The equatorial coordinates, $T_{\rm eff}$, $\log g$, M_*, *Gaia* Early Data Release 3 (ER3)-sourced RVs, parallaxes, proper motions, and V-band magnitudes of stars in the sample are listed in Table 4.2. Some sources are missing RV measurements because they do not meet the Gaia DR2/EDR3 RV criteria of G-band magnitudes less than ∼13, or were deemed inaccurate due to companion contamination (Boubert et al., 2019). Among the 36 systems, ten have existing high-precision abundance measurements (HAT-P-1, HD 20781-82, XO-2, WASP-94, HAT-P-4, HD 80606-07,

[2]`https://exoplanetarchive.ipac.caltech.edu/`

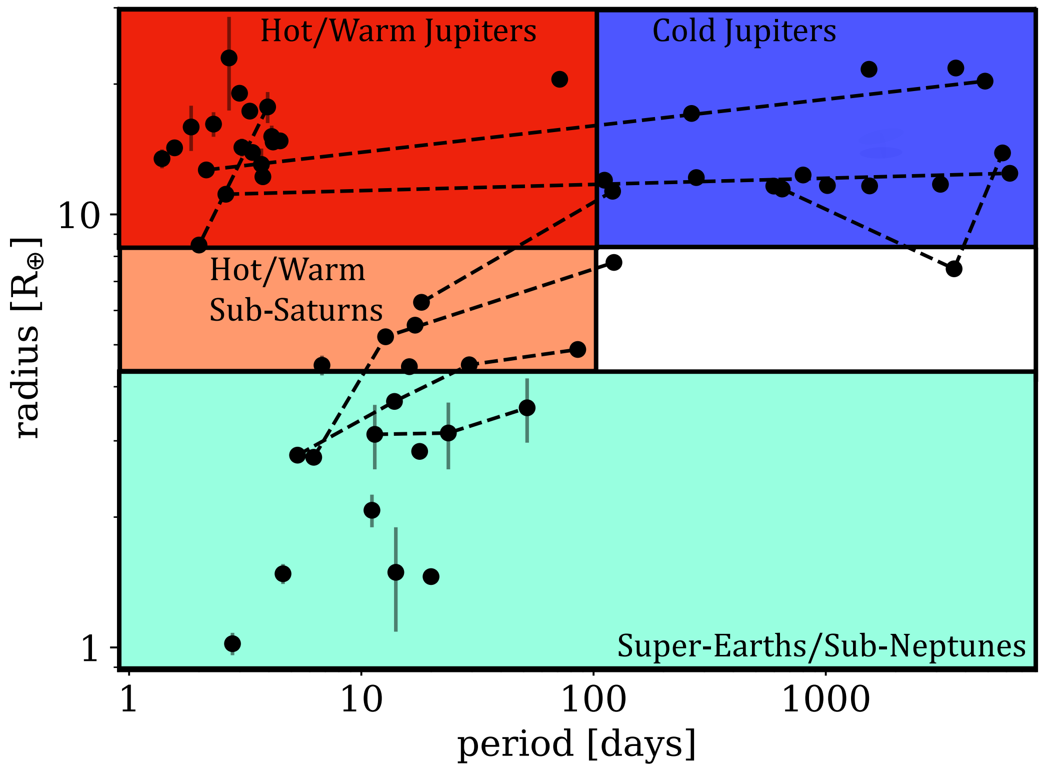


Figure 4.1: The radii vs. orbital period distribution for planets in our sample. Hot/warm Jupiters are defined as planets with $R > 8\ R_\oplus$ and $P < 100$ days, hot/warm sub-Saturns with $4\ R_\oplus < R < 8\ R_\oplus$ and $P < 100$ days, cold Jupiters with $R > 8\ R_\oplus$ and $P > 100$ days, and super-Earths/sub-Neptunes with $R < 4\ R_\oplus$. Planets that share the same host star are connected by dashed lines.

16-Cygni, HD 133131, HD 106515, WASP-160; Table 3.1) derived from the `MOOG` spectral synthesis code (Sneden, 1973; Sobeck et al., 2011) that can be compared with predictions from `SME`.

The engulfment sample systems span a wide range of planetary architectures that include super-Earths/sub-Neptunes, compact multi-planet systems, and giant planets at a range of orbital periods (Table 4.3). Figure 4.1 shows the radii versus rotation periods for all planets in the engulfment sample. For planets lacking reported radius measurements according to the *NASA Exoplanet Archive*[2], we derived radii from mass measurements with the following power-law mass-radius relation that assumes Earth-like compositions (Rubenzahl et al. *in prep.*):

$$M = CR^{\gamma} \tag{4.1}$$

where the C and γ were constrained to values of 0.83 and 3.52 using a sample of

122 confirmed exoplanets with Keck-HIRES spectra and precise radii measurements. For planets massive enough to host gaseous envelopes greater than 1% by mass, the envelope mass was accounted for by assuming a gas density of 0.417 g cm^{-3} as constrained with the Rubenzahl et al. (*in prep.*) planet sample. We include errors bars on planet radii measurements in Figure 4.1 if they are reported in the *NASA Exoplanet Archive*[2].

4.3 Stellar Abundance Analysis

We obtained spectra for these stars with HIRES at the Keck I 10 m telescope (Vogt et al., 1994) using procedures from the California Planet Search. Howard et al. (2010) provides descriptions of the observing and analysis procedures. We used the C2 decker for targets with V-band magnitudes fainter than 10 mag, and the B5 decker for targets with V-band magnitudes of 10 mag or brighter. The HIRES spectra are high-resolution (R ≈ 50,000) with high signal-to-noise ratios per pixel (SNR ≥ 40/pixel, with ~50% having SNR > 100/pixel). The wavelength range utilized spans 350 Å of the spectrum in specific segments between 5164 Å and 7800 Å, as described in Brewer, Fischer, et al. (2016) for their `SME` implementation. Our choice of SNR ≈ 40–400/pix for the engulfment sample HIRES observations was motivated by the expected `SME` prediction precisions as a function of SNR; for HIRES spectra with SNR = 40–100/pix, `SME` achieves precisions of 0.01–0.05 dex in [X/H] for the following elements: C, N, O, Na, Mg, Al, Si, Ca, Ti, V, Cr, Mn, Fe, Ni, Li, and Y (e.g., Brewer, Fischer, et al. 2016; Brewer and Fischer 2018). The refractory species alone (Fe, Ti, Al, etc.) achieve higher precision of 0.01–0.03 dex, which translates to detections at the ~1 $M_\oplus$ level according to the Oh, Price-Whelan, Brewer, et al. (2018) model used for their analysis of engulfment in the Kronos-Krios system. This precision is sufficient for detecting signatures of planet engulfment, i.e., refractory enhancements, at levels of >0.05 dex (e.g., Ramírez, Khanal, Lichon, et al. 2019). For reference, an abundance difference of Δ[X/H] = 0.05 dex corresponds to ~2 $M_\oplus$ of engulfed solid material assuming a solar-like convective zone mass of M_{cz} = 0.02 $M_\odot$ (Saffe, Jofré, Martioli, et al., 2017).

The `SME`-determined stellar parameters ($T_{\rm eff}$, log$g$) for the engulfment sample are provided in Table 4.2. The `SME` stellar parameters are more accurate than those initially used to select our engulfment sample, and seven stars (WASP-3 C, HD 23596 B, PR0211 B, HAT-P-41 B, Kepler-410 B, WASP-70 B, Kepler-1150 B) have `SME`-determined $T_{\rm eff}$ below our sample cutoff 4700 K. Thus, these systems were removed from our engulfment analysis, leaving 29 binaries in our sample that include all eleven twin systems. The SME-determined abundances are given in Table 4.4, and associated errors in Table 4.5. The abundance errors are estimated from two sources: the SNR of the HIRES spectra as mentioned above, and the scatter in measured abundances from different observations of the same target. To quantify how these error sources affect abundance predictions, Brewer and Fischer (2018) ran SME on a set of simulated solar and cool star spectra with varying amounts of added Gaussian random noise that mimic varying SNR levels (Table 2, Brewer and Fischer 2018). We conducted a similar investigation with real data using Keck-HIRES observations of eight bright stars spanning a range of T_{eff} and [Fe/H] at five different SNR levels, and found that the scatter in SME-determined abundances agrees with the abundance errors reported in Brewer and Fischer (2018). This error analysis is described more fully in the Appendix.

Lithium Measurements

Lithium abundances provide an independent line of evidence for planet engulfment. Unlike other refractory species, lithium is destroyed in thermonuclear reactions at comparatively low temperatures ($T \approx 3 \times 10^6$ K). This makes observable lithium enrichment signatures short-lived in stellar photospheres (Berger, Howard, and Boesgaard 2018), with signatures lifetimes varying as a function of stellar mass (~50 Myr–5.5 Gyr for 0.7–1.3 $M_{\odot}$, Sevilla, Behmard, and Fuller 2022). Thus, enhanced surface lithium in stars that are not particularly young may signify recent events that modified stellar chemistry beyond birth compositions, such as planet engulfment.

The Li I doublet at 6708 Å was used to measure Li abundances for our sample of planet host binaries. First, we derived Li equivalent width (EW) measurements. This was done with spectra that were continuum-normalized through removal of the blaze function, then Doppler-corrected through cross-correlation with the rest-wavelength, National Solar Observatory solar spectrum (Wallace et al., 2011) as implemented in the SpecMatch-Emp package (Yee, Petigura, and von Braun, 2017). We followed the procedure outlined in Berger, Howard, and Boesgaard (2018) to calculate Li EWs. In brief, the LMFIT (Newville et al., 2014) Levenberg-Marquardt minimization routine implemented in Python was used to fit a four component composite model to the Li I doublet region. The components consisted of a constant to accommodate the continuum, two Gaussians for the two Li I features at 6707.76 Å and 6707.91 Å, and another Gaussian for the nearby Fe I feature at 6707.44 Å. Only the continuum constant and two Li I Gaussians were considered in the Li

EW calculation. Li EW measurement uncertainties were taken as the quadratic sum of the statistical photometric error due to SNR/pixel, and the range in EW measurements when modifying the continuum placement (Cayrel, 1988; Bertran de Lis et al., 2015).

Li abundances were derived from the EW measurements with the `MOOG` (Sneden, 1973) spectral synthesis code. We chose `MOOG` over `SME` because the `SME` line list in our implementation from Brewer, Fischer, et al. (2016) does not include Li spectral features. Instead, we used the `MOOG` *blends* routine. `MOOG` was implemented via the Python wrappers `pymoog`[3] and `pymoogi`[4], where `pymoog` was used to select an appropriate model atmosphere from a provided library of Kurucz ATLAS9 model grids, and the Li abundances were calculated via the *blends* routine contained in `pymoogi` from Li EW measurements. In this step, the errors on stellar parameters were incorporated by simultaneously sampling from Gaussian distributions with widths equal to the uncertainties on $T_{\rm eff}$, $\log g$, and [Fe/H] 100 times. The scatter of the resulting abundance measurements was then added in quadrature with the difference in Li abundance from including the Li EW uncertainty discussed above. The result is our total Li abundance uncertainty. The engulfment sample Li EWs and abundances are provided in Table 4.6, along with their expected ΔA(Li) values only considering the difference in $T_{\rm eff}$ between the companions, which will be non-zero even in the absence of engulfment (Randich and Magrini, 2021). Only a ΔA(Li) value between companions that exceeds this expected value may indicate engulfment.

4.4 Engulfment Model

We present a framework similar to that of Oh, Price-Whelan, Brewer, et al. (2018) for estimating the remaining mass of bulk Earth composition (McDonough, 2003) material engulfed in one star given abundance measurements for a binary pair. We emphasize *remaining* here because the initial refractory enrichment in stellar photospheres following engulfment is depleted over time; once the system is observed, there will be less refractory material in the engulfing star photosphere than was immediately present after the engulfment event (see Section 4.5 for our analysis of engulfment signature timescales and depletion mechanisms).

From the stellar abundances of the engulfing star [X/H], we can express the mass fraction of each element X as:

[3] `https://github.com/MingjieJian/pymoog/`

[4] `https://github.com/madamow/pymoogi/`

$$f_{\mathrm{X,photo}} = \frac{10^{[\mathrm{X/H}]} m_{\mathrm{X}}}{\Sigma_{\mathrm{X}} 10^{[\mathrm{X/H}]} m_{\mathrm{X}}}, \tag{4.2}$$

where m_{X} is the mass of each element in atomic mass units. We note that this approach of computing mass fraction rather than number density fraction should be appropriate for our systems, namely binaries composed of stars with low Z. Assuming a total mass of accreted material M_{acc} and accreted mass fractions for each element $f_{\mathrm{X,acc}}$, the abundance difference is

$$\Delta[\mathrm{X/H}] = \log_{10} \frac{f_{\mathrm{X,photo}}\, f_{cz}\, M_* + f_{\mathrm{X,acc}}\, M_{\mathrm{acc}}}{f_{\mathrm{X,photo}}\, f_{cz}\, M_*}, \tag{4.3}$$

where f_{cz} is the mass fraction of the stellar convective zone. Similar calculations have been performed by, e.g., Chambers (2010) and Mack, Schuler, et al. (2014) and Mack, Stassun, et al. (2016). For more details on the engulfment model, see Oh, Price-Whelan, Brewer, et al. (2018). Because the modeled amount of polluting material derived from refractory enhancements depends on the convective zone mass M_{cz}, we adjusted M_{cz} to the stellar type of the engulfing star according to the T_{eff}-M_{cz} relation in Pinsonneault, DePoy, and Coffee (2001). We tested our model by applying it to the reported abundances of the Kronos-Krios system, which were also derived from Keck-HIRES spectra and `SME` (Brewer, Fischer, et al., 2016). The model recovered $13.68 \pm 1.93\ M_\oplus$ of bulk Earth composition engulfed mass (Figure 4.2), in good agreement with the reported engulfed mass of $\sim 15\ M_\oplus$ from Oh, Price-Whelan, Brewer, et al. (2018).

Our engulfment model employs the `dynesty` nested sampling code (Speagle, 2020) to determine the Bayesian evidence for the engulfment model or a flat model of differential abundances as a function of T_c, shown in Figure 3.2 as the long-dash line. The flat model represents the case of no engulfment. We found that the engulfment model is preferred over the flat model for the Kronos-Krios system with a Bayesian evidence difference of $\Delta\ln(Z)$[5] = 15.8.

Bayesian Evidence

To determine the Bayesian evidence difference $\Delta\ln(Z)$ that indicates a strong engulfment detection, we compared samples of simulated engulfment and non-engulfment

[5]The Bayesian evidence Z is defined as the ratio of probabilities of getting data D assuming the engulfment or flat model: $Z = \frac{P(D|M_{\mathrm{engulf}})}{P(D|M_{\mathrm{flat}})}$.

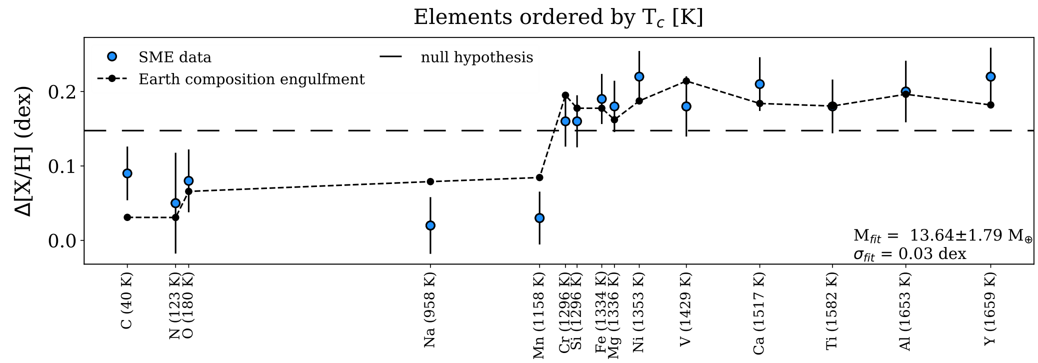


Figure 4.2: Fitted model to the differential abundance measurements between HD 240429 (Krios) and HD 240430 (Kronos) (Oh, Price-Whelan, Brewer, et al., 2018). Blue circles represent the abundance differences from Oh, Price-Whelan, Brewer, et al. (2018), and black dots are our model fit with 13.68 ± 1.93 $M_\oplus$ of bulk Earth composition engulfed material added to the convective zone of one star. The abundances are ranked by T_c of elements for solar-composition gas from Lodders (2003). The amount of modeled engulfed material and fitted scatter is provided in the lower right corner of the plot. The null hypothesis (no engulfment, but uniform abundance enrichment across all elements for one star) is shown by the long-dash line.

systems. The synthetic engulfment sample was constructed by randomly drawing 1000 systems from our twin binary systems. We drew from ten of our eleven twin systems because we excluded HAT-P-4 given its potential engulfment status (see Section 4.6). We took the planet host abundances for both stars to begin with as Δ[X/H] = 0 across all elements, and then we added 10 $M_\oplus$ of bulk Earth composition material into the convective zone of the planet host star. Intrinsic scatter was then added to the abundances of the companion star according to the observed abundance scatter of 20 chemically homogeneous (Δ[Fe/H] < 0.05 dex) wide binaries reported in Hawkins et al. (2020) (0.067 dex, 0.05 dex, 0.052 dex, 0.029 dex, 0.039 dex, 0.03 dex, 0.11 dex, 0.046 dex, 0.12 dex, 0.05 dex, 0.06 dex, 0.044 dex, 0.091 dex for C, Na, Mn, Cr, Si, Fe, Mg, Ni, V, Ca, Ti, Al, Y, respectively). Abundances for N and O were not provided in Hawkins et al. (2020), so we instead used the M67 open cluster scatter reported for these elements (0.015 dex and 0.022 dex for N and O, respectively, Bovy 2016). Further scatter was added to the companion star abundances as a function of SNR according to Brewer and Fischer (2018) to mimic observations. The simulated non-engulfment systems were constructed by again randomly drawing 1000 systems from the twin binaries, but again excluding HAT-P-4 given its potential engulfment status. The abundances of the stars were

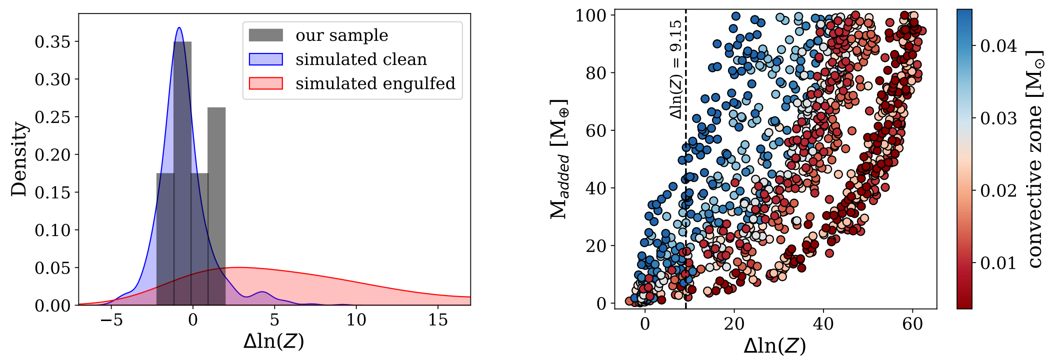


Figure 4.3: The left panel displays the difference in Bayesian evidence values between the engulfment and flat models $\Delta\ln(Z)$ for the eleven twin systems in our engulfment sample (black), with the two $\Delta\ln(Z)$ probability density functions for simulated systems that have (red) or have not (blue) undergone engulfment. These synthetic samples are each composed of 1000 systems randomly drawn with replacement from our ten twin systems (excluding HAT-P-4). For the simulated engulfment systems, we added 10 $M_\oplus$ of bulk Earth composition material to the planet host star assuming even distribution throughout the stellar convective zone, and computed abundances according to our engulfment model. The synthetic engulfment and non-engulfment distributions exhibit significant overlap. The right panel displays $\Delta\ln(Z)$ values corresponding to 0.1−100 $M_\oplus$ simulated engulfment systems. The colors represent the engulfing star convective zone mass, and the maximum $\Delta\ln(Z)$ value for the synthetic non-engulfment distribution of 9.15 is marked by the dashed line.

not modified at all because we assumed that these real observations correspond to non-engulfment systems, but we again included scatter according to the 20 chemically homogeneous Hawkins et al. (2020) wide binaries. We randomly chose the direction between the two companions when computing the differential abundances for the simulated non-engulfment pairs.

We then ran both samples through our engulfment model machinery to determine $\Delta\ln(Z)$ for each simulated system. The $\Delta\ln(Z)$ probability density distributions for the synthetic engulfment and non-engulfment samples are shown in the left panel of Figure 4.3. The synthetic engulfment and non-engulfment distributions exhibit significant overlap, with ~55% of engulfment systems overlapping with the non-engulfment distribution. We conclude that our spectroscopic measurements and $\Delta\ln(Z)$ analysis cannot identify nominal engulfment events (10 $M_\oplus$) with great confidence. We also constructed another synthetic engulfment sample drawn from our ten twin systems excluding HAT-P-4, but with 0.1−100 $M_\oplus$ added rather than 10 $M_\oplus$ (Figure 4.3, right panel). This illustrates the $\Delta\ln(Z)$ range resulting from a large set of different engulfed masses. Many of the simulated systems with $\leq 10\ M_\oplus$ engulfment reside to the left of the maximum $\Delta\ln(Z)$ value for simulated non-engulfment systems, marked by the dashed line ($\Delta\ln(Z) = 9.15$). This further underscores that many signatures resulting from nominal 10 $M_\oplus$ engulfment events will not be identifiable with our $\Delta\ln(Z)$ analysis. The scatter in simulated engulfed mass versus $\Delta\ln(Z)$ is due to the varying stellar types of our twin systems, which result in different convective zone volumes and refractory enrichment levels for each engulfed mass amount.

4.5 Engulfment Signature Timescales

Stellar interior mixing processes deplete refractory enrichments in convective zones and weaken engulfment signatures over time. The most efficient of these processes is thermohaline mixing, a form of double-diffusive convection that operates in the presence of an inverse mean-molecular-weight (μ) gradient (e.g., Ulrich 1972; Kippenhahn, Ruschenplatt, and Thomas 1980). Accreted planetary material is initially contained within the engulfing star's convective zone, and will create an inverse μ-gradient at the convective zone base by virtue of being relatively heavy. This allows thermohaline mixing to drag engulfed material across the boundary between the convective zone and radiative stellar interior, thus attenuating photosphere refractory enrichments that compose engulfment signatures.

We ran tests with the stellar evolution code `MESA` to constrain the timescales of observable engulfment signatures considering interior mixing processes such as thermohaline instabilities. The tests involved modeling stars with masses ranging from 0.7 to 1.2 $M_\odot$ up to the zero-age main sequence (ZAMS), simulating engulfment of 1, 10, or 50 $M_\oplus$ planets via rapid accretion of bulk Earth composition material (McDonough, 2003), and evolving the stars up to the end of their main sequence lifetimes. While the 0.7−1.2 $M_\odot$ mass range covers ~80% of our engulfment sample, there are 14 stars that have masses >1.2 $M_\odot$. We initially tested `MESA` models with masses of 1.3 $M_\odot$ and 1.4 $M_\odot$, but found their engulfment signature depletion behavior to be erratic and difficult to interpret, likely due to missing physics in `MESA` relevant for this mass regime. For this reason, we confined our `MESA` models to 0.7−1.2 $M_\odot$. For each engulfment model, we ran another model of the same stellar mass but lacking bulk Earth accretion. The differential abundances produced by `MESA` between the engulfment and non-engulfment models thus mimic those of our binary observations. Relevant mixing processes were applied throughout these MESA runs, namely convective overshoot, elemental diffusion, radiative levitation (though we do not expect it to matter at these low stellar masses, e.g., Deal et al. 2020), and thermohaline mixing. Thermohaline was included in these models according to the prescription of Brown, Garaud, and Stellmach (2013), which provides a more accurate estimate of mixing efficiency compared to previous implementations (e.g., Kippenhahn, Ruschenplatt, and Thomas 1980). For more details on our `MESA` modeling procedure, see Sevilla, Behmard, and Fuller (2022) and Behmard, Sevilla, and Fuller (2023).

We note that 10 $M_\oplus$ engulfment amounts can be considered nominal as runaway gas accretion is triggered by formation of a solid 10 $M_\oplus$ core according to the core accretion model of planet formation (Wuchterl et al., 2000). Thus, most planets are expected to contain $\lesssim$10 $M_\oplus$ of refractory material. For 0.7 $M_\odot$ stars, engulfment of a 10 $M_\oplus$ planet does not produce observable enrichment (Δ[X/H] > 0.05 dex) considering the chemical dispersion observed in coeval stellar populations (0.03−0.05 dex, e.g., De Silva, Freeman, Asplund, et al. 2007; Bovy 2016; Ness et al. 2018). This is due to their deep convective envelopes, which heavily dilute accreted refractory material. This effect is less pronounced for more massive stars with thinner convective envelopes; solar-like stars (0.8−1.2 $M_\odot$) exhibit enrichments of ~0.06−0.33 dex following engulfment of a 10 $M_\oplus$ planet. Stars in the 0.8−0.9 $M_\odot$ mass range still have moderately deep convective zones, so the initial enrichment is not significantly greater that 0.05 dex, and drops below this level after ~20 Myr have passed. 1 $M_\odot$ stars maintain >0.05 dex enrichment for a longer period of ~90 Myr. This timescale is still quite small compared to typical main sequence lifetimes, implying that it will be nearly impossible to detect engulfment in 1 $M_\odot$ stars even if it happened. Higher mass stars of 1.1−1.2 $M_\odot$ exhibit the largest and longest-lived signatures, which remain above 0.05 dex levels for ~2 Gyr. Thus, these stars are the best candidates for engulfment detections. The 1.2 $M_\odot$ model exhibits a spike in iron photospheric abundance back to observable levels ~5 Gyr after engulfment due to radiative levitation. This spike lasts for ~2 Gyr, so it is possible that engulfment could also be detected in 1.2 $M_\odot$ stars if they are observed within the window of ~5−7 Gyr post-engulfment. However radiative levitation may be quite sensitive to stellar metallicity and poorly understood mixing processes not included in `MESA` (e.g., turbulence and rotational mixing). Thus, the ~5−7 Gyr post-engulfment detection window for 1.2 $M_\odot$ stars may not be reliable. Refractory depletion behavior for 10 $M_\oplus$ engulfment across the 0.7−1.2 $M_\odot$ stellar mass regime is illustrated in Figure 4.4.

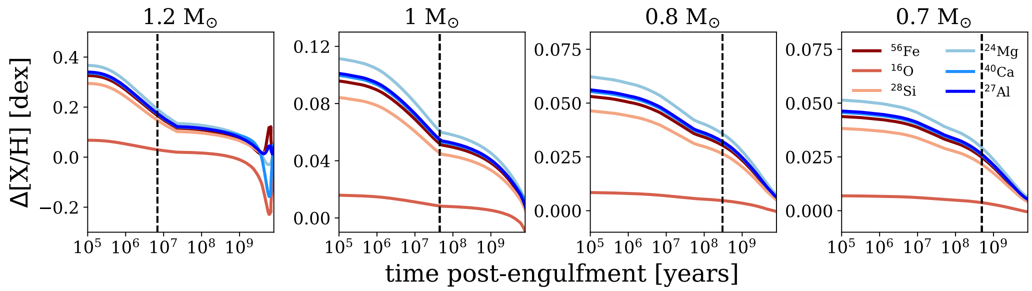


Figure 4.4: Evolution of the six most common isotopes in bulk Earth composition over time following engulfment of a 10 $M_\oplus$ bulk Earth composition planet by a 0.7−1.2 $M_\odot$ host star, as represented by abundances from MESA modeling. The abundances of a comparison model that did not undergo engulfment were subtracted off. The points at which the enrichments decrease to half their initial values post-engulfment range from ~6−500 Myr depending on the engulfing star mass. These half-life points are marked by the dashed vertical black lines.

For cases of 1 and 50 $M_\oplus$ engulfment, refractory depletion patterns across different stellar masses are similar to those of 10 $M_\oplus$ engulfment, but scaled down and up, respectively. For 1 $M_\oplus$ engulfment, stars with masses in the range 0.7−1.1 $M_\odot$ begin with enrichment levels at $\lesssim$0.05 dex, and thus never exhibit detectable engulfment signatures. However 1.2 $M_\oplus$ stars begin with >0.05 dex enrichment, and maintain this level for ~100 Myr. For 50 $M_\oplus$ engulfment, 0.7−1.2 $M_\odot$ stars maintain >0.05 dex enrichment for ~3−8 Gyr. However, planets containing up to 50 $M_\oplus$ of refractory material are predicted to be quite rare (Batygin, Bodenheimer, and Laughlin, 2016).

We ran additional MESA models with different engulfing star and accretion conditions, and found that observable signature timescales increase for sub-solar metallicities, or if engulfment occurs at later times post-ZAMS. Engulfment of a 10 $M_\oplus$ planet by a 1 $M_\odot$ sub-solar ($Z = 0.012$) metallicity star results in >0.05 dex refractory enrichment for ~3 Gyr. This is due to two effects: refractory enrichments are highlighted in low metallicity environments, and stars with low metallicities have thinner convective envelopes. Engulfment events occurring 300 Myr−3 Gyr post-ZAMS also yield signatures that remain observable on >1 Gyr timescales; 10 $M_\oplus$ engulfment by a 1 $M_\odot$ star at these times produces >0.05 dex enrichment that lasts for ~1.5 Gyr. Such late-stage engulfment results in longer observable signature timescales because refractory depletion via thermohaline is suppressed due to a counteracting positive μ-gradient from helium settling over time. Still, these timescales are short compared to main sequence lifetimes; our MESA results imply that enrichment from nominal

10 $M_\oplus$ engulfment events will rarely be observable in solar-like stars that are several Gyr old.

Twin Importance

As mentioned in Section 4.2, binary twin systems are well suited for engulfment surveys because twin companions are at the same evolutionary stage. Our `MESA` results underscore this; stars with different masses and evolutionary states exhibit different rates of refractory depletion, and Sevilla, Behmard, and Fuller (2022) found this to be true even in the absence of engulfment due to diffusion (Sevilla, Behmard, and Fuller 2022, Figure 9). This implies that non-twin binary pair stars will always have different refractory abundances, with differences increasing in time. Thus, only twin systems are capable of yielding reliable planet engulfment signatures. For a full description of our `MESA` modeling analysis and results, see Behmard, Sevilla, and Fuller (2023).

4.6 Engulfment or Primordial Differences

Before presenting our results, we outline our criteria for engulfment:

1. The stellar companions qualify as twins ($\Delta T_{\rm eff} < 200$ K, Andrews et al. 2019).
2. There is a large ($\geq$10 $M_\oplus$) amount of recovered engulfed mass from our model, with larger mass amounts considered more robust (see our section on Bayesian evidence).
3. The engulfment model shift (base of the T_c pattern across all abundances) lies above −0.05 dex. This is justified because the amount of primordial chemical dispersion between bound stellar companions is not expected to exceed 0.03−0.05 dex (e.g., De Silva, Freeman, Asplund, et al. 2007; Bovy 2016; Ness et al. 2018), and engulfment will result in a positive addition to the differential abundances.
4. These previous two conditions are satisfied across removal of each abundance, tested via applying the engulfment model after removing one abundance at a time. This leave-one-out test ensures that the T_c trends are not driven by any single abundance.
5. There is a positive ΔA(Li) between stellar companions, in the direction of potential engulfment.

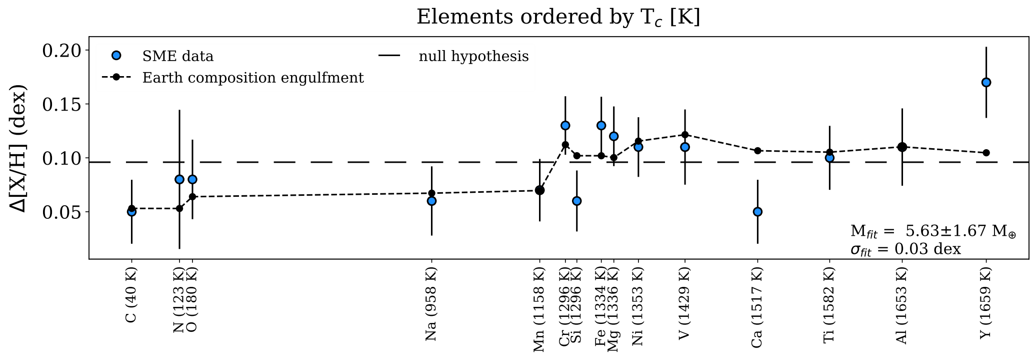


Figure 4.5: Fitted model to the differential abundance measurements between the hot Jupiter host HAT-P-4 and its companion. As in Figure 4.2, the blue circles represent the abundance differences, and the black dots are our model fit with 5.60 ± 1.64 $M_\oplus$ of bulk Earth composition engulfed material added to the convective zone of one star. The amount of modeled engulfed material and fitted scatter is provided in the lower right corner of the plot. The null hypothesis (no engulfment, but uniform abundance enrichment across all elements for one star) is shown by the long-dash line.

In light of our `MESA` results, we only considered the eleven twin systems in our sample as potential engulfment detections. However, we still applied our engulfment model to all 29 systems. Because we did not know which star in each pair may have undergone engulfment, both cases were considered for each system. All $\Delta\ln(Z)$ measurements for our engulfment sample are reported in Table 4.7. Among our eleven twin systems, only HAT-P-4 exhibits a positive Bayesian evidence difference ($\Delta\ln(Z)$ = 1.82) and an engulfment model shift that lies above −0.05 dex (Figure 4.5). The amount of recovered mass is 5.60 ± 1.64 $M_\oplus$, and remains above 5.11 ± 1.72 $M_\oplus$ across removal of each abundance. The HAT-P-4 $\Delta\ln(Z)$ value of 1.82 is well below our suggested cutoff of $\Delta\ln(Z)$ = 9.15 justified by our Bayesian evidence analysis (see our section on Bayesian evidence). Still, HAT-P-4 satisfies more of our engulfment claim criteria than any other system in our sample, making it the most promising potential engulfment detection. We note that there are five other systems (HD 99491-92, Kepler-477, Kepler-515, WASP-180, and WASP-94) with $\Delta\ln(Z)$ above the HAT-P-4 value of ∼1.82 (Table 4.7), but none satisfy the model shift above −0.05 dex criterion, and four do not qualify as twins (HD 99491-92, Kepler-477, Kepler-515, and WASP-180).

There are five systems in our sample with ΔLi > 0.1 dex and $\Delta\ln(Z)$ > 1.82, of which only two (HAT-P-4 and WASP-94) are twin binaries. HAT-P-4 and WASP-94 have Li abundances differences between the stellar companions of ΔA(Li) ≈ 0.38 ± 0.04

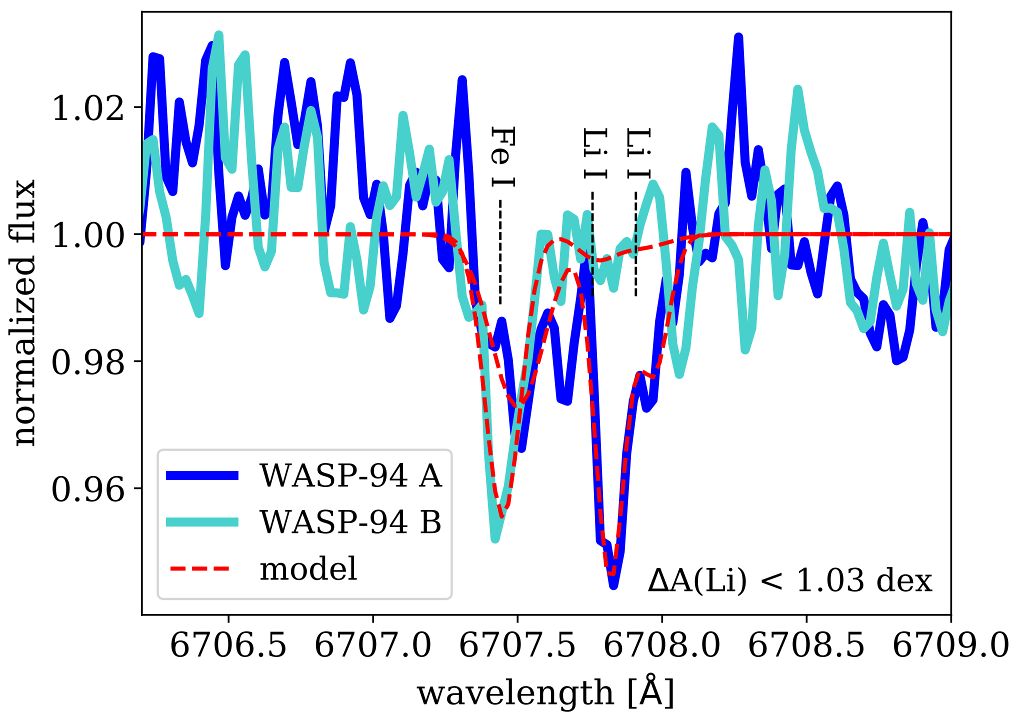


Figure 4.6: Li doublet region for WASP-94, with the normalized spectra of the stellar companions with lower and higher Li abundance plotted in light and dark blue, respectively. The model fits used to derive Li EWs and abundances are illustrated by the red dashed lines. The Fe I and Li I transitions are marked, and the differential Li abundance is provided in the lower right corner.

dex and ΔA(Li) <1.03 dex, respectively (we only report the upper limit ΔA(Li) value for WASP-94 because the Li EW is smaller than its associated error for WASP-94 B). The WASP-94 Li doublet appears quite weak (Figure 4.6). Thus, we argue that only HAT-P-4 has a ΔA(Li) potentially indicating engulfment. Kronos-Krios has a Li abundance difference of ΔA(Li) ≈ 0.51 ± 0.04 dex, which is comparable to the Li abundance difference of HAT-P-4. We plot the Li doublet regions for these systems in Figure 4.7.

To claim engulfment, we also need to verify that the differential abundance pattern supporting an engulfment scenario is not the result of primordial chemical differences between the two stellar companions. This was investigated via binary companion separations; chemical gradients could potentially increase with distance in molecular clouds, resulting in varied chemistry between widely separated stellar siblings. Thus, we must consider the possibility that large differential abundances in wide binary systems may result from primordial chemical differences rather than

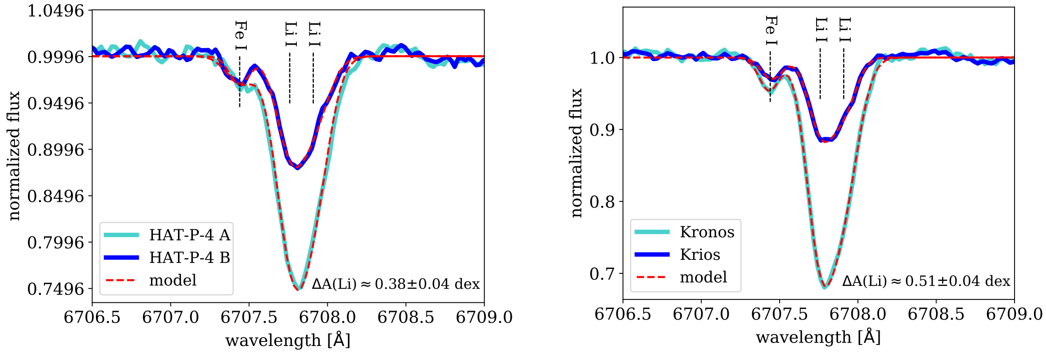


Figure 4.7: Li doublet regions for HAT-P-4 (left) and Kronos-Krios (right), with the normalized spectra of the stellar companions with lower and higher Li abundance plotted in light and dark blue, respectively. The model fits used to derive Li EWs and abundances are illustrated by the red dashed lines. The Fe I and Li I transitions are marked, and the differential Li abundances are provided in the lower right corners of the panels.

planet engulfment. There is some observational evidence for this possibility from open clusters, whose stars are widely separated by definition. Ness et al. (2018) examined pairs of red giants in seven open clusters, and found that a minority of pairs are highly chemically dissimilar according to a measure of chemical distance between the companions for 20 elements of $\chi^2 \approx 70$. For reference, most of the intra-cluster pairs are chemically homogeneous and exhibit $\chi^2 \approx 20$, corresponding to typical abundance dispersions of ~0.03 dex. Liu, Yong, Asplund, Ramírez, et al. (2016) put forward possibilities to explain such abundance differences in open clusters, such as supernova ejection in the proto-cluster cloud, or pollution of metal-poor gas. Both are contingent upon insufficient turbulent mixing within the cloud that would fail to smooth out chemical inhomogeneities.

To examine the possibility that the abundance differences of our systems are primordial, we calculated the projected separations for our 29 planet host binaries with `SME`-determined $T_{\rm eff}$ > 4700 K using *Gaia* Data Release 3 (DR3) astrometry. The errors on projected separations were taken as the scatter in calculated separations after sampling from the astrometric data uncertainty distributions 100 times for each system. These separations are reported in Table 4.7. The projected separation of HAT-P-4 is 30,000 ± 140 AU, which is larger than that of any other binary in our sample by an order of magnitude (Table 4.7). The projected separation can be considered a factor of $\sqrt{1.5}$ smaller than the true distance, and results in a value that exceeds typical turbulence scales in molecular clouds (0.05−0.2 pc, Brunt, Heyer, and Mac Low 2009 and references therein). This indicates that the HAT-P-4 stellar companions may have formed in distinct areas of chemodynamical space within their birth cloud. Thus, we regard the HAT-P-4 differential abundance pattern as potentially due to primordial chemical differences between the two stars rather than planet engulfment.

4.7 Assessment of Published Systems

There are ten planet host binary systems with high-precision abundances previously measured (HAT-P-1, HD 20781-82, XO-2, WASP-94, HAT-P-4, HD 80606-07, 16-Cygni, HD 133131, HD 106515, WASP-160; Table 4.1). Depending on the study, four to six of these systems are claimed as engulfment detections. Because no potential engulfment signatures were found in our sample aside from HAT-P-4, we were interested in testing if previously reported datasets for these ten systems yield robust signatures according to our engulfment model.

We found that six of the systems exhibit $\Delta\ln(Z) > 1.82$, above HAT-P-4 (16 Cygni, XO-2, HD 20781-82, HD 133131, WASP-94, and WASP-160). However this depends on the reported dataset; the 16 Cygni abundances derived by Tucci Maia, Meléndez, and Ramírez (2014), Tucci Maia, Meléndez, Lorenzo-Oliveira, et al. (2019), and Ryabchikova et al. (2022) are above this cutoff, but those of Ramírez, Meléndez, et al. (2011) yield a negative $\Delta\ln(Z)$. Likewise, the XO-2 abundances derived by Ramírez, Khanal, Aleo, et al. (2015) and Biazzo et al. (2015) pass the HAT-P-4 cutoff, but those of Teske, Ghezzi, et al. (2015) yield a negative $\Delta\ln(Z)$. The Ramírez, Meléndez, et al. (2011) and Teske, Ghezzi, et al. (2015) studies did not claim engulfment. Our fitted engulfment model to the Ryabchikova et al. (2022) 16 Cygni dataset also exhibits a shift below -0.05 dex, which violates our engulfment criteria. This is also true for the Mack, Schuler, et al. (2014) and Teske, Khanal, and Ramírez (2016) datasets for HD 20781-82 and WASP-94, respectively. The Teske, Shectman, et al. (2016) HD 133131 dataset passes this engulfment model shift criterion, but yields a small engulfed mass estimate ($M = 1.13 \pm 0.51\ M_\oplus$), and is not claimed as engulfment by Teske, Shectman, et al. (2016). This leaves the Jofré et al. (2021) WASP-160 dataset, which yields an estimated engulfed mass of $M = 7.73 \pm 1.59\ M_\oplus$ and $\Delta\ln(Z) = 9.37$. However WASP-160 is part of our sample, and our `SME` abundances do not clearly favor an engulfment scenario ($\Delta\ln(Z) = 0.98$). We conclude that there is no strong evidence for engulfment detections in the literature aside from potentially Kronos-Krios.

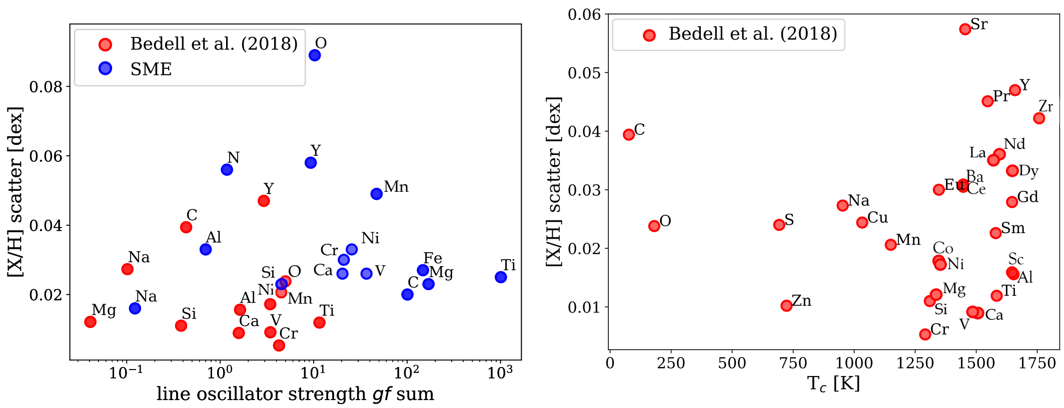


Figure 4.8: The left panel displays the scatter in abundance measurements vs. the sum of oscillator strengths gf for each element in the `SME` line list and our engulfment sample (blue points), and for the Bedell, Bean, et al. (2018) line list and solar twin sample (red points). The abundance scatter between companions in our binary sample and across the sample of solar twins increases as the included lines for each abundance become fewer and weaker for oscillator strength sums of $gf <$ 10. The right panel displays the scatter in abundance measurements vs. T_c for the Bedell, Bean, et al. (2018) sample considering all elements reported in their study.

Abundance Scatter

Abundance discrepancies between different studies of the same stars can be attributed to usage of different instruments (e.g., Bedell, Meléndez, et al. 2014); differences in the acquired spectra such as varying SNR levels (e.g., Liu, Yong, Asplund, Feltzing, et al. 2018); or to differences in abundance measurement pipelines that may employ different spectral synthesis codes, continuum placement, EW measurement procedures, and line lists (e.g., Schuler et al. 2011; Liu, Yong, Asplund, Feltzing, et al. 2018). A few studies that exemplify these discrepancy sources are Saffe, Flores, and Buccino (2015), Mack, Stassun, et al. (2016), and Liu, Yong, Asplund, Feltzing, et al. (2018), which all analyzed HD 80606-07, but derived widely varying abundance measurements. Saffe, Flores, and Buccino (2015) and Mack, Stassun, et al. (2016) used the same set of Keck-HIRES observations, but derived abundances that often do not agree within their combined uncertainties at the 1σ level. Liu, Yong, Asplund, Feltzing, et al. (2018) obtained higher SNR observations of HD 80606-07, and claimed that their abundance measurements are more reliable because their average uncertainties (~0.007 dex) are much smaller than those of Saffe, Flores, and Buccino (2015) and Mack, Stassun, et al. (2016) (0.02 dex and 0.027 dex, respectively).

These three studies also employed different line lists. The Saffe, Flores, and Buccino

(2015) list includes the highest number of lines at ∼500, followed by the Liu, Yong, Asplund, Feltzing, et al. (2018) list with ∼250 lines, then the Mack, Stassun, et al. (2016) list with ∼125 lines. To quantify the quality of these different line lists, we calculated the summed oscillator strength gf over each line corresponding to a single abundance. As expected, this quantity is a factor of 2−4 higher for the Liu, Yong, Asplund, Feltzing, et al. (2018) and Saffe, Flores, and Buccino (2015) line lists compared to the Mack, Stassun, et al. (2016) line list averaging across all abundances. This is likely responsible for the approximate abundance measurement agreement between Liu, Yong, Asplund, Feltzing, et al. (2018) and Saffe, Flores, and Buccino (2015), but not Mack, Stassun, et al. (2016). For comparison, the line list we employed in our `SME` analysis includes over 7500 lines, making the summed gf quantity ∼100 times higher than that of the Saffe, Flores, and Buccino (2015) line list. The average difference for our `SME`-derived HD 80606-07 abundances is +0.006, also in better agreement with Liu, Yong, Asplund, Feltzing, et al. (2018) and Saffe, Flores, and Buccino (2015) compared to Mack, Stassun, et al. (2016).

We were interested in quantifying how line lists affect abundance measurements by examining if abundance prediction scatter changes as a function of line number and strength. We tested our `SME` line list against the abundance scatter between companions in the ten twin systems excluding HAT-P-4 from our engulfment sample, and found that abundances with fewer and weaker lines according to oscillator strength gf (e.g., O, Y, N) exhibit larger abundance prediction scatter (Figure 4.8, left panel, blue points). This indicates that scatter is large for volatile and highly refractory abundances that anchor the lower and upper portions of the T_c trend, respectively. We carried out the same analysis for the Bedell, Bean, et al. (2018) sample of solar twins and the line list used in their `MOOG` analysis, and found the same trend of abundance scatter increasing with fewer and weaker lines per abundance (Figure 4.8, left panel, red points). We also examined the Bedell, Bean, et al. (2018) abundance scatter as a function of T_c, and found that abundances with low (e.g., C and O) and high T_c (e.g., Zr and Y) exhibit large scatter similar to our `SME` results (Figure 4.8, right panel). These findings show that large line lists with strong spectral features are necessary for measuring precise abundances, and elements that anchor the T_c trend lack an abundance of strong features and thus exhibit large scatter. This is unsurprising for the low T_c abundances; volatile elements like C, N, and O are often locked in molecular species that create blended features, making it difficult to identify strong, well-isolated lines. Because elements important for establishing a T_c trend tend to have large uncertainties, we expect that a T_c pattern can occur randomly in the absence of engulfment.

4.8 Discussion

We did not recover any strong planet engulfment detections in our planet host binary sample. HAT-P-4 is the only system whose abundances exhibit a possible engulfment signature. This binary is composed of two solar-like (G0V + G2V) stars, with the primary hosting a 0.68 $M_{\rm Jup}$ hot Jupiter at an orbital period of ~3 days (Kovács et al., 2007). Our engulfment model recovers 5.60 ± 1.64 $M_\oplus$ of engulfed mass by the planet host star. However, HAT-P-4's $\Delta \ln(Z)$ value of ~1.82 does not strongly support an engulfment claim, and the system sustains only $\Delta \ln(Z) = 1$ across the leave-one-out abundance test. For reference, these values are well below the maximum $\Delta \ln(Z)$ value of our synthetic non-engulfment systems (9.15, see the section on Bayesian evidence), indicating that the HAT-P-4 engulfment signature could be a false positive. Other signposts of engulfment in the HAT-P-4 systems include evidence of host star spin up (Oetjens et al., 2020) or increased luminosity (Yarza et al., 2022) post-engulfment. However, the difference in $v \sin i$ between HAT-P-4 A and B is only 1.59 ± 1.41 km/s, while engulfment of ~5 $M_\oplus$ planets is predicted to increase $v \sin i$ by at least 10 km/s (Oetjens et al., 2020). Similarly, increased host star luminosity post-engulfment is predicted to last for <1000 years (Yarza et al., 2022). We thus conclude that increases in host star luminosity or rotation cannot constitute evidence for engulfment in the HAT-P-4 system.

Alternatively, the abundance pattern of HAT-P-4 may be primordial rather than due to engulfment. HAT-P-4's projected separation (30,000 ± 140 AU) is an order of magnitude larger than that of any other binary in our sample (Table 4.7), and exceeds the lower bound of typical turbulence scales in molecular clouds (0.05−0.2 pc, Brunt, Heyer, and Mac Low 2009 and references therein). This suggests that HAT-P-4 A and B formed far from each other within their birth cloud, and were separated by large chemical gradients that gave rise to the differential abundance pattern we see today. It is possible that the Kronos-Krios abundance pattern is also primordial; we calculated the projected separation for this system to be 11,000 ± 12 AU. There is also a tentative trend of increasing abundance difference as a function of binary separation in our sample of eleven twin systems. To illustrate this, we plot their Δ[Fe/H] as a function of separation in Figure 4.9, along with those of the 25 wide binaries from Hawkins et al. (2020).

While a T_c-dependent abundance pattern is a signpost of planet engulfment, it is possible that the T_c-dependent patterns of HAT-P-4 and Kronos-Krios occurred in the absence of engulfment because of large uncertainties on abundances that anchor the upper and lower portions of the T_c trend. To test this, we simulated 1000 systems assuming the HAT-P-4 and Kronos-Krios companion masses and convective zones, but with abundances drawn from Gaussian distributions with widths equal to the average abundance scatter per element between the companions of our ten twin systems excluding HAT-P-4 (Figure 4.8, right panel, blue points). There are 33 simulated systems with $\Delta\ln(Z)$ values that exceed that of HAT-P-4 ($\Delta\ln(Z) > 1.82$), of which two also have recovered amounts of engulfed mass greater than HAT-P-4's value of 5.60 $M_\oplus$. However, there are no simulated systems with $\Delta\ln(Z)$ or recovered amounts of engulfed mass greater than those of Kronos-Krios (Figure 4.10). We conclude that the HAT-P-4 T_c-dependent abundance pattern can occur randomly in the absence of engulfment, but not that of Kronos-Krios. Thus, Kronos-Krios may be a true engulfment detection whereas HAT-P-4 is likely not.

The lack of clear engulfment detections in our sample can be explained by our `MESA` analysis (Behmard, Sevilla, and Fuller, 2023), which predicts that observable refractory enrichments from 10 $M_\oplus$ engulfment events occurring at ZAMS will become depleted on timescales of ~2 Myr−2 Gyr for solar-like (0.8−1.2 $M_\odot$) stars. The largest and longest-lived signatures are exhibited by 1.1−1.2 $M_\odot$ stars (~2 Gyr). We thus recommend these stars as the best candidates for engulfment detections. We also considered other engulfment scenarios assuming a 1 $M_\odot$ star, and found that engulfment signature timescales increase to ~1.5 Gyr for late-stage (300 Myr−3 Gyr post-ZAMS) engulfment, and ~3 Gyr for sub-solar (Z = 0.012) engulfing star metallicities. Most (~85% within mass measurement error) of the stars composing the 29 binaries in our sample assessed for engulfment signatures are in the solar-like mass range. In addition, there are only two systems younger than 2 Gyr (HD 202772 and WASP-180), and only 1 system younger than 3 Gyr with sub-solar metallicities (Kepler-477). Thus, the timescales of observable signatures from nominal 10 $M_\oplus$ engulfment are short compared to the system lifetimes. Our `MESA` results also show that refractory enhancements exhibit half-lives of ~6−500 Myr (Figure 4.4). This implies that unless the engulfment event happened recently, we can only recover clear engulfment signatures by taking observations soon after the engulfment event. Perhaps this is the case for Kronos-Krios assuming it is a true engulfment detection.

Our `MESA` results also underscore the importance of using stellar twin binaries for planet engulfment surveys. Refractory depletion rates vary as a function of

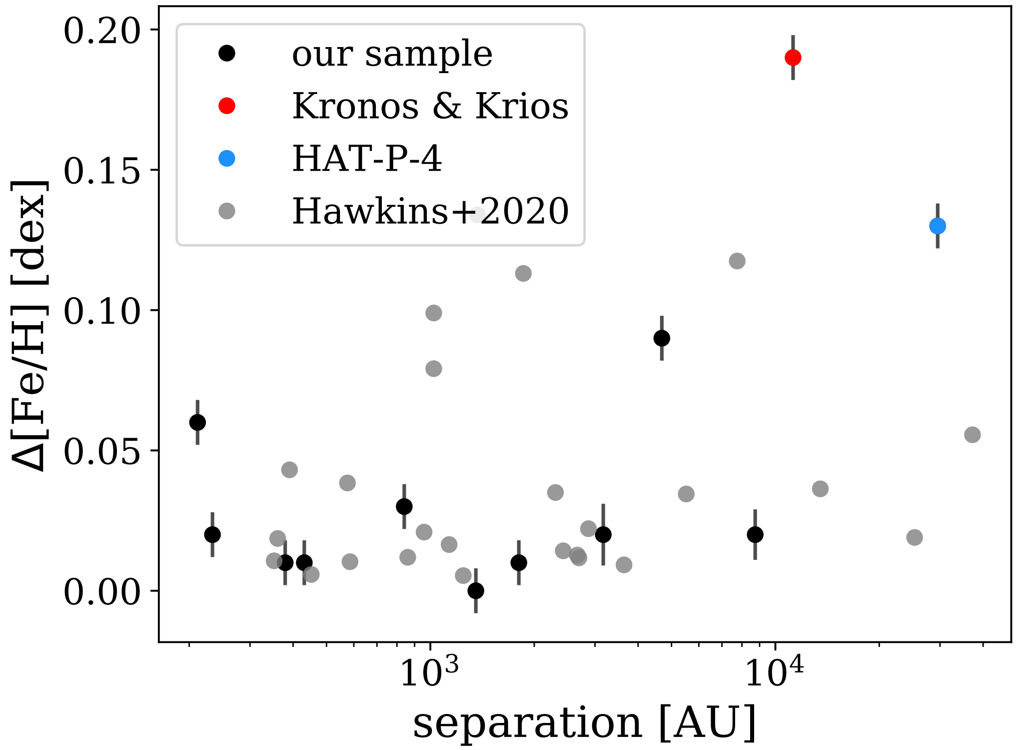


Figure 4.9: Δ[Fe/H] (right) vs. projected binary separation of the eleven twin systems in our sample assessed for engulfment signatures (black). The 25 wide binaries from Hawkins et al. (2020) are plotted for comparison (gray), along with the points representing Kronos-Krios (red) and HAT-P-4 (blue). There appears to be a trend of increasing Δ[Fe/H] as a function of separation across all samples. The Spina et al. (2021) systems are not explicitly shown because most of their systems that qualify as twins are drawn from the Hawkins et al. (2020) sample, and the remaining do not have reported separations.

engulfing star mass and spectral type, even in the absence of planet engulfment (Sevilla, Behmard, and Fuller 2022, Figure 9). Thus, non-twin stellar siblings will always exhibit different photospheric abundances. As mentioned in Section 4.2, only eleven of the 36 binaries in our sample qualify as twins. This is another potential contributing factor to our lack of engulfment detections. We thus recommend that future engulfment surveys focus solely on stellar twin systems. Considering the eleven twin systems in our sample, we calculated an upper limit engulfment detection rate for our study using the observable signature timescales from our `MESA` analysis. This rate was taken as the average in log space of signature timescales (which varies as a function of engulfing star mass) over system age ratios for the eleven twin systems. The resulting rate is ~4.9%, though we note that the true rate

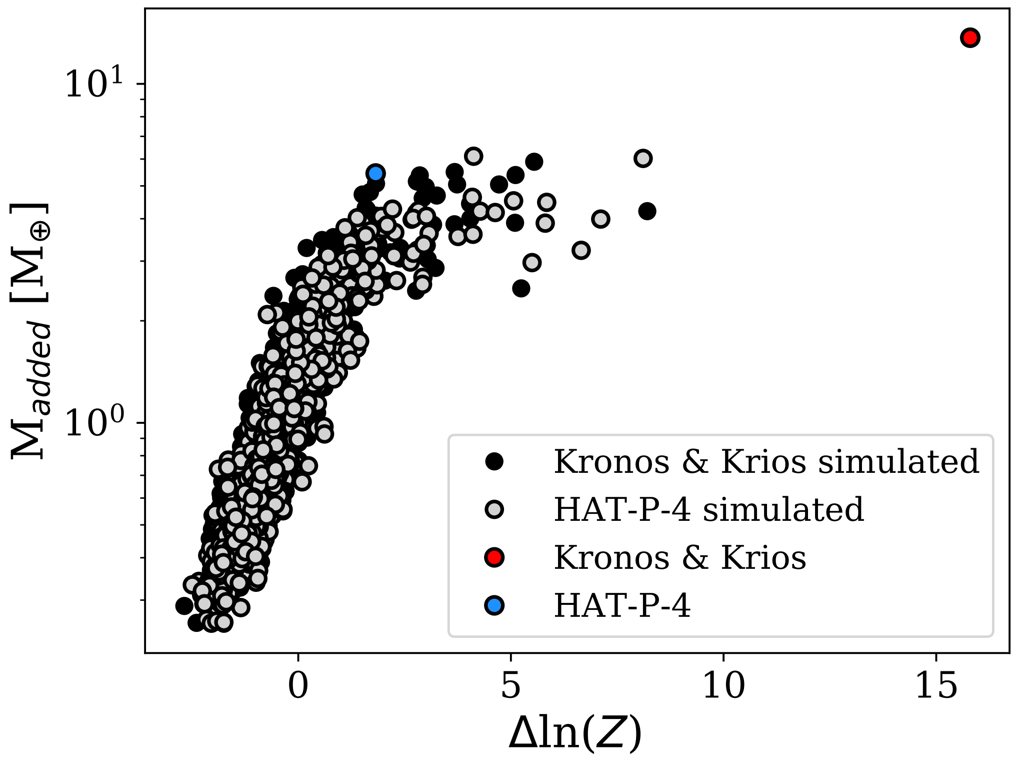


Figure 4.10: Estimated amounts of engulfed material from our engulfment model fits vs. $\Delta\ln(Z)$ values for the 1000 simulated systems assuming the HAT-P-4 companion masses and convective zones (gray), and the 1000 simulated systems assuming those of Kronos-Krios (black). The abundances of the simulated systems were drawn from Gaussian distributions with widths equal to the average abundance scatter per element of our ten twin systems excluding HAT-P-4. The real data for Kronos-Krios and HAT-P-4 are also shown for comparison as the red and blue dots, respectively. The simulated systems can mimic the T_c trend of HAT-P-4, but not that of Kronos-Krios.

will be much lower since it should be multiplied by a factor corresponding to the intrinsic engulfment rate, which is unknown.

Our results are in contradiction with previous studies that report high rates of engulfment detections. For example, Spina et al. (2021) claim an engulfment rate of ∼20−35% for their sample of 107 binary systems. They based this on a large fraction of systems ("chemically anomalous" pairs) with high [Fe/C] ratios, Δ[Fe/H], and ΔA(Li). No other abundances were examined and thus there is no analysis of T_c trends. In addition, the elemental abundances of the 107 systems were derived from multiple literature sources that took observations with different instruments, and employed different spectral synthesis pipelines and line lists (Desidera, Gratton,

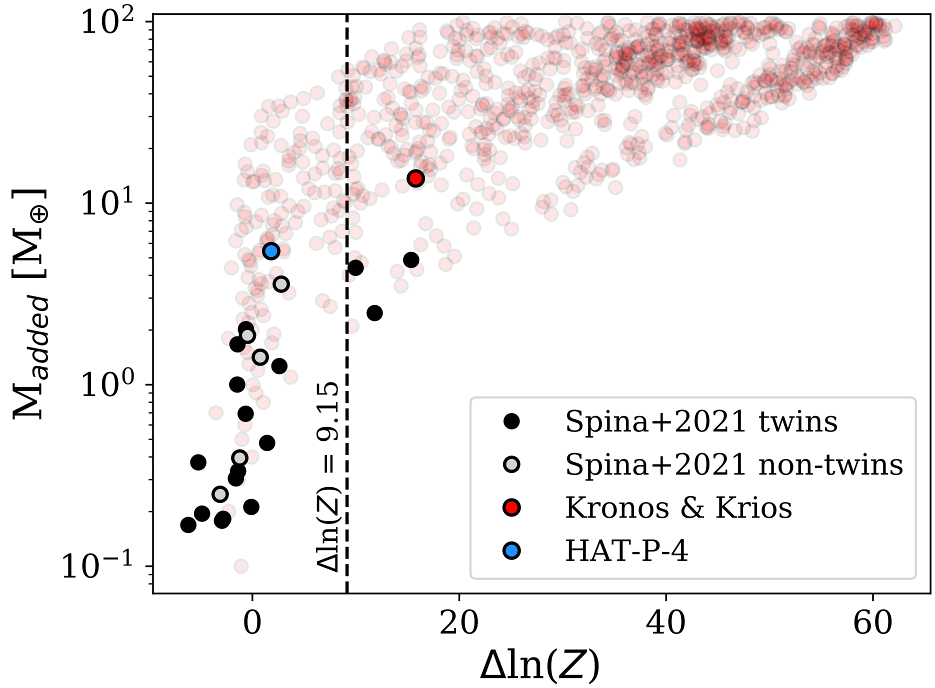


Figure 4.11: Estimated amounts of engulfed material from our engulfment model fits vs. $\Delta\ln(Z)$ values for the "chemically anomalous" systems reported Spina et al. (2021). The twin ($\Delta T_{\text{eff}} < 200$ K, Andrews et al. 2019) pairs are shown in black while the non-twin pairs are shown in gray. The points corresponding to 0.1−100 $M_\oplus$ simulated engulfment systems are represented by the transparent red dots in the background for comparison. Kronos-Krios and HAT-P-4 are also shown for comparison as the red and blue dots, respectively. The maximum $\Delta\ln(Z)$ value for the synthetic non-engulfment distribution of 9.15 is marked by the dashed line.

Scuderi, et al., 2004; Desidera, Gratton, Lucatello, et al., 2006; Hawkins et al., 2020; Nagar, Spina, and Karakas, 2020). Such heterogeneous methods can introduce systematic bias into abundance samples (e.g., Schuler et al. 2011; Liu, Yong, Asplund, Feltzing, et al. 2018). Finally, many of the binaries employed in this study do not qualify as stellar twins; Spina et al. (2021) imposed a ΔT_{eff} cutoff of 600 K. Thus, we argue that Spina et al. (2021) lack sufficient evidence for their ~20−35% engulfment rate claim.

Spina et al. (2021) based their claim on 33 "chemically anomalous" pairs among their total sample of 107 binaries. Eleven of these 33 pairs were observed by Spina et al. (2021) with the HARPS spectrograph and analyzed with `MOOG`. The abundance measurements of the remaining pairs are drawn from other catalogs; another eleven systems are from Hawkins et al. (2020), four from Desidera, Gratton, Lucatello, et al. (2006), two from Nagar, Spina, and Karakas (2020), and one from Desidera, Gratton, Scuderi, et al. (2004). The last four systems are included in our sample (Kronos-Krios, HAT-P-4, 16 Cygni, XO-2). As discussed earlier, we only consider Kronos-Krios and HAT-P-4 as potential engulfment detections. We assessed the other "chemically anomalous" pairs as follows. The Desidera, Gratton, Scuderi, et al. (2004), Desidera, Gratton, Lucatello, et al. (2006), and Nagar, Spina, and Karakas (2020) studies do not provide abundances beyond Fe, but Spina et al. (2021) and Hawkins et al. (2020) measured a set of abundances spanning a wide range of T_c (e.g., C, N, Mn, Cr, Si, Fe, Mg, Ni, V, Ca, Ti, Al, Y). We analyzed the 22 Spina et al. (2021) and Hawkins et al. (2020) "chemically anomalous" pairs with our engulfment model considering the abundances listed above. The fitted amounts of engulfed material from our engulfment model vs. $\Delta\ln(Z)$ values for these pairs are shown in Figure 4.11, with twin ($\Delta T_{\rm eff} < 200$ K, Andrews et al. 2019) pairs represented by black dots and non-twin pairs represented by gray dots. Three systems exhibit $\Delta\ln(Z)$ values above 9.15, the maximum $\Delta\ln(Z)$ of our synthetic non-engulfment systems. All three systems qualify as binary twins. They have $2.48-4.87\ M_\oplus$ fitted amounts of engulfed material from our engulfment model, and $\Delta\ln(Z)$ values ranging from 10.0−15.4. We conclude that these three systems are potential engulfment detections.

We carried out a similar analysis to estimate the mass of engulfed material for the remaining seven Desidera, Gratton, Scuderi, et al. (2004), Desidera, Gratton, Lucatello, et al. (2006), and Nagar, Spina, and Karakas (2020) "chemically anomalous" pairs considering just Fe and its abundance in bulk Earth compositions, and estimated $1.27-12.34\ M_\oplus$ amounts of engulfed material. Only three of these seven systems qualify as twins. If we consider just the twin pairs, the amount of engulfed material drops to $1.27-2.50\ M_\oplus$, and there are no $\Delta\ln(Z)$ values to provide further evidence for these systems as engulfment detections. We thus conclude that there are only five potential detections (Kronos-Krios, HAT-P-4, and three additional systems) in the Spina et al. (2021) sample of 107 systems, yielding an engulfment rate of ~4.7%. However, Kronos-Krios, HAT-P-4, 16 Cygni, and XO-2 were likely included because they are reported as possible engulfment detections in previous studies (Table 4.1). If we remove these systems, there are only three potential detections out of 103 systems, yielding an engulfment rate of ~2.9%. This is much lower than the ~20−35% Spina et al. (2021) engulfment rate claim.

We conclude that engulfment detections are rare, and put forward the possibility that the abundance differences of HAT-P-4 are primordial. Those of Kronos-Krios may also be primordial, but there is evidence that this system is a true engulfment detection because its strong T_c trend is not produced randomly from large uncertainties on low and high T_c abundances. Considering the HAT-P-4 case, if large (Δ[X/H] > 0.05 dex) primordial abundance differences between binary companions are common, it may not be safe to assume that stellar siblings born from the same molecular cloud are always chemically homogeneous. This could undermine the validity of galactic archaeological tools used to trace stars back to their parent clouds, namely chemical tagging. There are hints that chemical tagging may have limitations. As mentioned in Section 4.6, Ness et al. (2018) found a small population of chemically inhomogeneous red giant pairs in open clusters. Similarly, the Hawkins et al. (2020) study of 25 wide binaries reported that while 80% are homogeneous to 0.02 dex levels, six pairs exhibit Δ[Fe/H] > 0.05 dex. Larger wide binary pair samples could be used to place upper limits on abundance differences as a function of separation, and may be aided by ongoing high-resolution spectroscopic surveys such as APOGEE (Abdurro'uf et al., 2022) and GALAH (Buder et al., 2021).

4.9 Summary

We carried out a Keck-HIRES survey of 36 planet host binaries and examined their differential stellar abundances for evidence of planet engulfment. However we reiterate that only eleven of these 36 binaries qualify as stellar twins ($\Delta T_{\text{eff}} < 200$ K, Andrews et al. 2019), and our `MESA` results show that reliable engulfment signatures can only be detected in twin systems because refractory depletion rates vary as a function of engulfing star type. None of the systems in our sample exhibit clear engulfment signatures, which dovetails with our `MESA` results that show observable signatures in solar-like (0.8−1.2 $M_\odot$) stars are depleted below observable levels (Δ[X/H] > 0.05 dex) within ~2 Gyr after the engulfment event (Behmard, Sevilla, and Fuller, 2023). Only one of our twin binary systems, HD 202772, has an age below 2 Gyr (1.8 Gyr, Wang et al. 2019).

Among our twin systems, only HAT-P-4 exhibits a possible engulfment signature. If engulfment occurred in this system, it must have happened within the last 2 Gyr, which is less than half of HAT-P-4's estimated age (4.2 Gyr, Ment et al. 2018). This makes the engulfment scenario somewhat unlikely. Alternatively, HAT-P-4's abundance differences may be primordial as evidenced by the large projected separation (30,000 ± 140 AU) between the binary companions. This projected separation is larger than that of any other system in our sample by an order of magnitude (Table 4.7). Similarly, we suggest that the Kronos-Krios abundances differences may be primordial based on the large projected separation of the system (11,000 ± 12 AU).

We used our engulfment model to analyze previously published datasets for ten planet host binary systems (HAT-P-1, HD 20781-82, XO-2, WASP-94, HAT-P-4, HD 80606-07, 16-Cygni, HD 133131, HD 106515, WASP-160; Table 4.1), of which four to six are claimed as engulfment detections depending on the study. None of the systems can be claimed as detections according to our criteria for engulfment, outlined in Section 4.6. We also examined how abundance scatter depends on line lists employed in spectral synthesis pipelines, and found that abundance precision increases with larger numbers of strong spectral features per chemical species (Figure 4.8, left panel). Elements with low T_c (e.g., volatiles such as C, N, O), and high T_c (e.g., Y) lack an abundance of strong features and thus exhibit large scatter. Because these abundances are important for anchoring T_c trends, we conclude that T_c patterns can randomly result from poorly measured abundances in the absence of engulfment (Figure 4.8, right panel). We tested if the HAT-P-4 and Kronos-Krios T_c trends can be randomly produced from large uncertainties on low and high T_c abundances, and found that this is the case for HAT-P-4, but not Kronos-Krios. We conclude that Kronos-Krios may still be a true engulfment detection, but HAT-P-4 is likely not.

Our results contradict previous studies that report high rates of engulfment, namely Spina et al. (2021) that claimed an engulfment rate of ~20−35% for their sample of 107 binary systems. We analyzed the abundance patterns of their "chemically anomalous" systems with our engulfment model, and determined that the true engulfment rate is closer to ~2.9%. This is comparable to the upper limit engulfment detection rate we calculated from our `MESA` engulfment signature timescales (~4.9%). Our results suggest that reported detections of planet engulfment may instead be due to primordial chemical differences between stellar companions. To confirm this, the homogeneity of bound stellar siblings as a function of binary separation should be investigated further in future studies.

4.10 Appendix: Abundance Error Analysis

We obtained Keck-HIRES observations of eight bright stars (HIP 38931, HIP 44137, HIP 47288, HIP 16107, HIP 14300, HIP 15099, HIP 14241, HIP 21272) at five different SNR levels, and calculated the variance in their `SME` abundance predictions. These eight stars span the $T_{\rm eff}$ range of our engulfment sample, and a wide [Fe/H] range of −0.39 to +0.37 dex, making the results of this test relevant for a diverse set of stars. We collected 3–6 spectra per star and SNR level, and found that the variance in measured abundances is $\lesssim$0.03 dex for refractory species, and higher for volatile species with variance up to ~0.1 dex. As expected, the variance decreases dramatically as a function of SNR for all abundances and stars, with the exception of HIP 38931 which exhibits large scatter in the volatile abundances even as the highest SNR level (200/pix) is approached. This is likely due to its low temperature ($T_{\rm eff}$ = 4680 K) and low metallicity ([Fe/H] = −0.17 dex) which together create a favorable environment for forming volatile-bearing molecular species whose spectral features are difficult to model with `SME`. Because of this, we consider `SME`-determined abundances for targets with $T_{\rm eff} < 4700$ K and sub-solar metallicities to be suspect.

Excluding HIP 38931, we found that the abundance scatter of the remaining seven bright stars agrees with the abundance errors reported in Brewer and Fischer (2018). To illustrate this, we plotted the standard deviation of C, N, O, and Fe `SME` abundance predictions for these seven stars against the Brewer and Fischer (2018) solar spectra abundance scatter for SNR levels of 40, 60, 80, and 100 in Figure 4.12. We chose these abundances because spectral synthesis codes like `SME` struggle to model the features of volatile species like C, N, and O due to molecular lines, and Fe provides a good comparison point by possessing many easily modeled lines. As expected, the abundance scatter trends as a function of SNR are approximately monotonic, though the scatter of HIP 47288, HIP 14300, and HIP 14241 noticeably deviate for C and N. This is likely because these are the most metal poor stars remaining in our now seven bright star sample, and are thus more likely to host volatile-bearing molecules in their photospheres.

It can be seen that the Brewer and Fischer (2018) predictions (gray circles) well-represent the abundance scatter across all seven stars that we observed (colored circles). The average absolute difference between the abundance scatter reported by Brewer and Fischer (2018) and those of our seven bright stars is ~0.012 dex, and the Brewer and Fischer (2018) scatter is larger ~57% of the time across all SNR levels and abundances. Thus, the Brewer and Fischer (2018) abundance scatter is a good approximation of `SME` abundance errors for our engulfment sample, and we derived our errors by linearly interpolating through SNR in Table 2 of Brewer and Fischer (2018) to match the individual SNR of each star (Table 4.5).

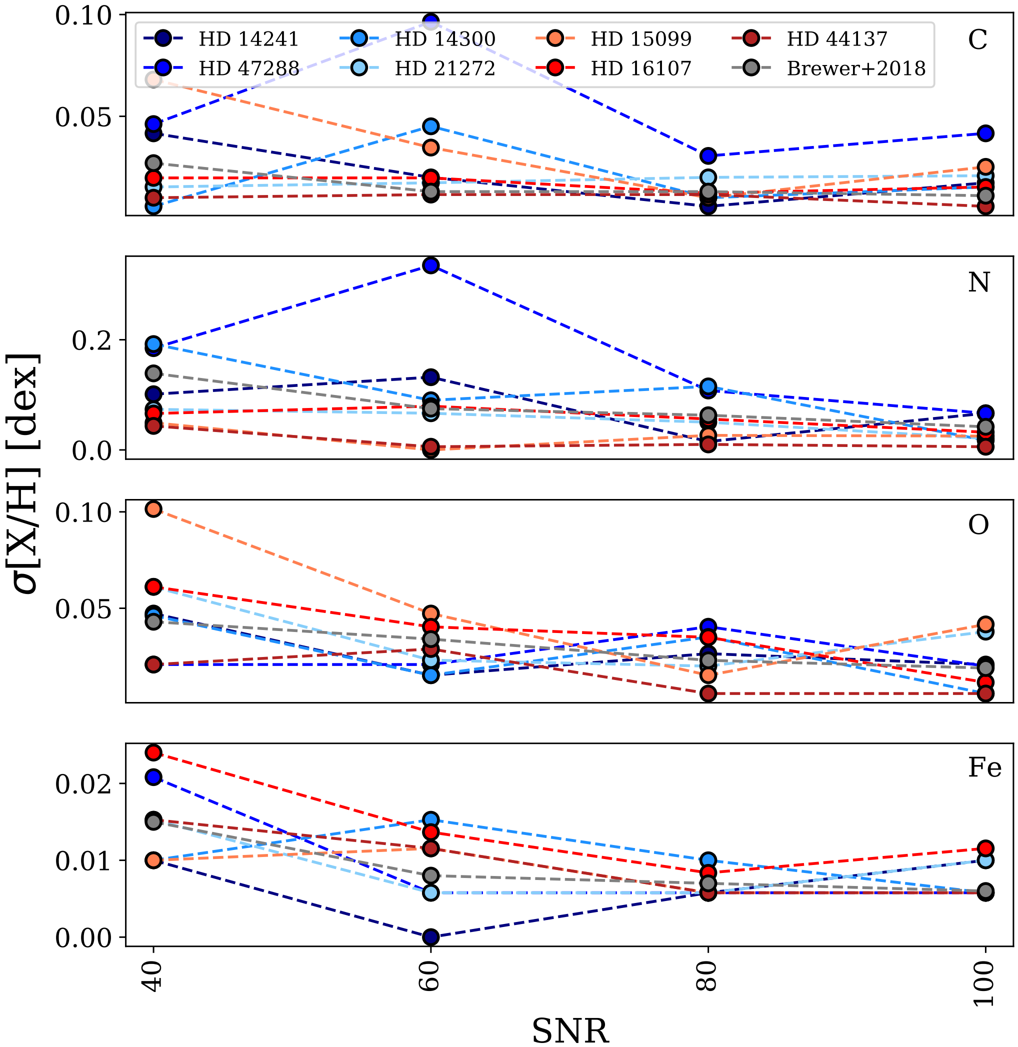


Figure 4.12: Standard deviation in SME abundance predictions from multiple HIRES observations of seven bright stars (colored circles), and the Brewer and Fischer (2018) scatter in SME abundance predictions for a solar spectrum with varying amount of added Gaussian random noise to mimic varying SNR levels (gray). The seven bright stars are colored in order of increasing metallicity (dark blue to dark red). The abundances displayed are C, N, O, and Fe from top to bottom, as a function of SNR = 40, 60, 80, and 100.

References

Abdurro'uf et al. (Apr. 2022). "The Seventeenth Data Release of the Sloan Digital Sky Surveys: Complete Release of MaNGA, MaStar, and APOGEE-2 Data". In: *The Astrophysical Journal Supplement Series* 259.2, 35, p. 35. DOI: 10.3847/1538-4365/ac4414. arXiv: 2112.02026 [astro-ph.GA].

Andrews, Jeff J. et al. (Jan. 2019). "Using APOGEE Wide Binaries to Test Chemical Tagging with Dwarf Stars". In: *The Astrophysical Journal* 871.1, 42, p. 42. DOI: 10.3847/1538-4357/aaf502. arXiv: 1811.12032 [astro-ph.SR].

Baraffe, Isabelle et al. (May 2015). "New evolutionary models for pre-main sequence and main sequence low-mass stars down to the hydrogen-burning limit". In: *Astronomy and Astrophysics* 577, A42, A42. DOI: 10.1051/0004-6361/201425481. arXiv: 1503.04107 [astro-ph.SR].

Batygin, Konstantin, Peter H. Bodenheimer, and Gregory P. Laughlin (Oct. 2016). "In Situ Formation and Dynamical Evolution of Hot Jupiter Systems". In: *The Astrophysical Journal* 829.2, 114, p. 114. DOI: 10.3847/0004-637X/829/2/114. arXiv: 1511.09157 [astro-ph.EP].

Bedell, Megan, Jacob L. Bean, et al. (Sept. 2018). "The Chemical Homogeneity of Sun-like Stars in the Solar Neighborhood". In: *The Astrophysical Journal* 865.1, 68, p. 68. DOI: 10.3847/1538-4357/aad908. arXiv: 1802.02576 [astro-ph.SR].

Bedell, Megan, Jorge Meléndez, et al. (Nov. 2014). "Stellar Chemical Abundances: In Pursuit of the Highest Achievable Precision". In: *The Astrophysical Journal* 795.1, 23, p. 23. DOI: 10.1088/0004-637X/795/1/23. arXiv: 1409.1230 [astro-ph.SR].

Behmard, Aida, Jason Sevilla, and Jim Fuller (Feb. 2023). "Planet engulfment signatures in twin stars". In: *Monthly Notices of the Royal Astronomical Society* 518.4, pp. 5465–5474. DOI: 10.1093/mnras/stac3435. arXiv: 2210.11679 [astro-ph.SR].

Berger, Travis A., Andrew W. Howard, and Ann Merchant Boesgaard (Mar. 2018). "Identifying Young Kepler Planet Host Stars from Keck-HIRES Spectra of Lithium". In: *The Astrophysical Journal* 855.2, 115, p. 115. DOI: 10.3847/1538-4357/aab154. arXiv: 1802.09529 [astro-ph.EP].

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*. Ed. by Masanori Iye and Alan F. M. Moorwood. Vol. 4841. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, pp. 1694–1704. DOI: 10.1117/12.461502.

Bertran de Lis, S. et al. (Apr. 2015). "Oxygen abundances in G- and F-type stars from HARPS. Comparison of [OI] 6300 Å and OI 6158 Å". In: *Astronomy and Astrophysics* 576, A89, A89. DOI: 10.1051/0004-6361/201424633. arXiv: 1501.05805 [astro-ph.SR].

Biazzo, K. et al. (Nov. 2015). "The GAPS programme with HARPS-N at TNG. X. Differential abundances in the XO-2 planet-hosting binary". In: *Astronomy and Astrophysics* 583, A135, A135. DOI: 10.1051/0004-6361/201526375. arXiv: 1506.01614 [astro-ph.SR].

Bland-Hawthorn, Joss, Mark R. Krumholz, and Ken Freeman (Apr. 2010). "The Long-term Evolution of the Galactic Disk Traced by Dissolving Star Clusters". In: *The Astrophysical Journal* 713.1, pp. 166–179. DOI: 10.1088/0004-637X/713/1/166. arXiv: 1002.4357 [astro-ph.GA].

Booth, Richard A. and James E. Owen (Apr. 2020). "Fingerprints of giant planets in the composition of solar twins". In: *Monthly Notices of The Royal Astronomical Society* 493.4, pp. 5079–5088. DOI: 10.1093/mnras/staa578. arXiv: 2002.11135 [astro-ph.EP].

Boubert, D. et al. (June 2019). "Lessons from the curious case of the 'fastest' star in Gaia DR2". In: *Monthly Notices of The Royal Astronomical Society* 486.2, pp. 2618–2630. DOI: 10.1093/mnras/stz253. arXiv: 1901.10460 [astro-ph.SR].

Bovy, Jo (Jan. 2016). "The Chemical Homogeneity of Open Clusters". In: *The Astrophysical Journal* 817.1, 49, p. 49. DOI: 10.3847/0004-637X/817/1/49. arXiv: 1510.06745 [astro-ph.GA].

Brewer, J. M. and D. A. Fischer (Aug. 2018). "Spectral Properties of Cool Stars: Extended Abundance Analysis of Kepler Objects of Interest". In: *The Astrophysical Journal Supplement Series* 237, 38, p. 38. DOI: 10.3847/1538-4365/aad501. arXiv: 1804.00673 [astro-ph.SR].

Brewer, J. M., D. A. Fischer, et al. (Aug. 2016). "Spectral Properties of Cool Stars: Extended Abundance Analysis of 1,617 Planet-search Stars". In: *The Astrophysical Journal Supplement Series* 225, 32, p. 32. DOI: 10.3847/0067-0049/225/2/32. arXiv: 1606.07929 [astro-ph.SR].

Brown, Justin M., Pascale Garaud, and Stephan Stellmach (Apr. 2013). "Chemical Transport and Spontaneous Layer Formation in Fingering Convection in Astrophysics". In: *The Astrophysical Journal* 768.1, p. 34. ISSN: 1538-4357. DOI: 10.1088/0004-637x/768/1/34. arXiv: 1212.1688. URL: http://dx.doi.org/10.1088/0004-637X/768/1/34.

Brunt, C. M., M. H. Heyer, and M. -M. Mac Low (Sept. 2009). "Turbulent driving scales in molecular clouds". In: *Astronomy and Astrophysics* 504.3, pp. 883–890. DOI: 10.1051/0004-6361/200911797. arXiv: 0910.0398 [astro-ph.GA].

Buder, Sven et al. (Sept. 2021). "The GALAH+ survey: Third data release". In: *Monthly Notices of The Royal Astronomical Society* 506.1, pp. 150–201. DOI: 10.1093/mnras/stab1242. arXiv: 2011.02505 [astro-ph.GA].

Cayrel, R. (Jan. 1988). "Data Analysis". In: *The Impact of Very High S/N Spectroscopy on Stellar Physics*. Ed. by G. Cayrel de Strobel and Monique Spite. Vol. 132, p. 345.

Chambers, J. E. (Nov. 2010). "Stellar Elemental Abundance Patterns: Implications for Planet Formation". In: *The Astrophysical Journal* 724, pp. 92–97. DOI: 10.1088/0004-637X/724/1/92.

Chene, André-Nicolas et al. (July 2014). "GRACES: Gemini remote access to CFHT ESPaDOnS spectrograph through the longest astronomical fiber ever made: experimental phase completed". In: *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation*. Vol. 9151. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 915147, p. 915147. DOI: 10.1117/12.2057417. arXiv: 1409.7448 [astro-ph.IM].

Cosentino, Rosario et al. (2012). "Harps-N: the new planet hunter at TNG". In: *Ground-based and Airborne Instrumentation for Astronomy IV*. Ed. by Ian S. McLean, Suzanne K. Ramsay, and Hideki Takami. Vol. 8446. International Society for Optics and Photonics. SPIE, pp. 657–676. DOI: 10.1117/12.925738. URL: https://doi.org/10.1117/12.925738.

Currie, Thayne et al. (Nov. 2007). "Discovery of Gas Accretion onto Stars in 13 Myr old h and χ Persei". In: *The Astrophysical Journal Letters* 669.1, pp. L33–L36. DOI: 10.1086/523595. arXiv: 0709.1847 [astro-ph].

De Silva, G. M., K. C. Freeman, M. Asplund, et al. (Mar. 2007). "Chemical Homogeneity in Collinder 261 and Implications for Chemical Tagging". In: *The Astronomical Journal* 133.3, pp. 1161–1175. DOI: 10.1086/511182. arXiv: astro-ph/0611832 [astro-ph].

De Silva, G. M., K. C. Freeman, and J. Bland-Hawthorn (Apr. 2009). "Reconstructing Fossil Sub-structures of the Galactic Disk: Clues from Abundance Patterns of Old Open Clusters and Moving Groups". In: *Publications of the Astronomical Society of the Pacific* 26, pp. 11–16. DOI: 10.1071/AS08019. arXiv: 0810.2287.

Deal, M. et al. (Jan. 2020). "Chemical mixing in low mass stars. I. Rotation against atomic diffusion including radiative acceleration". In: *Astronomy and Astrophysics* 633, A23, A23. DOI: 10.1051/0004-6361/201936666. arXiv: 1910.14335 [astro-ph.SR].

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: *Optical and IR Telescope Instrumentation and Detectors*. Ed. by Masanori Iye and Alan F. Moorwood. Vol. 4008. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, pp. 534–545. DOI: 10.1117/12.395512.

Desidera, S., R. G. Gratton, S. Lucatello, et al. (Aug. 2006). "Abundance difference between components of wide binaries. II. The southern sample". In: *Astronomy and Astrophysics* 454.2, pp. 581–593. DOI: 10.1051/0004-6361:20064896.

Desidera, S., R. G. Gratton, S. Scuderi, et al. (June 2004). "Abundance difference between components of wide binaries". In: *Astronomy and Astrophysics* 420, pp. 683–697. DOI: 10.1051/0004-6361:20041242. arXiv: astro-ph/0403051 [astro-ph].

Dressing, Courtney D. et al. (Feb. 2015). "The Mass of Kepler-93b and the Composition of Terrestrial Planets". In: *The Astrophysical Journal* 800.2, p. 135. ISSN: 1538-4357. DOI: 10.1088/0004-637x/800/2/135. URL: http://dx.doi.org/10.1088/0004-637X/800/2/135.

Farihi, J. (Apr. 2016). "Circumstellar debris and pollution at white dwarf stars". In: *New Astronomy Review* 71, pp. 9–34. DOI: 10.1016/j.newar.2016.03.001. arXiv: 1604.03092 [astro-ph.EP].

Gaia Collaboration et al. (Aug. 2018). "Gaia Data Release 2. Summary of the contents and survey properties". In: *Astronomy and Astrophysics* 616, A1, A1. DOI: 10.1051/0004-6361/201833051. arXiv: 1804.09365 [astro-ph.GA].

Hawkins, Keith et al. (Feb. 2020). "Identical or fraternal twins? The chemical homogeneity of wide binaries from Gaia DR2". In: *Monthly Notices of The Royal Astronomical Society* 492.1, pp. 1164–1179. DOI: 10.1093/mnras/stz3132. arXiv: 1912.08895 [astro-ph.SR].

Howard, A. W. et al. (Oct. 2010). "The California Planet Survey. I. Four New Giant Exoplanets". In: *The Astrophysical Journal* 721, pp. 1467–1481. DOI: 10.1088/0004-637X/721/2/1467. arXiv: 1003.3488 [astro-ph.EP].

Hughes, D. W., R. Rosner, and N. O. Weiss (2007). *The Solar Tachocline*.

Jofré, Emiliano et al. (Dec. 2021). "The Peculiar Chemical Pattern of the WASP-160 Binary System: Signatures of Planetary Formation and Evolution?" In: *The Astronomical Journal* 162.6, 291, p. 291. DOI: 10.3847/1538-3881/ac25ef. arXiv: 2109.04590 [astro-ph.EP].

Kippenhahn, R., G. Ruschenplatt, and H. -C. Thomas (Nov. 1980). "The time scale of thermohaline mixing in stars". In: *Astronomy and Astrophysics* 91.1-2, pp. 175–180.

Klein, B. et al. (Feb. 2010). "Chemical Abundances in the Externally Polluted White Dwarf GD 40: Evidence of a Rocky Extrasolar Minor Planet". In: *The Astrophysical Journal* 709.2, pp. 950–962. DOI: 10.1088/0004-637X/709/2/950. arXiv: 0912.1422 [astro-ph.EP].

Koester, D., B. T. Gänsicke, and J. Farihi (June 2014). "The frequency of planetary debris around young white dwarfs". In: *Astronomy and Astrophysics* 566, A34, A34. DOI: 10.1051/0004-6361/201423691. arXiv: 1404.2617 [astro-ph.SR].

Kolbl, Rea et al. (Jan. 2015). "Detection of Stars Within ~0.8 in of Kepler Objects of Interest". In: *The Astronomical Journal* 149.1, 18, p. 18. DOI: 10.1088/0004-6256/149/1/18. arXiv: 1412.5259 [astro-ph.SR].

Kovács, G. et al. (Nov. 2007). "HAT-P-4b: A Metal-rich Low-Density Transiting Hot Jupiter". In: *The Astrophysical Journal Letters* 670.1, pp. L41–L44. DOI: 10.1086/524058. arXiv: 0710.0602 [astro-ph].

Krumholz, Mark R., Christopher F. McKee, and Joss Bland-Hawthorn (Aug. 2019). "Star Clusters Across Cosmic Time". In: *Annual Review of Astronomy and Astrophysics* 57, pp. 227–303. DOI: 10.1146/annurev-astro-091918-104430. arXiv: 1812.01615 [astro-ph.GA].

Li, Min and Lin Xiao (Mar. 2016). "Lifetimes and Accretion Rates of Protoplanetary Disks". In: *The Astrophysical Journal* 820.1, 36, p. 36. DOI: 10.3847/0004-637X/820/1/36.

Liu, F., M. Asplund, I. Ramírez, et al. (May 2014). "A high-precision chemical abundance analysis of the HAT-P-1 stellar binary: constraints on planet formation". In: *Monthly Notices of the Royal Astronomical Society: Letters* 442.1, pp. L51–L55. ISSN: 1745-3925. DOI: 10.1093/mnrasl/slu055. eprint: https://academic.oup.com/mnrasl/article-pdf/442/1/L51/14117201/slu055.pdf. URL: https://doi.org/10.1093/mnrasl/slu055.

Liu, F., M. Asplund, D. Yong, et al. (Nov. 2016). "The chemical compositions of solar twins in the open cluster M67". In: *Monthly Notices of The Royal Astronomical Society* 463.1, pp. 696–704. DOI: 10.1093/mnras/stw2045. arXiv: 1608.03788 [astro-ph.SR].

Liu, F., D. Yong, M. Asplund, S. Feltzing, et al. (July 2018). "Detailed chemical compositions of the wide binary HD 80606/80607: revised stellar properties and constraints on planet formation". In: *Astronomy and Astrophysics* 614, A138, A138. DOI: 10.1051/0004-6361/201832701. arXiv: 1802.09306 [astro-ph.SR].

Liu, F., D. Yong, M. Asplund, I. Ramírez, et al. (Apr. 2016). "The Hyades open cluster is chemically inhomogeneous". In: *Monthly Notices of The Royal Astronomical Society* 457.4, pp. 3934–3948. DOI: 10.1093/mnras/stw247. arXiv: 1601.07354 [astro-ph.SR].

Liu, Fan, Bertram Bitsch, et al. (Sept. 2021). "Detailed elemental abundances of binary stars: Searching for signatures of planet formation and atomic diffusion". In: *Monthly Notices of The Royal Astronomical Society*. DOI: 10.1093/mnras/stab2471. arXiv: 2108.11001 [astro-ph.SR].

Lodders, Katharina (July 2003). "Solar System Abundances and Condensation Temperatures of the Elements". In: *The Astrophysical Journal* 591.2, pp. 1220–1247. DOI: 10.1086/375492.

Mack Claude E., III, Simon C. Schuler, et al. (June 2014). "Detailed Abundances of Planet-hosting Wide Binaries. I. Did Planet Formation Imprint Chemical Signatures in the Atmospheres of HD 20782/81?" In: *The Astrophysical Journal* 787.2, 98, p. 98. DOI: 10.1088/0004-637X/787/2/98. arXiv: 1404.1967 [astro-ph.EP].

Mack Claude E., III, Keivan G. Stassun, et al. (Feb. 2016). "Detailed Abundances of Planet-hosting Wide Binaries. II. HD80606+HD80607". In: *The Astrophysical Journal* 818.1, 54, p. 54. DOI: 10.3847/0004-637X/818/1/54. arXiv: 1601.00018 [astro-ph.SR].

McDonough, W. F. (Dec. 2003). “Compositional Model for the Earth’s Core”. In: *Treatise on Geochemistry* 2, p. 568. DOI: 10.1016/B0-08-043751-6/02015-6.

Meléndez, J. et al. (Oct. 2009). “The Peculiar Solar Composition and Its Possible Relation to Planet Formation”. In: *The Astrophysical Journal Letters* 704.1, pp. L66–L70. DOI: 10.1088/0004-637X/704/1/L66. arXiv: 0909.2299 [astro-ph.SR].

Ment, Kristo et al. (Nov. 2018). “Radial Velocities from the N2K Project: Six New Cold Gas Giant Planets Orbiting HD 55696, HD 98736, HD 148164, HD 203473, and HD 211810”. In: *The Astronomical Journal* 156.5, 213, p. 213. DOI: 10.3847/1538-3881/aae1f5. arXiv: 1809.01228 [astro-ph.EP].

Mugrauer, M (Nov. 2019). “Search for stellar companions of exoplanet host stars by exploring the second ESA-Gaia data release”. In: *Monthly Notices of the Royal Astronomical Society* 490.4, pp. 5088–5102. ISSN: 0035-8711. DOI: 10.1093/mnras/stz2673. eprint: https://academic.oup.com/mnras/article-pdf/490/4/5088/30656870/stz2673.pdf. URL: https://doi.org/10.1093/mnras/stz2673.

Nagar, Tushar, Lorenzo Spina, and Amanda I. Karakas (Jan. 2020). “The Chemical Signatures of Planetary Engulfment Events in Binary Systems”. In: *The Astrophysical Journal Letters* 888.1, L9, p. L9. DOI: 10.3847/2041-8213/ab5dc6.

Ness, M. et al. (Feb. 2018). “Galactic Doppelgängers: The Chemical Similarity Among Field Stars and Among Stars with a Common Birth Origin”. In: *The Astrophysical Journal* 853.2, 198, p. 198. DOI: 10.3847/1538-4357/aa9d8e. arXiv: 1701.07829 [astro-ph.SR].

Newville, Matthew et al. (Sept. 2014). *LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python*. Zenodo. Version 0.8.0. DOI: 10.5281/zenodo.11813.

Noguchi, Kunio et al. (Dec. 2002). “High Dispersion Spectrograph (HDS) for the Subaru Telescope”. In: *Publications of the Astronomical Society of Japan* 54.6, pp. 855–864. ISSN: 0004-6264. DOI: 10.1093/pasj/54.6.855. eprint: https://academic.oup.com/pasj/article-pdf/54/6/855/17446863/pasj54-0855.pdf. URL: https://doi.org/10.1093/pasj/54.6.855.

Oetjens, A. et al. (Nov. 2020). “The influence of planetary engulfment on stellar rotation in metal-poor main-sequence stars”. In: *Astronomy and Astrophysics* 643, A34, A34. DOI: 10.1051/0004-6361/202038653. arXiv: 2009.03624 [astro-ph.SR].

Oh, Semyeong, Adrian M. Price-Whelan, John M. Brewer, et al. (Feb. 2018). “Kronos and Krios: Evidence for Accretion of a Massive, Rocky Planetary System in a Comoving Pair of Solar-type Stars”. In: *The Astrophysical Journal* 854.2, 138, p. 138. DOI: 10.3847/1538-4357/aaab4d. arXiv: 1709.05344 [astro-ph.SR].

Oh, Semyeong, Adrian M. Price-Whelan, David W. Hogg, et al. (June 2017). “Co-moving Stars in Gaia DR1: An Abundance of Very Wide Separation Comoving Pairs”. In: *The Astronomical Journal* 153.6, 257, p. 257. DOI: 10.3847/1538-3881/aa6ffd. arXiv: 1612.02440 [astro-ph.SR].

Petigura, E. A. (2015). “Prevalence of Earth-size Planets Orbiting Sun-like Stars”. PhD thesis. University of California, Berkeley.

Petit, P., J. -F. Donati, and Espadons Project Team (Jan. 2003). “Stellar polarimetry with ESPaDOnS”. In: *EAS Publications Series*. Ed. by J. Arnaud and N. Meunier. Vol. 9. EAS Publications Series, p. 97. arXiv: astro-ph/0403118 [astro-ph].

Pinsonneault, M. H., D. L. DePoy, and M. Coffee (July 2001). “The Mass of the Convective Zone in FGK Main-Sequence Stars and the Effect of Accreted Planetary Material on Apparent Metallicity Determinations”. In: *The Astrophysical Journal Letters* 556.1, pp. L59–L62. DOI: 10.1086/323531. arXiv: astro-ph/0105257 [astro-ph].

Putirka, Keith D. and Siyi Xu (2021). “Polluted white dwarfs reveal exotic mantle rock types on exoplanets in our solar neighborhood”. In: *Nature Communications* 12.1, p. 6168. DOI: 10.1038/s41467-021-26403-8. URL: https://doi.org/10.1038/s41467-021-26403-8.

Ramírez, I., S. Khanal, P. Aleo, et al. (July 2015). “The Dissimilar Chemical Composition of the Planet-hosting Stars of the XO-2 Binary System”. In: *The Astrophysical Journal* 808, 13, p. 13. DOI: 10.1088/0004-637X/808/1/13. arXiv: 1506.01025 [astro-ph.SR].

Ramírez, I., S. Khanal, S. J. Lichon, et al. (Dec. 2019). “The chemical composition of HIP 34407/HIP 34426 and other twin-star comoving pairs”. In: *Monthly Notices of The Royal Astronomical Society* 490.2, pp. 2448–2457. DOI: 10.1093/mnras/stz2709. arXiv: 1909.07460 [astro-ph.SR].

Ramírez, I., J. Meléndez, et al. (Oct. 2011). “Elemental Abundance Differences in the 16 Cygni Binary System: A Signature of Gas Giant Planet Formation?” In: *The Astrophysical Journal* 740.2, 76, p. 76. DOI: 10.1088/0004-637X/740/2/76. arXiv: 1107.5814 [astro-ph.SR].

Randich, Sofia and Laura Magrini (Mar. 2021). “Light Elements in the Universe”. In: *Frontiers in Astronomy and Space Sciences* 8, 6, p. 6. DOI: 10.3389/fspas.2021.616201. arXiv: 2103.11000 [astro-ph.GA].

Rasio, Frederic A. and Eric B. Ford (1996). “Dynamical Instabilities and the Formation of Extrasolar Planetary Systems”. In: *Science* 274.5289, pp. 954–956. ISSN: 0036-8075. DOI: 10.1126/science.274.5289.954. eprint: https://science.sciencemag.org/content/274/5289/954.full.pdf. URL: https://science.sciencemag.org/content/274/5289/954.

Ryabchikova, T. et al. (June 2022). "Detailed abundances of the wide pairs of stars with and without planets: the binary systems 16 Cyg and HD 219542". In: *arXiv e-prints*, arXiv:2206.07463, arXiv:2206.07463. arXiv: 2206.07463 [astro-ph.SR].

Saffe, C., M. Flores, and A. Buccino (Oct. 2015). "HD 80606: searching for the chemical signature of planet formation". In: *Astronomy and Astrophysics* 582, A17, A17. DOI: 10.1051/0004-6361/201526644. arXiv: 1507.08125 [astro-ph.SR].

Saffe, C., E. Jofré, E. Martioli, et al. (July 2017). "Signatures of rocky planet engulfment in HAT-P-4. Implications for chemical tagging studies". In: *Astronomy and Astrophysics* 604, L4, p. L4. DOI: 10.1051/0004-6361/201731430. arXiv: 1707.02180 [astro-ph.SR].

Saffe, C., E. Jofré, P. Miquelarena, et al. (May 2019). "High-precision analysis of binary stars with planets. I. Searching for condensation temperature trends in the HD 106515 system". In: *Astronomy and Astrophysics* 625, A39, A39. DOI: 10.1051/0004-6361/201935352. arXiv: 1904.01955 [astro-ph.SR].

Schuler, Simon C. et al. (May 2011). "Abundances of Stars with Planets: Trends with Condensation Temperature". In: *The Astrophysical Journal* 732.1, 55, p. 55. DOI: 10.1088/0004-637X/732/1/55. arXiv: 1103.0757 [astro-ph.SR].

Sevilla, Jason, Aida Behmard, and Jim Fuller (Nov. 2022). "Long-term lithium abundance signatures following planetary engulfment". In: *Monthly Notices of The Royal Astronomical Society* 516.3, pp. 3354–3365. DOI: 10.1093/mnras/stac2436. arXiv: 2207.13232 [astro-ph.SR].

Sneden, C. A. (1973). "Carbon and Nitrogen Abundances in Metal-Poor Stars." PhD thesis. THE UNIVERSITY OF TEXAS AT AUSTIN.

Sobeck, Jennifer S. et al. (Apr. 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* 141.6, p. 175. DOI: 10.1088/0004-6256/141/6/175. URL: https://doi.org/10.1088%2F0004-6256%2F141%2F6%2F175.

Speagle, Joshua S. (Apr. 2020). "DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences". In: *Monthly Notices of the Royal Astronomical Society* 493.3, pp. 3132–3158. DOI: 10.1093/mnras/staa278. arXiv: 1904.02180 [astro-ph.IM].

Spina, Lorenzo et al. (Aug. 2021). "Chemical evidence for planetary ingestion in a quarter of Sun-like stars". In: *Nature Astronomy*. DOI: 10.1038/s41550-021-01451-8. arXiv: 2108.12040 [astro-ph.SR].

Teske, Johanna K., Luan Ghezzi, et al. (Mar. 2015). "Abundance Differences between Exoplanet Binary Host Stars XO-2N and XO-2S−Dependence on Stellar Parameters". In: *The Astrophysical Journal Letters* 801.1, L10, p. L10. DOI: 10.1088/2041-8205/801/1/L10. arXiv: 1501.02167 [astro-ph.EP].

Teske, Johanna K., Sandhya Khanal, and Ivan Ramírez (Mar. 2016). "The Curious Case of Elemental Abundance Differences in the Dual Hot Jupiter Hosts WASP-94A and B". In: *The Astrophysical Journal* 819.1, 19, p. 19. DOI: 10.3847/0004-637X/819/1/19. arXiv: 1601.01731 [astro-ph.SR].

Teske, Johanna K., Stephen A. Shectman, et al. (Dec. 2016). "The Magellan PFS Planet Search Program: Radial Velocity and Stellar Abundance Analyses of the 360 au, Metal-poor Binary "Twins" HD 133131A & B". In: *The Astronomical Journal* 152.6, 167, p. 167. DOI: 10.3847/0004-6256/152/6/167. arXiv: 1608.06216 [astro-ph.EP].

Tucci Maia, M., J. Meléndez, D. Lorenzo-Oliveira, et al. (Aug. 2019). "Revisiting the 16 Cygni planet host at unprecedented precision and exploring automated tools for precise abundances". In: *Astronomy and Astrophysics* 628, A126, A126. DOI: 10.1051/0004-6361/201935952. arXiv: 1906.04195 [astro-ph.SR].

Tucci Maia, M., J. Meléndez, and I. Ramírez (Aug. 2014). "High Precision Abundances in the 16 Cyg Binary System: A Signature of the Rocky Core in the Giant Planet". In: *The Astrophysical Journal Letters* 790, L25, p. L25. DOI: 10.1088/2041-8205/790/2/L25. arXiv: 1407.4132 [astro-ph.SR].

Tull, R. G. (June 1972). "The coudé spectrograph and echelle scanner of the 2.7 m telescope at McDonald Observatory." In: *Auxiliary Instrumentation for Large Telescopes*, pp. 259–274.

Ulrich, Roger K. (Feb. 1972). "Thermohaline Convection in Stellar Interiors." In: *The Astrophysical Journal* 172, p. 165. DOI: 10.1086/151336.

Valenti, J. A. and N. Piskunov (Sept. 1996). "Spectroscopy made easy: A new tool for fitting observations with synthetic spectra." In: *Astronomy and Astrophysics, Supplement* 118, pp. 595–603.

Vogt, S. S. et al. (June 1994). "HIRES: the high-resolution echelle spectrometer on the Keck 10-m Telescope". In: *Instrumentation in Astronomy VIII*. Ed. by D. L. Crawford and E. R. Craine. Vol. 2198. Proceedings of SPIE, p. 362. DOI: 10.1117/12.176725.

Wallace, L. et al. (July 2011). "An Optical and Near-infrared (2958-9250 Å) Solar Flux Atlas". In: *The Astrophysical Journal Supplement Series* 195.1, 6, p. 6. DOI: 10.1088/0067-0049/195/1/6.

Wang, Songhu et al. (Feb. 2019). "HD 202772A b: A Transiting Hot Jupiter around a Bright, Mildly Evolved Star in a Visual Binary Discovered by TESS". In: *The Astronomical Journal* 157.2, 51, p. 51. DOI: 10.3847/1538-3881/aaf1b7. arXiv: 1810.02341 [astro-ph.EP].

Weidenschilling, Stuart J. and Francesco Marzari (Dec. 1996). "Gravitational scattering as a possible origin for giant planets at small stellar distances". In: *Nature* 384.6610, pp. 619–621. DOI: 10.1038/384619a0.

White, Russel J. and Lynne A. Hillenbrand (Mar. 2005). "A Long-lived Accretion Disk around a Lithium-depleted Binary T Tauri Star". In: *The Astrophysical Journal Letters* 621.1, pp. L65–L68. DOI: 10.1086/428752. arXiv: astro-ph/0501307 [astro-ph].

Wuchterl, G et al. (2000). "Protostars and Planets IV". In: *ed. Mannings, V., Boss, AP & Russell, SS, University of Arizona Press: Tucson*, pp. 1081–1110.

Xu, Siyi et al. (Nov. 2019). "Compositions of Planetary Debris around Dusty White Dwarfs". In: *The Astronomical Journal* 158.6, p. 242. DOI: 10.3847/1538-3881/ab4cee. URL: https://doi.org/10.3847%2F1538-3881%2Fab4cee.

Yarza, Ricardo et al. (Mar. 2022). "Hydrodynamics and survivability during post-main-sequence planetary engulfment". In: *arXiv e-prints*, arXiv:2203.11227, arXiv:2203.11227. DOI: 10.48550/arXiv.2203.11227. arXiv: 2203.11227 [astro-ph.SR].

Yee, S. W., E. A. Petigura, and K. von Braun (Feb. 2017). "Precision Stellar Characterization of FGKM Stars using an Empirical Spectral Library". In: *The Astrophysical Journal* 836, 77, p. 77. DOI: 10.3847/1538-4357/836/1/77. arXiv: 1701.00922 [astro-ph.SR].

Zuckerman, B., D. Koester, et al. (Dec. 2007). "The Chemical Composition of an Extrasolar Minor Planet". In: *The Astrophysical Journal* 671.1, pp. 872–877. DOI: 10.1086/522223. arXiv: 0708.0198 [astro-ph].

Zuckerman, B., C. Melis, et al. (Oct. 2010). "Ancient Planetary Systems are Orbiting a Large Fraction of White Dwarf Stars". In: *The Astrophysical Journal* 722.1, pp. 725–736. DOI: 10.1088/0004-637X/722/1/725. arXiv: 1007.2252 [astro-ph.SR].

Table 4.1: Planet Host Binaries with Previously Measured High-Precision Abundances

System	$\|\Delta T_{\mathrm{eff}}\|$ K	$\|\Delta \log g\|$ dex	sep AU	$\|\Delta[\mathrm{Fe/H}]\|$ dex	Instrument	T_c trend	Source
HAT-P-1	17	0.07	1550	0.009 ± 0.009	Keck-HIRES[a]	no	Liu, Asplund, Ramírez, et al. (2014)
HD 20781-82	465	0.10	9000	0.060 ± 0.010	Magellan-MIKE[b]	yes	Mack, Schuler, et al. (2014)
XO-2	60	0.02	4500	0.054 ± 0.005	Subaru-HDS[c]; Keck-HIRES[a] HARPS-N[d]	maybe/yes	Teske, Ghezzi, et al. (2015); Ramírez, Khanal, Aleo, et al. (2015); Biazzo et al. (2015)
WASP-94	82	0.09	2700	0.015 ± 0.004	Magellan-MIKE[b]	yes	Teske, Khanal, and Ramírez (2016)
HAT-P-4	10	0.05	28446	0.105 ± 0.006	Gemini-GRACES[e]	yes	Saffe, Jofré, Martioli, et al. (2017)
HD 80606-07	52	0.04	1200	0.000 ± 0.040	Keck-HIRES[a]	no	Saffe, Flores, and Buccino (2015) and Mack, Stassun, et al. (2016); Liu, Yong, Asplund, Feltzing, et al. (2018)
16 Cygni	79	0.05	860	0.047 ± 0.005	McDonald-RGT[f] CFHT-ESPaDOnS[g] Subaru-HDS[c]	yes/no	Ramírez, Meléndez, et al. (2011); Tucci Maia, Meléndez, and Ramírez (2014); Tucci Maia, Meléndez, Lorenzo-Oliveira, et al. (2019) Ryabchikova et al. (2022)
HD 133131	5	0.0	360	0.032 ± 0.015	Magellan-MIKE[b], VLT-UVES[h]	maybe	Teske, Shectman, et al. (2016) and Liu, Bitsch, et al. (2021)
HD 106515	250	0.09	860	0.008 ± 0.01	Keck-HIRES[a], VLT-UVES[h]	no	Saffe, Jofré, Miquelarena, et al. (2019) and Liu, Bitsch, et al. (2021)
WASP-160	8060	0.05	860	0.012 ± 0.017	Gemini-GRACES[e]	yes	Jofré et al. (2021)

(a) Keck High Resolution Echelle Spectrograph (Vogt et al., 1994), (b) Magellan Inamori Kyocera Echelle high-resolution spectrograph (Bernstein et al., 2003), (c) High Dispersion Spectrograph (Noguchi et al., 2002), (d) Telescopio Nazionale Galileo HARPS-N Spectrograph (Cosentino et al., 2012), (e) Gemini Remote Access to CFHT ESPaDOnS Spectrograph (Chene et al., 2014), (f) McDonald Observatory R.G. Tull spectrograph (Tull, 1972), (g) Canada-France-Hawaii Telescope ESPaDOnS spectrograph (Petit, Donati, and Espadons Project Team, 2003), (h) ESO Very Large Telescope UV-visual echelle spectrograph (Dekker et al., 2000).

Table 4.2: Observational Properties of Engulfment Sample Stars

Name	RA deg:mm:ss	Dec deg:mm:ss	T_{eff} K	$\log g$ dex	M_* $M_\odot$	RV km s^{-1}	π mas	μ_α mas yr^{-1}	μ_δ mas yr^{-1}	G mag
HAT-P-4 A*	15:19:57.89	36:13:46.35	5903 ± 7	4.14	1.31	−1.67	3.11	−21.51	−24.25	11.1
HAT-P-4 B	15:19:59.98	36:12:18.13	5919 ± 7	4.17	1.10†	−1.94	3.08	−21.42	−24.18	11.4
HD 132563 A	14:58:21.43	44:02:34.25	6158 ± 7	4.18	1.16	–	9.41	−62.79	−67.68	8.9
HD 132563 B*	14:58:21.05	44:02:34.74	6032 ± 7	4.32	1.09	−5.98	9.47	−57.48	−70.15	9.3
HD 133131 A*	15:03:36.00	−27:50:29.81	5827 ± 7	4.50	0.86	−15.34	19.41	159.01	−139.13	8.3
HD 133131 B*	15:03:35.63	−27:50:35.36	5815 ± 7	4.48	0.86	−16.63	19.43	156.23	−133.77	8.3
ω Ser A*	15:50:17.58	02:11:46.67	4900 ± 7	2.87	1.97	−3.65	13.10	30.26	−47.59	4.9
ω Ser B	15:50:13.27	02:12:24.42	5252 ± 7	4.54	0.88†	−3.24	12.90	30.55	−48.53	10.1
HD 178911 A	19:09:04.45	34:36:04.56	5849 ± 7	4.20	1.39	−38.09	20.23	76.62	207.13	6.6
HD 178911 B*	19:09:03.17	34:36:02.61	5563 ± 7	4.39	1.03	–	24.41	57.18	195.81	7.9
16 Cyg A	19:41:48.71	50:31:27.68	5781 ± 7	4.28	1.02	−27.21	47.32	−148.03	−159.03	5.8
16 Cyg B*	19:41:51.75	50:31:00.49	5746 ± 7	4.37	0.98	−27.73	47.33	−134.48	−162.70	6.1
HD 202772 A*	21:18:47.90	−26:36:58.98	6255 ± 7	3.91	1.48	−17.71	6.14	23.25	−57.67	8.2
HD 202772 B	21:18:47.81	−26:36:58.44	6103 ± 7	4.14	1.26	–	6.33	28.91	−56.51	10.0
HAT-P-1 A	22 57 45.96	38:40:26.53	6069 ± 7	4.12	1.23†	−3.02	6.24	32.08	−42.08	9.6
HAT-P-1 B*	22:57:46.89	38:40:29.69	5966 ± 7	4.32	1.13	−2.98	6.24	32.42	−41.95	10.2
Kepler-25 A*	19:06:33.21	39:29:16.46	6214 ± 7	4.12	1.14	−7.59	4.15	−0.30	6.11	10.6
Kepler-25 B	19:06:32.52	39:29:19.10	4825 ± 11	4.47	0.80	–	4.11	0.32	6.18	13.2
WASP-94 A*	20:55:07.98	−34:08:08.73	6042 ± 11	4.16	1.36	−8.30	4.75	26.50	−44.97	10.0
WASP-94 B	20:55:09.19	−34:08:08.63	5987 ± 11	4.23	1.24	−8.45	4.72	26.19	−44.70	10.4
HD 20781*	03:20:03.37	−28:47:02.86	5232 ± 7	4.45	0.86	40.31	27.81	348.87	−66.61	7.2
HD 20782*	03:20:04.00	−28 51 15.71	5760 ± 7	4.36	0.93	39.89	27.88	349.05	−65.31	8.2
HD 40979 A*	06:04:30.08	44:15:35.15	6137 ± 7	4.36	1.23	32.47	29.43	95.07	−152.65	6.6
HD 40979 B	06:04:13.16	44:16:38.63	4896 ± 7	4.54	0.85	33.02	29.46	94.28	−153.19	8.8
KELT-2 A*	06:10:39.37	30:57:25.68	6142 ± 7	3.96	1.48	−47.22	7.43	16.73	−2.15	8.6
KELT-2 B	06:10:39.28	30:57:27.79	4847 ± 11	4.41	0.80†	–	7.29	17.86	−3.59	12.0
WASP-173 A*	23:36:40.49	−34:36:40.70	5796 ± 11	4.49	1.10	–	4.24	87.91	−8.71	11.4
WASP-173 B	23:36:40.96	−34:36:42.82	5441 ± 11	4.43	0.95†	–	4.27	87.41	−8.95	12.0
WASP-180 A*	08:13:34.14	−01:58:58.04	6316 ± 11	4.41	1.22	27.73	3.98	−13.89	−2.82	10.9
WASP-180 B	08:13:34.35	−01:59:01.70	5808 ± 14	4.53	1.07	27.99	3.84	−13.23	−2.79	11.8
Kepler-515 A*	19:21:58.64	52:03:18.98	5197 ± 11	4.52	0.80	−8.05	3.05	−23.45	−71.26	13.2
Kepler-515 B	19:21:58.42	52:03:19.08	4798 ± 11	4.52	0.71	–	3.07	−24.18	−71.90	13.8

This is a subset of a table that lists the equatorial coordinates, T_{eff}, $\log g$, M_*, *Gaia* EDR3-sourced RVs, parallaxes, proper motions, and G-magnitudes for stars in the engulfment sample. T_{eff} and $\log g$ were calculated by applying `SME` to the Keck-HIRES spectra. M_* were generated via `SpecMatch-Syn` (Petigura, 2015), except for targets marked with †, which were obtained from Mugrauer (2019) or the *NASA Exoplanet Archive*[2]. The brighter component of each binary pair is denoted as ‘A’, and the fainter component as ‘B’. The planet hosts are marked with *.

Table 4.3: Sample Architectures

Architecture	Number
Hot/Warm Jupiters	15
Hot/Warm sub-Saturns	11
Cold Jupiters	15
Cold sub-Saturns	2
Super-Earths/Sub-Neptunes	11

Table 4.4: Abundances of Engulfment Sample Stars

Name	[C/H] dex	[N/H] dex	[O/H] dex	[Na/H] dex	[Mn/H] dex	[Cr/H] dex	[Si/H] dex	[Fe/H] dex	[Mg/H] dex	[Ni/H] dex	[V/H] dex	[Ca/H] dex	[Ti/H] dex	[Al/H] dex	[Y/H] dex
HAT-P-4 A*	0.17	0.31	0.27	0.21	0.23	0.30	0.25	0.33	0.28	0.21	0.28	0.24	0.29	0.30	0.39
HAT-P-4 B	0.12	0.23	0.19	0.15	0.16	0.17	0.19	0.20	0.16	0.10	0.17	0.19	0.19	0.19	0.22
HD 132563 A	−0.11	0.18	0.07	−0.19	−0.31	−0.12	−0.10	−0.11	−0.13	−0.20	−0.14	−0.09	−0.07	−0.26	−0.10
HD 132563 B*	−0.10	−0.09	−0.02	−0.19	−0.28	−0.13	−0.11	−0.12	−0.12	−0.19	−0.15	−0.09	−0.09	−0.25	−0.10
HD 133131 A*	−0.19	−0.33	−0.17	−0.26	−0.22	−0.22	−0.21	−0.23	−0.19	−0.20	−0.28	−0.37	−0.26	−0.27	−0.32
HD 133131 B*	−0.19	−0.25	−0.15	−0.27	−0.22	−0.23	−0.22	−0.24	−0.20	−0.24	−0.29	−0.39	−0.27	−0.29	−0.33
ω Ser A*	−0.24	0.17	−0.14	0.00	0.15	0.04	−0.18	0.13	−0.07	0.14	−0.03	0.08	0.09	0.00	0.41
ω Ser B	−0.16	−0.31	−0.07	−0.21	−0.25	−0.17	−0.14	−0.18	−0.14	−0.21	−0.15	−0.13	−0.13	−0.17	−0.13
HD 178911 A	0.13	0.34	0.23	0.31	0.23	0.20	0.15	0.20	0.12	0.20	0.28	0.25	0.20	0.23	0.17
HD 178911 B*	0.18	0.24	0.17	0.28	0.26	0.21	0.22	0.21	0.18	0.24	0.21	0.21	0.20	0.22	0.14
16 Cyg A	0.07	0.05	0.13	0.10	0.08	0.08	0.08	0.09	0.08	0.10	0.11	0.10	0.09	0.12	0.03
16 Cyg B*	0.04	0.05	0.06	0.08	0.05	0.06	0.06	0.06	0.06	0.07	0.06	0.07	0.08	0.09	0.03
HD 202772 A*	0.15	0.55	0.47	0.31	0.24	0.37	0.27	0.35	0.19	0.31	0.20	0.41	0.38	0.25	0.57
HD 202772 B	0.16	0.45	−0.01	0.29	0.33	0.26	0.25	0.29	0.17	0.28	0.17	0.32	0.25	0.27	0.33
HAT-P-1 A	0.07	0.13	0.21	0.13	0.05	0.14	0.13	0.15	0.09	0.11	0.12	0.21	0.15	0.11	0.17
HAT-P-1 B*	0.11	0.15	0.17	0.12	0.14	0.16	0.16	0.16	0.14	0.16	0.14	0.18	0.15	0.15	0.24
Kepler-25 A*	−0.06	0.06	0.16	−0.07	−0.14	0.00	−0.01	0.01	−0.05	−0.09	−0.08	0.05	0.05	−0.19	−0.03
Kepler-25 B	−0.01	−0.33	−0.07	−0.03	0.01	0.04	0.03	0.03	−0.04	−0.02	0.00	0.08	0.00	0.00	−0.08
WASP-94 A*	0.21	0.39	0.32	0.34	0.37	0.34	0.28	0.34	0.28	0.37	0.22	0.35	0.34	0.26	0.44
WASP-94 B	0.24	0.35	0.34	0.37	0.43	0.30	0.30	0.32	0.29	0.35	0.20	0.38	0.32	0.34	0.35
HD 20781*	−0.06	−0.16	0.03	−0.15	−0.11	−0.05	−0.07	−0.05	−0.04	−0.10	−0.05	−0.04	−0.03	−0.05	−0.18
HD 20782*	−0.06	−0.13	−0.01	−0.17	−0.16	−0.07	−0.08	−0.06	−0.06	−0.11	−0.08	−0.05	−0.05	−0.06	−0.14
HD 40979 A*	0.17	0.37	0.29	0.29	0.29	0.27	0.25	0.27	0.21	0.25	0.24	0.29	0.23	0.18	0.23
HD 40979 B	0.24	0.13	0.19	0.31	0.27	0.22	0.25	0.27	0.19	0.26	0.23	0.30	0.24	0.25	0.05
KELT-2 A*	0.05	0.36	0.38	0.19	0.08	0.18	0.15	0.20	0.07	0.14	0.08	0.22	0.16	0.07	0.18
KELT-2 B	0.20	−0.13	0.10	0.30	0.20	0.06	0.36	0.21	0.07	0.16	0.12	0.25	0.10	0.36	-0.14
WASP-173 A*	0.15	0.28	0.25	0.26	0.30	0.24	0.18	0.23	0.16	0.22	0.17	0.27	0.26	0.21	0.22
WASP-173 B	0.07	0.18	0.03	0.16	0.19	0.16	0.15	0.17	0.14	0.18	0.15	0.16	0.19	0.17	0.14
WASP-180 A*	−0.09	0.31	0.13	−0.08	−0.02	0.08	0.06	0.10	0.00	−0.04	−0.20	0.17	0.07	−0.37	0.18
WASP-180 B	−0.06	−0.02	0.15	−0.08	−0.12	0.02	−0.02	0.01	−0.07	−0.12	−0.07	0.09	0.06	−0.09	0.06
Kepler-515 A*	−0.14	−0.22	−0.07	−0.23	−0.19	−0.17	−0.11	−0.16	−0.10	−0.18	−0.09	−0.16	−0.09	−0.16	−0.19
Kepler-515 B	0.04	−0.20	0.13	−0.16	−0.26	−0.18	−0.03	-0.20	−0.13	−0.17	−0.12	−0.09	−0.11	−0.08	−0.43
Kepler-477 A	−0.15	−0.49	0.07	−0.47	−0.53	−0.40	−0.27	−0.39	−0.34	−0.43	−0.28	−0.33	−0.26	−0.29	−0.65
Kepler-477 B*	−0.28	−0.56	−0.11	−0.50	−0.57	−0.42	−0.37	−0.44	−0.34	−0.43	−0.36	−0.35	−0.33	−0.38	−0.56
Kepler-1063 A*	0.14	0.16	0.09	0.24	0.20	0.17	0.15	0.16	0.14	0.19	0.12	0.21	0.14	0.18	0.10
Kepler-1063 B	0.20	0.18	0.20	0.28	0.28	0.24	0.22	0.23	0.19	0.25	0.21	0.27	0.23	0.27	0.22
WASP-3 A*	−0.14	0.40	0.12	−0.21	−0.24	−0.02	−0.01	−0.01	−0.09	−0.14	−0.05	0.01	0.01	−0.41	−0.09
WASP-3 C	−0.11	−0.61	−0.27	−0.02	−0.14	−0.12	0.05	−0.07	−0.13	−0.08	−0.09	0.05	−0.07	0.02	−0.08

This is a subset of a table that lists the SME-determined elemental abundances for stars in the engulfment sample. $T_{\rm eff}$ and log$g$ were calculated by applying SME to the Keck-HIRES spectra. The brighter component of each binary pair is denoted as ‘A’, and the fainter component as ‘B’. The planet hosts are marked with *.

Table 4.5: Abundance Errors of Engulfment Sample Stars

Name	SNR/pix	σ[C/H] dex	σ[N/H] dex	σ[O/H] dex	σ[Na/H] dex	σ[Mn/H] dex	σ[Cr/H] dex	σ[Si/H] dex	σ[Fe/H] dex	σ[Mg/H] dex	σ[Ni/H] dex	σ[V/H] dex	...
HAT-P-4 A*	150	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HAT-P-4 B	139	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HD 132563 A	200	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HD 132563 B*	200	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HD 133131 A*	201	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HD 133131 B*	200	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
ω Ser A*	253	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
ω Ser B	200	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HD 178911 A	202	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HD 178911 B*	253	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
16 Cyg A	205	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
16 Cyg B*	200	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HD 202772 A*	141	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HD 202772 B	141	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HAT-P-1 A	219	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HAT-P-1 B*	98	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
Kepler-25 A*	167	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
Kepler-25 B	40	0.014	0.082	0.035	0.021	0.016	0.014	0.016	0.008	0.013	0.012	0.031	...
WASP-94 A*	58	0.013	0.069	0.028	0.020	0.014	0.010	0.014	0.008	0.012	0.010	0.027	...
WASP-94 B	57	0.013	0.070	0.029	0.020	0.014	0.010	0.014	0.008	0.012	0.010	0.027	...
HD 20781*	200	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HD 20782*	202	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HD 40979 A*	253	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
HD 40979 B	141	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
KELT-2 A*	142	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
KELT-2 B	51	0.013	0.068	0.028	0.021	0.015	0.011	0.013	0.007	0.012	0.009	0.026	...
WASP-173 A*	51	0.013	0.072	0.031	0.019	0.014	0.010	0.014	0.008	0.013	0.011	0.027	...
WASP-173 B	51	0.013	0.071	0.030	0.020	0.014	0.010	0.014	0.008	0.013	0.010	0.027	...
WASP-180 A*	62	0.013	0.066	0.026	0.022	0.015	0.011	0.013	0.007	0.011	0.008	0.026	...
WASP-180 B	40	0.016	0.092	0.036	0.025	0.020	0.019	0.020	0.009	0.011	0.012	0.037	...
Kepler-515 A*	49	0.013	0.073	0.032	0.019	0.014	0.010	0.014	0.008	0.013	0.011	0.027	...
Kepler-515 B	51	0.013	0.070	0.030	0.020	0.014	0.010	0.014	0.008	0.012	0.010	0.027	...
Kepler-477 A	40	0.017	0.097	0.037	0.026	0.021	0.019	0.020	0.010	0.012	0.013	0.038	...
Kepler-477 B*	40	0.017	0.097	0.037	0.026	0.021	0.019	0.020	0.010	0.012	0.013	0.038	...
Kepler-1063 A*	51	0.013	0.073	0.032	0.019	0.014	0.010	0.014	0.008	0.013	0.011	0.027	...
Kepler-1063 B	51	0.013	0.073	0.032	0.019	0.014	0.010	0.014	0.008	0.013	0.011	0.027	...
WASP-3 A*	170	0.011	0.042	0.019	0.014	0.010	0.007	0.009	0.006	0.008	0.008	0.017	...
WASP-3 C	40	0.014	0.083	0.035	0.022	0.017	0.014	0.017	0.008	0.012	0.012	0.032	...
WASP-160 A	51	0.013	0.071	0.031	0.020	0.014	0.010	0.014	0.008	0.013	0.011	0.027	...
WASP-160 B*	51	0.013	0.071	0.030	0.020	0.014	0.010	0.014	0.008	0.013	0.010	0.027	...

This is a subset of a table that lists the SME-determined elemental abundance errors for stars in the engulfment sample. The brighter component of each binary pair is denoted as ‘A’, and the fainter component as ‘B’. The planet hosts are marked with *.

Table 4.6: Abundance Errors of Engulfment Sample Stars

Name	$EW_{\rm Li}$ mÅ	A(Li) dex	ΔA(Li) dex	Expected ΔA(Li) dex
HD 23596 A*	73.18 ± 2.55	2.68 ± 0.03	2.83 ± 0.08	~1.5
HD 23596 B	17.53 ± 2.69	−0.14 ± 0.07	–	–
WASP-3 A*	18.27 ± 1.06	2.26 ± 0.03	> 2.83	~1.3
WASP-3 C	2.16 ± 4.06	< −0.57	–	–
KELT-4 A*	24.64 ± 1.04	2.39 ± 0.03	2.68 ± 0.18	~2.0
KELT-4 B	2.03 ± 1.01	−0.29 ± 0.18	–	–
Kepler-410 A*	13.98 ± 2.08	2.12 ± 0.07	> 2.50	~2.0
Kepler-410 B	0.00 ± 1.36	< −0.38	–	–
Kepler-25 A*	23.69 ± 1.07	2.31 ± 0.03	> 2.32	~2.0
Kepler-25 B	3.24 ± 3.58	< −0.02	–	–
HD 40979 A*	79.33 ± 0.61	2.86 ± 0.02	2.24 ± 0.06	~2.0
HD 40979 B	11.30 ± 1.12	0.62 ± 0.05	–	–
Kepler-104 A*	17.07 ± 0.82	1.84 ± 0.03	> 1.87	~1.5
Kepler-104 B	0.00 ± 0.70	< −0.03	–	–
WASP-70 A*	3.29 ± 2.85	1.03 ± 0.28	> 1.71	~1.5
WASP-70 B	1.20 ± 2.65	< −0.68	–	–
HAT-P-41 A*	1.44 ± 0.76	1.22 ± 0.19	> 1.70 ± 0.20	~2.0
HAT-P-41 B	0.00 ± 2.49	< −0.48	–	–
WASP-127 A*	27.13 ± 1.56	2.03 ± 0.03	> 1.60	~0.3
WASP-127 B	0.00 ± 2.23	< 0.43	–	–
WASP-173 A*	0.00 ± 2.70	< 1.53	< 1.37	~0.5
WASP-173 B	0.94 ± 2.67	< 0.16	–	–
Kepler-99 B*	0.00 ± 2.78	< 0.15	< 1.07	~0.3
Kepler-99 A	0.48 ± 1.18	< −0.92	–	–
†WASP-94 A*	9.73 ± 3.33	1.75 ± 0.13	> 1.03	< 0.1
†WASP-94 B	1.07 ± 3.11	< 0.72	–	—

This is a subset of a table that lists the $EW_{\rm Li}$ and A(Li) measurements for stars in the engulfment sample, ranked by their ΔA(Li). In cases where the Li EW is smaller than the associated uncertainty, A(Li) is reported as an upper limit. The brighter component of each binary pair is denoted as 'A', and the fainter component as 'B'. The planet hosts are marked with *, and twin systems are marked with †. We provide the expected ΔA(Li) between companions as a function of their differing $T_{\rm eff}$ not considering engulfment (Randich and Magrini, 2021) in the last column.

Table 4.7: Engulfment Model Parameterss

Binary System	sep AU	M $M_\oplus$	$\sigma_{\rm jit}$	shift dex	flat model shift dex	$\Delta\ln(Z)$
WASP-94*†	3200 ± 17	2.95 ± 1.55	0.03 ± 0.01	−0.03 ± 0.01	0.00 ± 0.01	1.96
HAT-P-4*†	30000 ± 140	5.60 ± 1.64	0.03 ± 0.01	0.05 ± 0.01	0.10 ± 0.01	1.82
HD 133131*†	380 ± 0.62	0.50 ± 0.23	0.00 ± 0.00	0.00 ± 0.01	0.01 ± 0.00	1.04
WASP-64†*	8700 ± 34	1.83 ± 0.89	0.00 ± 0.00	0.00 ± 0.01	0.01 ± 0.00	0.66
HD 106515*†	230 ± 0.25	0.97 ± 0.84	0.00 ± 0.00	−0.02 ± 0.01	−0.01 ± 0.00	0.01
HD 132563*†	430 ± 0.52	0.42 ± 0.29	0.02 ± 0.01	−0.02 ± 0.01	−0.01 ± 0.00	−0.37
HAT-P-1*†	1800 ± 4.1	0.65 ± 0.61	0.03 ± 0.01	0.02 ± 0.01	0.03 ± 0.01	−0.76
XO-2*†	4700 ± 11	9.46 ± 4.97	0.04 ± 0.01	0.04 ± 0.01	0.06 ± 0.01	−0.83
16 Cyg*†	840 ± 0.28	0.29 ± 0.22	0.00 ± 0.00	0.02 ± 0.00	0.03 ± 0.00	−1.50
HD 202772*†	210 ± 1.7	1.13 ± 1.05	0.14 ± 0.02	0.03 ± 0.03	0.08 ± 0.03	−2.20
HD 99491-92	510 ± 0.30	11.73 ± 2.98	0.02 ± 0.01	−0.08 ± 0.01	−0.04 ± 0.01	5.21
Kepler-477*	560 ± 5.9	3.06 ± 0.85	0.03 ± 0.01	−0.10 ± 0.02	−0.05 ± 0.01	4.53
Kepler-515*	650 ± 2.2	8.62 ± 2.92	0.08 ± 0.01	−0.08 ± 0.03	−0.02 ± 0.02	3.98
WASP-180	1200 ± 6.2	4.93 ± 2.23	0.09 ± 0.02	−0.07 ± 0.02	−0.03 ± 0.02	2.31
Kepler-25	2000 ± 5.5	8.36 ± 5.04	0.10 ± 0.02	−0.03 ± 0.03	0.01 ± 0.02	1.46
WASP-160*	8300 ± 28	4.36 ± 2.70	0.01 ± 0.01	0.02 ± 0.01	0.04 ± 0.01	0.98
WASP-173	1400 ± 6.9	6.18 ± 3.20	0.02 ± 0.01	−0.10 ± 0.01	−0.07 ± 0.01	−0.12
K2-27*	8100 ± 31	0.86 ± 0.70	0.02 ± 0.01	−0.02 ± 0.01	−0.02 ± 0.00	−0.25
HD 178911*	650 ± 0.39	1.62 ± 1.81	0.04 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	−0.27
HD 40979	6500 ± 4.5	2.45 ± 3.10	0.07 ± 0.01	−0.03 ± 0.02	−0.02 ± 0.01	−0.40
WASP-127*	6500 ± 20	0.62 ± 0.42	0.03 ± 0.01	0.03 ± 0.01	0.04 ± 0.01	−0.55
Kepler-99*	3100 ± 8.5	1.16 ± 1.51	0.04 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	−0.55
HD 80606-07*	1400 ± 1.6	0.66 ± 0.71	0.01 ± 0.00	−0.01 ± 0.00	−0.01 ± 0.00	−0.64
KELT-2	320 ± 0.83	1.33 ± 1.36	0.20 ± 0.03	0.00 ± 0.03	0.02 ± 0.04	−0.75
Kepler-104*	6900 ± 27	0.68 ± 0.45	0.13 ± 0.02	−0.06 ± 0.02	−0.07 ± 0.02	−1.44
HD 20781-82*	9100 ± 7.9	0.25 ± 0.23	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	−1.70
Kepler-1063	580 ± 27	1.50 ± 1.12	0.00 ± 0.00	0.06 ± 0.01	0.07 ± 0.00	−1.79
KELT-4	340 ± 1.2	3.55 ± 4.74	0.31 ± 0.04	−0.03 ± 0.03	−0.09 ± 0.06	−2.05
ω Ser*	5700 ± 35	48.09 ± 18.58	0.20 ± 0.03	0.03 ± 0.03	0.20 ± 0.04	−4.07

This table lists the binary separation, modeled amount of engulfed planetary mass, fitted jitter term $\sigma_{\rm jit}$, engulfment model shift, flat model shift, and difference in engulfment model and flat model Bayesian evidence $\Delta\ln(Z)$ for each of the remaining 29 binary pairs in the engulfment sample. We separate stellar twin systems ($\Delta T_{\rm eff}$ < 200 K, marked with †) from the non-twin systems. For each pair, we chose either the planet host or the non-planet host to be the engulfing star based on which order yielded the largest $\Delta\ln(Z)$. Pairs where the planet host was assumed to be the engulfing star are marked with *. Within the twin and non-twin systems, binary pairs are sorted by $\Delta\ln(Z)$.

Chapter 5

PLANET ENGULFMENT SIGNATURES IN TWIN STARS

Behmard, Aida, Jason Sevilla, and Jim Fuller (Feb. 2023). "Planet engulfment signatures in twin stars". In: *Monthly Notices of the Royal Astronomical Society* 518.4, pp. 5465–5474. DOI: `10.1093/mnras/stac3435`. arXiv: `2210.11679 [astro-ph.SR]`.

Abstract

Planet engulfment can be inferred from enhancement of refractory elements in the photosphere of the engulfing star following accretion of rocky planetary material. Such refractory enrichments are subject to stellar interior mixing processes, namely thermohaline mixing induced by an inverse mean-molecular-weight gradient between the convective envelope and radiative core. Using `MESA` stellar models, we quantified the strength and duration of engulfment signatures following planet engulfment. We found that thermohaline mixing dominates during the first ∼5−45 Myr post-engulfment, weakening signatures by a factor of ∼2 before giving way to depletion via gravitational settling on longer timescales. Solar metallicity stars in the 0.5–1.2 $M_\odot$ mass range have observable signature timescales of ∼1 Myr−8 Gyr, depending on the engulfing star mass and amount of material engulfed. Early type stars exhibit larger initial refractory enhancements but more rapid depletion. Solar-like stars (M = 0.9−1.1 $M_\odot$) maintain observable signatures (>0.05 dex) over timescales of ∼20 Myr−1.7 Gyr for nominal 10 $M_\oplus$ engulfment events, with longer-lived signatures occurring for low-metallicity and/or hotter stars (1 $M_\odot$, ∼2−3 Gyr). Engulfment events occurring well after the zero-age main sequence produce larger signals due to suppression of thermohaline mixing by gravitational settling of helium (1 $M_\odot$, ∼1.5 Gyr). These results indicate that it may be difficult to observe engulfment signatures in solar-like stars that are several Gyr old.

5.1 Introduction

The formation and evolution of planetary systems can alter the primordial elemental abundances of planet host stars. For example, planet engulfment can produce refractory element enhancements within the engulfing star convective region due to ingestion of rocky planetary material (e.g., Oh et al. 2018). However, such refractory enhancements may be weakened by internal mixing processes over time. While gravitational settling alone is not sufficiently rapid to diminish these enhancements (e.g., Pinsonneault, DePoy, and Coffee 2001), thermohaline mixing may more effectively deplete overlying refractory material by allowing it to sink below the convective zone via "metallic fingers" that stretch towards the deep stellar interior (e.g., Ulrich 1972; Kippenhahn, Ruschenplatt, and Thomas 1980; Vauclair 2004; Théado and Vauclair 2012; Bauer and Bildsten 2019). As a form of double-diffusive mixing driven by heat diffusion, thermohaline mixing acts in the presence of a mean-molecular-weight (μ) gradient, and is theorized to be particularly effective at removing accreted refractory material from the thin convective zones of hot stars due to more rapid cooling at the bottom of the convective envelope.

Vauclair (2004) was the first to consider thermohaline mixing in stellar interiors following accretion of planetary material. They analytically determined that thermohaline instabilities are capable of depleting accreted rocky matter out of convective envelopes on timescales of $\sim$1000 years, leaving only a small μ-gradient at the end of the mixing process. Later work by Garaud (2011) presented numerical and semi-analytical estimations of thermohaline mixing using the coefficient derived by Traxler, Garaud, and Stellmach (2011) from empirical fitting of 3D numerical simulation results, and found that the refractory depletion timescale following engulfment of a 1 M_J planet depends strongly on increasing stellar mass, with enhancements dropping to 10% within 600 Myr for a 1.3 $M_\odot$ star, 60 Myr for a 1.4 $M_\odot$ star, etc.

More recently, Théado and Vauclair (2012) modeled engulfment with planet masses ranging from 1 $M_\oplus$ to 1.5 M_J using the Toulouse-Geneva Evolution Code (Richard, Théado, and Vauclair, 2004; Hui-Bon-Hoa, 2008) considering thermohaline mixing and atomic diffusion. They found that the μ-gradient is softened on timescales of a few to tens of Myr depending on the accretion scenario, and noted that their depletion timescale is much larger than those of previous studies because they correctly assumed the characteristic mixing length to be the entire mixed region rather than just the "metallic finger" dimensions. Théado and Vauclair (2012) showed how different mixing prescriptions could lead to different results, especially in the case of a small composition gradient (near the minimum for thermohaline instability to operate), as often occurs in models of planet engulfment. This highlights the need for stellar models that include all relevant interior processes for accurately constraining depletion timescales. Such timescales are important for investigations of engulfment because they determine how long engulfment signatures remain observable, which is necessary for placing engulfment events within planetary system histories.

Stellar abundance patterns suggestive of planet engulfment have been observed in several systems, indicating that engulfment may be more common than predictions from theory suggest (Galactic occurrence rates of 0.1–1 yr^{-1}, Metzger, Giannios, and Spiegel 2012). More specifically, there are several individual binary systems reported in the literature with significant (>0.05 dex[1]) refractory differences between the two stars (Ramírez, Meléndez, et al., 2011; Mack et al., 2014; Tucci Maia, Meléndez, and Ramírez, 2014; Teske, Ghezzi, et al., 2015; Ramírez, Khanal, Aleo, et al., 2015; Biazzo et al., 2015; Saffe, Flores, et al., 2016; Teske, Khanal, and Ramírez, 2016; Adibekyan et al., 2016; Saffe, Jofré, et al., 2017; Tucci Maia, Meléndez, Lorenzo-Oliveira, et al., 2019; Ramírez, Khanal, Lichon, et al., 2019; Nagar, Spina, and Karakas, 2020; Galarza et al., 2021; Jofré et al., 2021). Such abundances differences between stellar siblings suggest post-birth chemodynamical processes such as engulfment because bound stellar companions are born from the same natal gas cloud and thus assumed to be chemically homogeneous at birth to within ~0.05 dex (e.g., De Silva, Freeman, and Bland-Hawthorn 2009). Larger binary samples exhibit abundance differences as well. For example, Hawkins et al. (2020) assessed 25 comoving, wide binaries and found that 5 systems exhibit Δ[Fe/H] $\sim$ 0.10 dex.

To determine how planet engulfment signatures are affected by stellar mixing processes, namely thermohaline instabilities, we ran tests with the `MESA` (Modules for Experiments in Stellar Astrophysics) stellar evolution code (Paxton, Bildsten, et al., 2011; Paxton, Cantiello, et al., 2013; Paxton, Marchant, et al., 2015; Paxton, Schwab, et al., 2018; Paxton, Smolec, et al., 2019). Our `MESA` stellar models and implementation of non-standard mixing processes such as thermohaline instabilities are outlined in Section 5.2. The results of our `MESA` models considering different engulfment conditions are presented in Section 5.3. We discuss the implications of these results for planet engulfment signatures in Section 5.4, and provide a summary in Section 5.5.

5.2 Stellar Models

We computed our stellar models using the open-source 1D stellar evolution code `MESA`. These models are non-rotating with masses of 0.5–1.2 $M_\odot$ and solar metal-

[1]In this work, we adopt the standard "bracket" chemical abundance notation [X/H] = A(X) - $A(\mathrm{X})_\odot$, where $A(\mathrm{X}) = \log(n_\mathrm{X}/n_H) + 12$ and n_X is the number density of species X in the star's photosphere.

licities of $Z = 0.017$. We did not include more massive stars, where gravitational settling and radiative levitation can potentially produce huge changes in surface composition, but are complicated by poorly understood additional mixing processes not included in our models (e.g., Eddington-Sweet circulation and wave mixing).

Our modeling procedure follows that of Sevilla, Behmard, and Fuller (2022). In brief, the models were run in three stages, where the first stage evolved the stars up to the zero-age main sequence (ZAMS), the second stage simulated planet engulfment of a 1, 10, or 50 $M_\oplus$ planet via accretion of bulk-Earth composition material, and the third stage evolved the stars up to the end of their main sequence (MS) lifetimes. We then computed the refractory species mass fraction remaining in their convective zones as a function of stellar age to assess the timescales over which engulfment signatures remained observable. Throughout these `MESA` runs we applied relevant mixing processes, namely convective overshoot, elemental diffusion, radiative levitation, thermohaline instabilities, and additional mixing required to reproduce observations of surface lithium abundances. These processes are discussed in more detail below.

Input Physics

We applied convective overshoot via the prescription taken from Herwig (2000), with `MESA` implementation detailed in Paxton, Bildsten, et al. (2011):

$$D_{\mathrm{ov}} = D_{\mathrm{conv,\,0}} \exp\left(-\frac{2z}{f\lambda_{P,0}}\right), \tag{5.1}$$

where $D_{\mathrm{conv,\,0}}$ is the diffusion coefficient at a location f_0 scale heights inside the convective zone, $\lambda_{P,0}$ is the pressure scale height at that location, z is the distance away from this location, and f is an adjustable parameter that sets the characteristic size of the region undergoing convective overshoot. We set f and f_0 to 0.02 and 0.005, respectively.

We implemented elemental diffusion via the prescription detailed in Paxton, Marchant, et al. (2015) and Paxton, Schwab, et al. (2018) that solves the Burgers' equations in a similar fashion to Thoul, Bahcall, and Loeb (1994). In `MESA`, diffusion is carried out for a set of chemical species organized in "classes", where a single class may contain several isotopes, and a chosen representative isotope is used to solve the diffusion equations and set the diffusion velocities for all species in that class. In our case, we created individual classes for each of the 13 most abundant species included in our bulk-Earth accretion that simulates engulfment, so each species is its own representative isotope.

We also applied radiative levitation according to Hu et al. (2011), implemented in `MESA` as described in Paxton, Marchant, et al. (2015). The Hu et al. (2011) prescription is based on that of Thoul, Bahcall, and Loeb (1994), with a few modifications. These include an additional force term in the basic diffusion equations that describes the radiative acceleration on a chemical species as a function of its average ion charge. We only applied radiative levitation to our stellar models with masses of 0.7 $M_\odot$ and above. This is justified because radiative levitation dominates atomic diffusion for stars with temperatures above 6000 K (e.g., Richer et al. 1998; Deal et al. 2020; Campilho, Deal, and Bossini 2022), and is not expected to operate efficiently in low mass stars.

Thermohaline mixing acts in the presence of a mean molecular weight (μ) inversion in regions that satisfy the Ledoux criterion for convective stability, such that

$$\nabla_T - \nabla_{\mathrm{ad}} \leq B \leq 0. \tag{5.2}$$

Here, B is the Ledoux term that accounts for composition gradients (e.g., Brassard et al. 1991), and $\nabla_T - \nabla_{\mathrm{ad}}$ is the difference between the actual temperature gradient and the adiabatic temperature gradient, also referred to as the superadiabaticity. We included thermohaline mixing in our `MESA` models according to the prescription of Brown, Garaud, and Stellmach (2013) derived via 3D numerical simulations and linear stability analyses. This prescription is more accurate than previous thermohaline implementations, such as that of Kippenhahn, Ruschenplatt, and Thomas (1980) which overestimates mixing efficiency in certain fingering convection configurations (Prat, Lignières, and Lagarde, 2015). It also corrects inconsistencies in the models of Denissenkov (2010) and Traxler, Garaud, and Stellmach (2011).

There are other mixing processes at play that we did not explicitly model because their effects are poorly understood, e.g., rotationally induced mixing. To account for this, we included a `min_D_mix = 700` command in our `MESA` inlist. This amount of extra mixing allowed us to reproduce the observed lithium abundances for $\sim$1 $M_\odot$ stars from Sestito and Randich (2005). For more details on the comparison with lithium observations, see Sevilla, Behmard, and Fuller (2022).

Bulk-Earth Accretion

To simulate planet engulfment, we began accreting bulk-Earth composition material once the `MESA` stellar models reached ZAMS, consisting of the top 13 most abundant elements (McDonough, 2003). For a complete analysis of the lithium enrichment

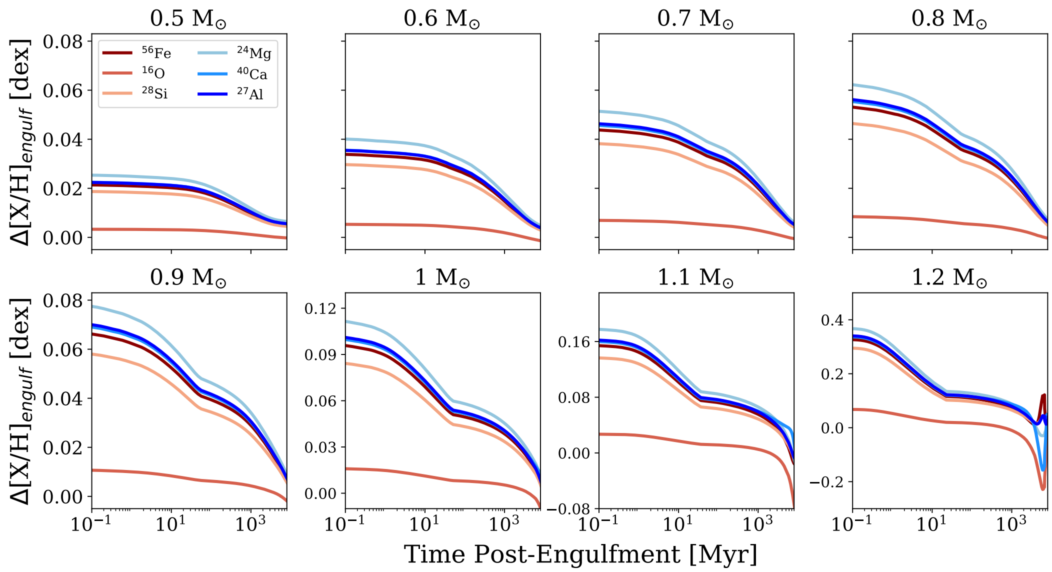


Figure 5.1: Post-engulfment abundances for the top six most abundant bulk-Earth elements over time for our complete stellar mass model range (0.5−1.2 $M_\odot$) following near-instantaneous accretion of a 10 $M_\oplus$ bulk-Earth composition planet. The abundances of comparison `MESA` models with no accretion were subtracted off. Each element is represented by its most abundant isotope because `MESA` chemical networks are written in terms of isotopic species.

evolution, see Sevilla, Behmard, and Fuller (2022). Because `MESA` chemical networks are written in terms of isotopic species, we accreted the most abundant isotope of each element included in our bulk-Earth accretion scheme.

`MESA` does not allow for instantaneous engulfment of a planetary mass body. Instead, planet engulfment was simulated through rapid accretion of bulk-Earth composition material. We chose accretion rates that ensured the total planetary mass was accreted within 10,000 years for all combinations of 1, 10, and 50 $M_\oplus$ engulfed masses and host stars of 0.5−1.2 $M_\odot$. An accretion period of 10,000 years is appropriately short for simulating engulfment events, and should not affect our results because our observed refractory depletion occurs on timescales of >1 Myr (Figure 5.1).

We also tested two alternative accretion prescriptions that mimic late heavy bombardment (LHB) and disk accretion scenarios. LHB-like accretion begins when the star reaches an age of 500 Myr and lasts for 300 Myr (e.g., Bottke and Norman 2017), while disk accretion begins 10 Myr after the star is born and lasts for 3 Myr, the nominal disk lifetime for MS stars (e.g., Hartmann, Herczeg, and Calvet 2016). We tested accretion of 1 and 10 $M_\oplus$ of planetary material for both scenarios. After

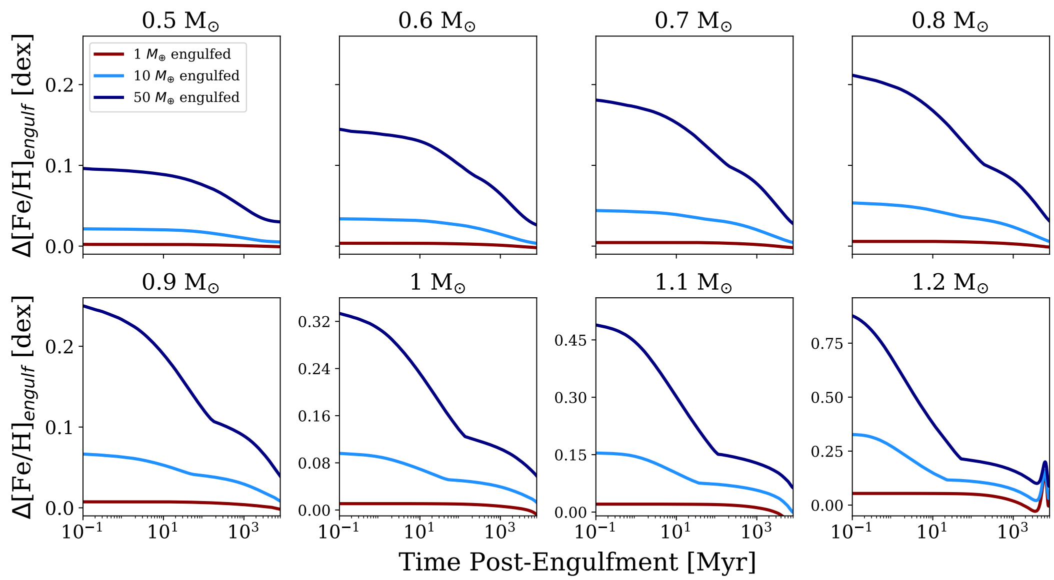


Figure 5.2: Post-engulfment [Fe/H] abundances over time for our complete stellar mass model range (0.5−1.2 $M_\odot$) following near-instantaneous accretion of a 1, 10, or 50 $M_\oplus$ bulk-Earth composition planet, marked in red, blue, and navy, respectively. The abundances of comparison `MESA` models with no accretion were subtracted off.

engulfment according to near-instantaneous, LHB, or disk accretion scenarios, we evolved the stellar models to the ends of their MS lifetimes. We utilized the default definition for the end of the MS, which is when the core hydrogen mass fraction drops below 10^{-10}.

Near-Instantaneous Engulfment

We first examined how refractory enrichment resulting from engulfment changes as a function of engulfed planet mass and stellar type. This was carried out by running `MESA` models that included accretion of 1, 10, or 50 $M_\oplus$ of planetary material with bulk-Earth compositions, and models with no accretion as comparison points. We took the differential abundances between the engulfment and non-engulfment models as our potential engulfment signatures. These differential measurements mimic the signatures found in observations because refractory enrichments from engulfment are detected with respect to pristine, chemically homogeneous stars, e.g., a bound companion or fellow cluster member.

As shown in Figure 5.1, for low-mass stars (0.5−0.7 $M_\odot$), engulfment of a 10 $M_\oplus$ planet does not produce observable enrichment at >0.05 dex levels because the accreted material is heavily diluted within the deep stellar convective envelope.

The enrichment is more pronounced for more massive stars with thinner convective envelopes; for solar-like stars (0.8−1.2 $M_\odot$), engulfing 10 $M_\oplus$ of planetary material initially produces enrichment at levels of ∼0.06−0.33 dex. Because the 0.8−0.9 $M_\odot$ stars still have comparatively deep envelopes, the initial enrichment is not well above 0.05 dex to begin with, and drops below this after by ∼20 Myr. However the higher mass stars in this range (1−1.2 $M_\odot$) maintain enrichment at >0.05 dex for ∼90 Myr−2 Gyr.

Figure 5.1 shows that for models with higher mass ($M \gtrsim 1\,M_\odot$), the refractory abundances of a star that engulfed a planet are predicted to eventually decrease *below* those of a star that did not engulf a planet within 8 Gyr post-engulfment. The strength of this counter-intuitive effect increases with stellar mass. Inspection of our 1.2 $M_\odot$ models with engulfment reveals that they are typically ∼200 K cooler and ∼1.5% larger in radius than their reference models. Surprisingly, the models with engulfment have thinner surface convection zones, leading to faster gravitational settling of heavy elements and hence lower refractory abundances long after the engulfment event. The exact cause of this structural difference is not clear and should be examined in future work. There is an exception for certain isotopic species that are subject to radiative levitation (e.g, ^{56}Fe, ^{27}Al, ^{28}Si), whose abundances increase at ∼5 Gyr in the 1.2 $M_\odot$ model (Figures 5.1 and 5.2).

Apart from this case, we found little difference between the chemical species in our models; the relative abundance of each decreases on nearly identical timescales. For thermohaline mixing, this is expected as it affects all elements equally. In principle, gravitational settling causes heavier elements to sink below the convective zone faster, but the difference is not discernible in our results. Note that the enrichment signature provided by oxygen is much weaker than those of other elements due to the large amount of oxygen already present in the star.

Dependence on Planet Mass

We also examined near-instantaneous engulfment events of 1 and 50 $M_\oplus$ of planetary material (Figure 5.2). For 1 $M_\oplus$ engulfment, refractory depletion behavior across different stellar mass regimes is similar to that of 10 $M_\oplus$ engulfment, but the initial enrichment levels are lower. Consequently, stars with masses in the range 0.5−1.1 $M_\odot$ do not exhibit enrichment above ∼0.05 dex immediately post-engulfment, rendering their engulfment signatures undetectable. The 1.2 $M_\odot$ model begins with ∼0.05 dex enrichment and sustains this level of enrichment for ∼100 Myr.

For 50 $M_{\oplus}$ engulfment, the initial enrichment is comparatively high as expected from accretion of more planetary material. Stars in the 0.5−0.7 $M_{\odot}$ mass regime begin with enrichment at 0.10−0.19 dex levels, and maintain enrichment above ~0.05 dex for ~900 Myr−3 Gyr. Solar-like stars (0.8−1.2 $M_{\odot}$) begin with enrichment at ~0.25−1.0 dex, and maintain ~0.05 dex enrichment on timescales of ~4−8 Gyr.

LHB and Disk Accretion

The LHB is considered to be the last dramatic dynamical event in the history of the Solar System, when giant planet migration disrupted the orbits of inward-lying planetesimals and perhaps shepherded some into the Sun. We modeled LHB-like events to compare how the resulting engulfment signatures may differ from those of engulfment at ZAMS. To mimic LHB-like accretion, we ran a 1 $M_{\odot}$ model undergoing gradual accretion of 1 $M_{\oplus}$ of bulk-Earth composition planetesimals from a stellar age of 500−800 Myr, thus beginning ~470 Myr after ZAMS. We found that engulfment signatures from an LHB-like event are almost completely depleted post-engulfment because refractory material is continuously depleted over the longer accretion period. The engulfment signature after LHB-like accretion is complete is shown in Figure 5.3.

We were also interested in modeling a protoplanetary disk accretion scenario, which by definition occurs while the disk is still present. To mimic these conditions, we began accretion once the star reached an age of 10 Myr, for a duration of 3 Myr as mentioned in the section outlining our bulk Earth accretion prescription. At 10 Myr, the star is still in the pre-MS phase, ~20 Myr away from reaching ZAMS and almost fully convective. Thus, any accreted material will become heavily diluted within a large volume of the stellar interior. We tested accretion of 1 and 10 $M_{\oplus}$, and found that the resulting engulfment signatures are slightly stronger than those resulting from LHB scenarios, but weaker compared to those of near-instantaneous engulfment scenarios because the refractory material is comparatively diluted (Figure 5.3).

Late-Stage Engulfment

Finally, we tested near-instantaneous engulfment occurring at 300 Myr, 1 Gyr, or 3 Gyr post-ZAMS rather than exactly at ZAMS to examine refractory enrichment evolution from late-stage accretion. For these cases we considered engulfment of a 10 $M_{\oplus}$ mass by a 1 $M_{\odot}$ star. Notably, later engulfment results in systematically larger surface refractory enhancements; for the 3 Gyr model, the enhancement is

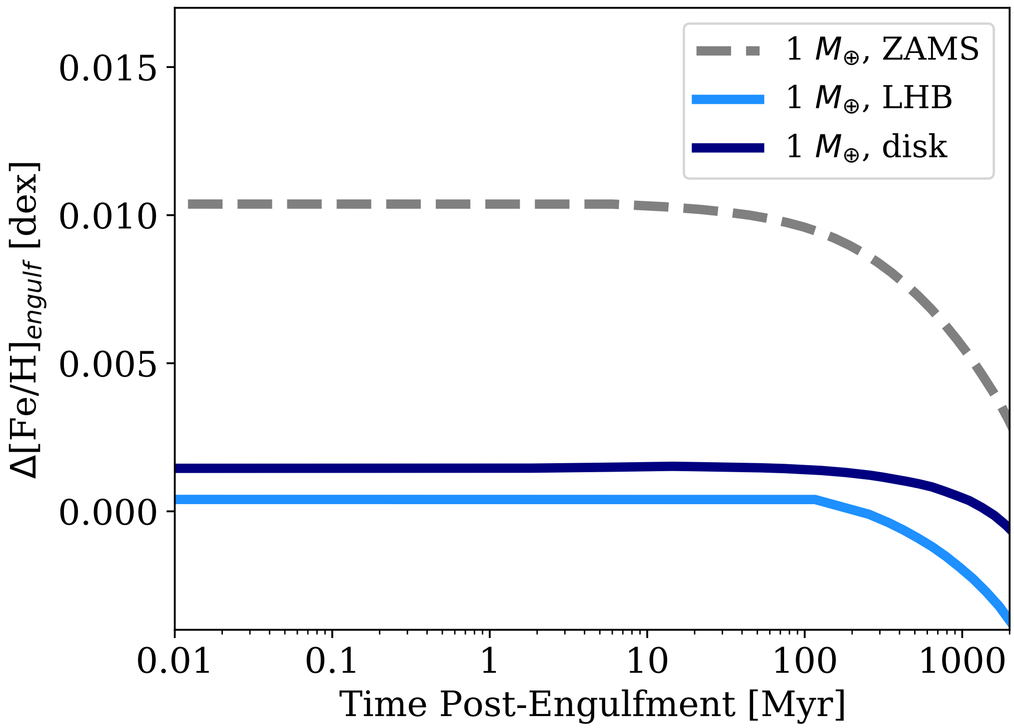


Figure 5.3: Post-engulfment [Fe/H] abundances over time for a 1 $M_\odot$ `MESA` model that underwent accretion of a 1 $M_\oplus$ planet under different conditions: the standard post-ZAMS near-instantaneous engulfment scenario (gray, dashed), an LHB-like scenario (blue), and a disk accretion-like scenario (navy). The abundances of comparison `MESA` models with no accretion were subtracted off.

$\sim$1.5$-$2 times that of the ZAMS accretion model at $\gtrsim$10 Myr post-engulfment times (Figure 5.4). This effect is also shown in the trends of different engulfed masses; we tested engulfment of a 1, 10, or 50 $M_\oplus$ bulk-Earth planetary companion by a 1 $M_\odot$ star at 1 Gyr post-ZAMS (Figure 5.5), and found that the 1 Gyr post-ZAMS cases all exhibit slower depletion of surface refractory enhancement compared to ZAMS engulfment (Figure 5.2).

The primary reason for the larger signal in late-stage engulfment is a smaller amount of thermohaline mixing, which results in less depletion after engulfment as shown in Figures 5.4 and 5.5. The weakened thermohaline mixing arises due to its interaction with gravitational settling. As stars age, gravitational settling slowly creates a stabilizing μ-gradient as helium settles downwards. This inhibits thermohaline mixing when planetary material is engulfed, thus weakening refractory depletion and leading to longer-lived engulfment signatures. This effect is discussed further in Théado and Vauclair (2012) and Sevilla, Behmard, and Fuller (2022).

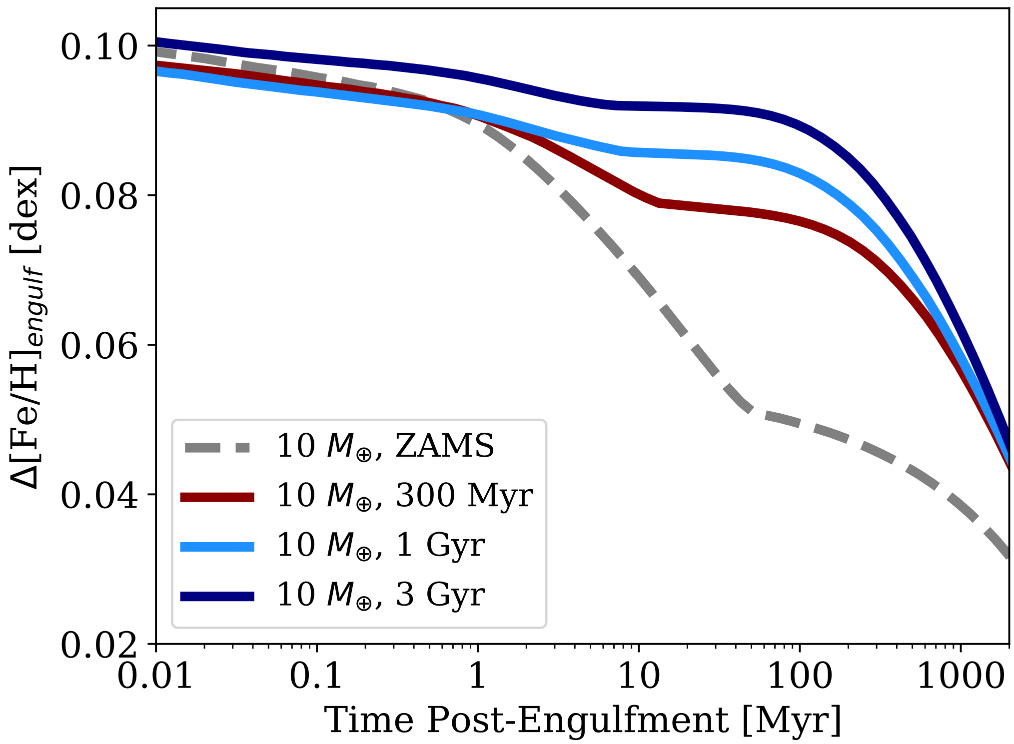


Figure 5.4: Post-engulfment [Fe/H] abundances over time for a 1 $M_\odot$ `MESA` model that underwent accretion of a 10 $M_\oplus$ planet at different times: ZAMS (gray, dashed), 300 Myr post-ZAMS (red), 1 Gyr post-ZAMS (blue), and 3 Gyr post-ZAMS (navy). The abundances of comparison `MESA` models with no accretion were subtracted off.

To further illustrate how this manifests in our `MESA` models, we plotted [Fe/H], [He/H], and the compositional component of the Brunt-Väisälä frequency ($N^2_{\rm comp}$) as a function of mass coordinate for the 1 $M_\odot$ model with 10 $M_\oplus$ engulfment considering both ZAMS and 1 Gyr post-ZAMS engulfment (Figure 5.6). The [Fe/H] panels of Figure 5.6 showcase that some gravitational settling has already occurred in the 1 Gyr post-ZAMS model compared to the ZAMS model, increasing the value of [Fe/H] at larger depths. Similarly, helium settling creates a stabilizing composition gradient beneath the convection zone, as can be seen by the increasing value of [He/H] and [Fe/H] with depth below the convective zone in the post-engulfment models (though note that nuclear burning also increases [He/H] near the core).

We further examined this effect on the μ-gradient by plotting $N^2_{\rm comp}$, a measure of stability to convection in a fluid medium. A positive $N^2_{\rm comp}$ indicates a stable μ-gradient, while a negative value indicates that thermohaline mixing can occur.

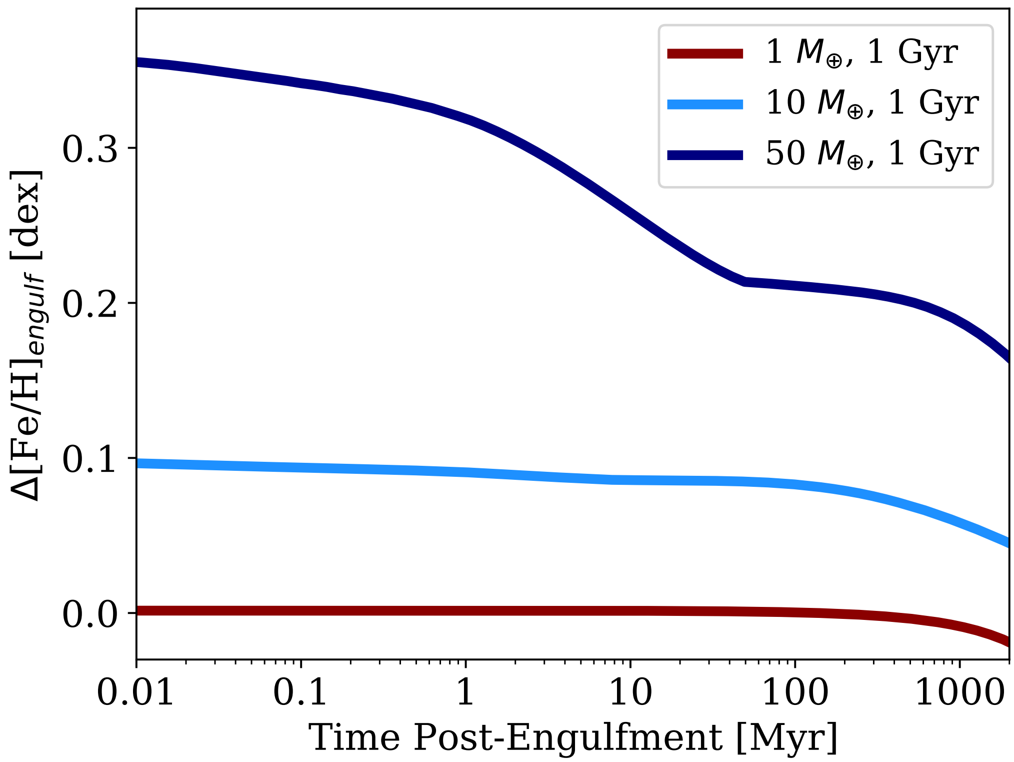


Figure 5.5: Post-engulfment [Fe/H] abundances over time for a 1 $M_\odot$ `MESA` model that underwent accretion of a 1 (red), 10 (blue), or 50 (navy) $M_\oplus$ planet at 1 Gyr post-ZAMS. The abundances of comparison `MESA` models with no accretion were subtracted off.

In both models, the bottom panels of Figure 5.6 show that $N^2_{\rm comp}$ is negative just below the convective zone of the post-engulfment model. However, at larger depths, $N^2_{\rm comp}$ has lower values for engulfment at ZAMS compared to engulfment at 1 Gyr post-ZAMS, which indicates that the 1 Gyr post-ZAMS μ-gradient is more stable to begin with. Thermohaline mixing cannot penetrate as deeply into the interior in the 1 Gyr post-ZAMS model, which thus experiences less refractory depletion via thermohaline mixing over time.

Engulfing Star Metallicity

While we defaulted to solar metallicities (Z = 0.017) in our `MESA` models, we also examined how engulfing star metallicities affect engulfment signatures. We ran 1 $M_\odot$ models of Z = 0.012 and 0.022 with accretion of 10 $M_\oplus$ planets, and found that engulfment signature strength is strongly inversely correlated with host star metallicity (Figure 5.7). This is due to two effects: lower metallicity stars have fewer metals leading to stronger relative refractory enrichments, and lower

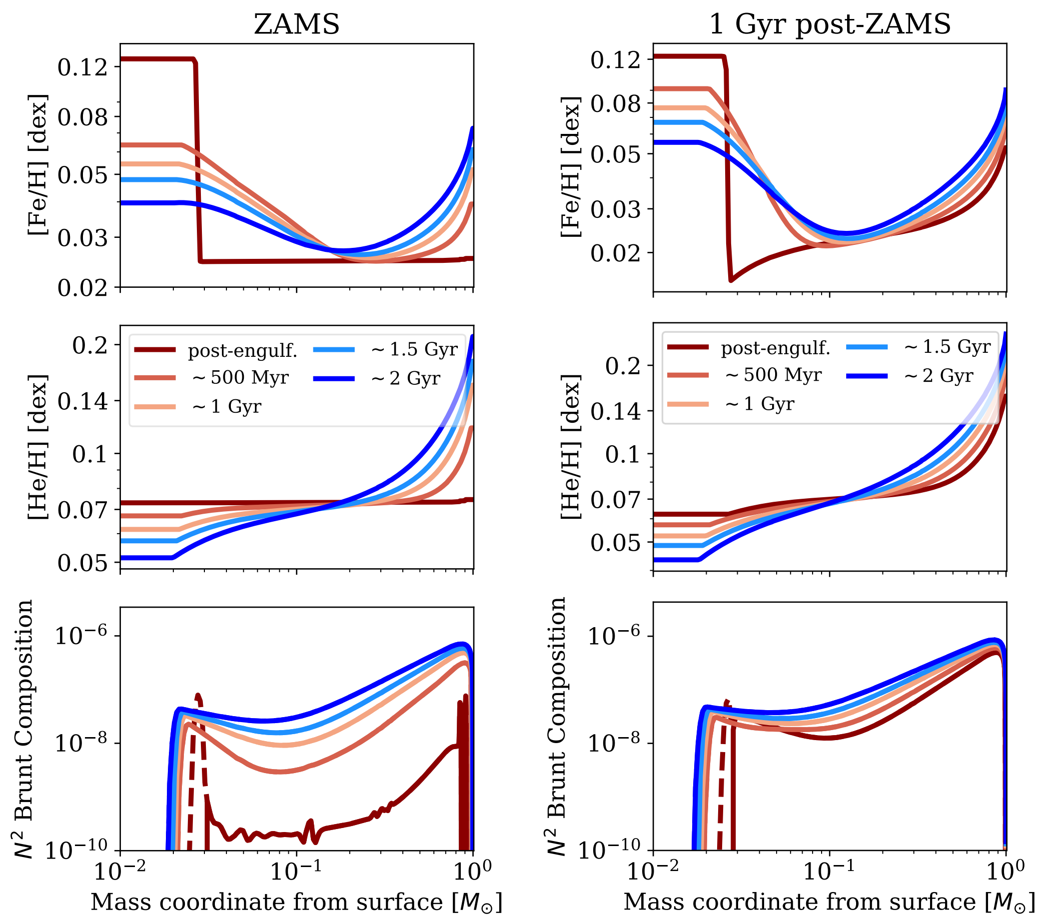


Figure 5.6: Plots of [Fe/H], [He/H], and $N^2_{\rm comp}$ as a function of mass coordinate for a 1 $M_\odot$ MESA stellar model. Negative $N^2_{\rm comp}$ values are shown via dashed line (absolute values are used to enable visualization in log scale). The left panel corresponds to the case of a 1 $M_\odot$ model engulfing 10 $M_\oplus$ of planetary material exactly at ZAMS, while the right panel corresponds to an equivalent case, but with engulfment occurring at 1 Gyr post-ZAMS. We sampled these quantities at five different time points (measured relative to the ZAMS age) from engulfment to an age of ~2 Gyr, as depicted by different line colors.

metallicity stars have thinner convective envelopes. The contributions of both these factors to convective zone metal content before engulfment is

$$M_{\rm Z,CZ} = f_{\rm Z,CZ}\, M_{\rm CZ}, \tag{5.3}$$

where $f_{\rm Z,CZ}$ is the mass fraction of refractory species in the convective zone, and $M_{\rm CZ}$ is the convective zone mass. We computed $M_{\rm CZ}$ for the three metallicity cases to be ~0.014 $M_\odot$, ~0.024 $M_\odot$, and ~0.030 $M_\odot$ for Z = 0.012, 0.017, and

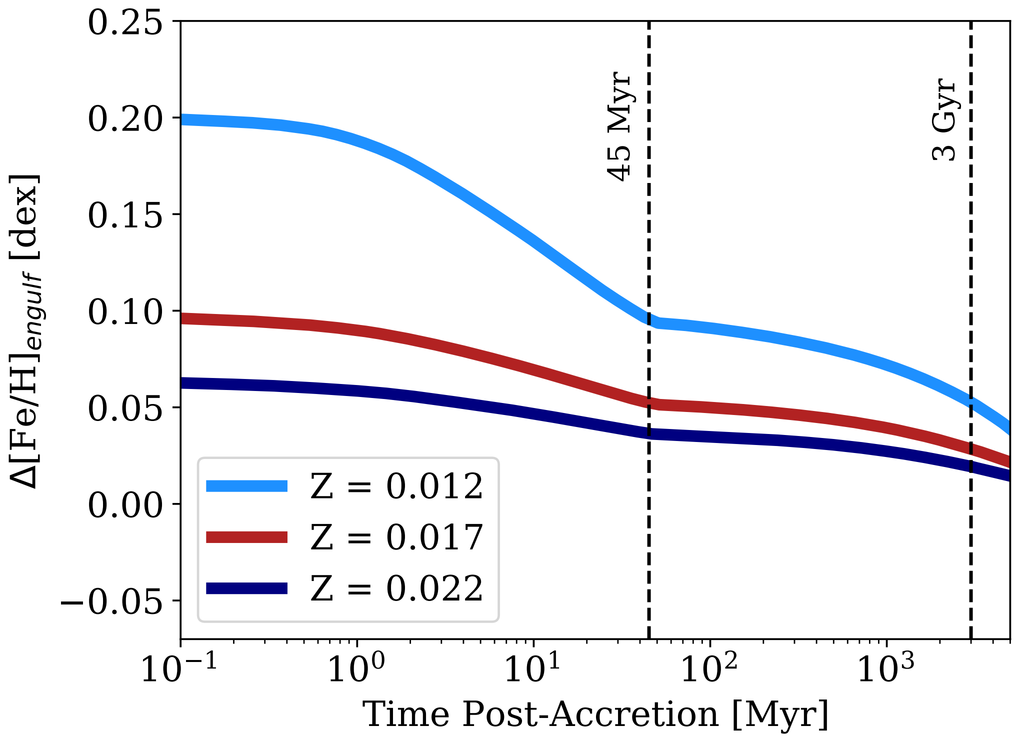


Figure 5.7: Post-engulfment [Fe/H] abundances over time for 1 $M_{\odot}$ models following near-instantaneous accretion of a 10 $M_{\oplus}$ bulk-Earth composition planetary companion at ZAMS. The three 1 $M_{\odot}$ models have different metallicities of slightly sub-solar (Z = 0.012, blue), solar (Z = 0.017, red), and slightly super-solar (Z = 0.022, navy). The abundances of comparison `MESA` models with no accretion were subtracted off. The refractory enhancements drop by a factor of ∼2 after ∼45 Myr. For the sub-solar metallicity model, observable signatures last for ∼3 Gyr.

0.022, respectively. The convective zone refractory mass fraction $f_{Z,\mathrm{CZ}}$ can be approximated as Z, so M_{CZ} increases slightly faster with metallicity than $f_{Z,\mathrm{CZ}}$, but both factors are nearly equally important for setting the pre-engulfment metal content. Because the total metal content in the convective envelope is much larger at higher metallicity, an engulfed planet has a proportionally smaller effect.

Notably, the lowest metallicity model with Z = 0.012 exhibits an engulfment signature that remains above the 0.05 dex observable level for ∼3 Gyr. This implies that stars with sub-solar metallicities may exhibit engulfment signatures that are fairly long lived. For context, the Z distribution of ∼300 G dwarfs in the Solar neighborhood ranges from ∼0.001−0.043, and peaks at ∼0.009 dex (Rocha-Pinto and Maciel, 1996). A more recent study of ∼17,000 K and ∼24,000 G dwarfs outside the Solar vicinity at 0.2−2.3 kpc from the Galactic plane derived similar Z

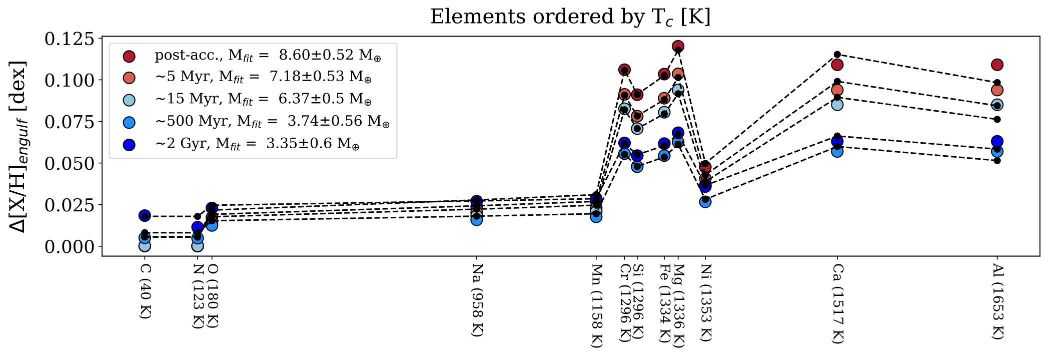


Figure 5.8: Differential abundances between a 1 $M_\odot$ model and a 1 $M_\odot$ model that accreted 10 $M_\oplus$ of bulk-Earth composition material (colored circles) for several post-engulfment times. The abundances are ranked by T_c of elements for solar-composition gas from Lodders (2003). The modeled abundances are represented by the black dots. The amount of apparent engulfed material at each time is provided in the upper left corner of the plot.

distributions, ranging from ~0.0002−0.03 and peaking at ~0.01 (Schlesinger et al., 2012). Thus, the distribution of solar-like stars in our neighborhood leans relatively metal-poor, indicating that observable engulfment signatures may persist on ~3 Gyr timescales for many nearby $\sim 1\ M_\odot$ stars.

Binary Observation Simulations

As mentioned in the section on near-instantaneous engulfment, engulfment signatures are detected in observations as patterns in the differential abundances of an engulfing star and a comparison stellar companion. The characteristic pattern indicative of an engulfment event is a trend of increasing differential abundances as a function of the elemental condensation temperature T_c (e.g., Oh et al. 2018). We expect a T_c-dependent abundance pattern following engulfment; elements with higher T_c are more likely to remain in the condensed phase throughout the disk and become locked in solid planetary material. Thus, planet compositions will exhibit higher abundances of species in order of T_c, which will be reflected in the photosphere of a star that has engulfed a planet.

We simulated how engulfment signatures would appear in binary system observations by computing differential abundances for elements that span a range of T_c between two G-type (1 $M_\odot$) stars, one of which underwent engulfment of a 10 $M_\oplus$ bulk-Earth composition planet at ZAMS. The differential abundances from `MESA` display the characteristic T_c trend that signifies engulfment (Figure 5.8, colored points). We also constructed an engulfment model similar to that of Oh et al. (2018) for estimating the mass of bulk-Earth composition (McDonough, 2003) material engulfed in one star given abundance measurements for a binary pair. This model estimates the apparent engulfed mass, which is the quantity inferred from stellar surface measurements assuming the planetary material is mixed throughout the convective zone, but not deeper.

Assuming solar abundances for the engulfing star [X/H] (Asplund et al., 2009), the model computes the mass fraction of each element X as follows:

$$f_{\mathrm{X,photo}} = \frac{10^{[\mathrm{X/H}]} m_{\mathrm{X}}}{\Sigma_{\mathrm{X}} 10^{[\mathrm{X/H}]} m_{\mathrm{X}}}, \tag{5.4}$$

where m_{X} is the mass of each element. Given a total mass of accreted material M_{acc} and accreted mass fractions for each element $f_{\mathrm{X,acc}}$, the abundance difference between the stars is

$$\Delta[\mathrm{X/H}] = \log_{10} \frac{f_{\mathrm{X,photo}}\, f_{\mathrm{CZ}}\, M_{\mathrm{star}} + f_{\mathrm{X,acc}}\, M_{\mathrm{acc}}}{f_{\mathrm{X,photo}}\, f_{\mathrm{CZ}}\, M_{\mathrm{star}}}, \tag{5.5}$$

where f_{CZ} is the mass fraction of the outer convective zone. For more details on the engulfment model, see Oh et al. (2018). Our implementation makes use of the `dynesty` nested sampling code (Speagle, 2020) to fit the estimated amount of engulfed mass. We also added abundance scatter estimated as a function of observation signal-to-noise ratio (SNR) as reported by Brewer and Fischer (2018), assuming an SNR level of 200/pix to mimic high quality observations.

We fit for the apparent amount of engulfed mass according to our model at five different time points that span from immediately post-engulfment to ~2 Gyr after the engulfment event (Figure 5.8, black points connected via dashed line). The fitted mass begins at 8.60 ± 0.52 $M_\oplus$ and falls to 3.35 ± 0.60 $M_\oplus$ over this time period. Our model underestimates the initial engulfed mass because it assumes all accreted material is initially contained in the convective zone, which is not completely accurate; overshoot mixing pulls engulfed material underneath the convective zone to an additional depth of ~10% the convective zone mass. We note that while the signature appears clear even at the ~2 Gyr post-engulfment point, abundance measurements from real observations will exhibit scatter due to instrumental effects, observing conditions, etc., making refractory enhancements below the 0.05 dex level at this point nearly undetectable.

5.3 Discussion

The strength and duration of planet engulfment signatures are affected by the amount of mass engulfed, the properties of the engulfing star, and the conditions of the engulfment event, e.g., when the event occurred within the lifetime of the system. We discuss how these factors influence engulfment signature evolution here.

We considered engulfed masses of 1, 10, and 50 $M_\oplus$. 10 $M_\oplus$ engulfment can be considered nominal as formation of a solid ~10 $M_\oplus$ core triggers runaway gas accretion according to the core accretion model of planet formation. However some giant planets may possess up to ~100 $M_\oplus$ worth of refractory material, perhaps due to protoplanet merger events that occurred prior to disk outgassing (Ginzburg and Chiang, 2020). We tested planet masses up to 50 $M_\oplus$ to account for occasional engulfment of such "heavy metal Jupiters". Larger amounts of engulfed mass result in higher initial enrichment, but also stronger μ-gradients that enhance thermohaline mixing and subsequent refractory depletion. These effects are illustrated throughout different engulfing star mass regimes (Figure 5.2). For low-mass stars (0.5−0.7 $M_\odot$), observable enrichment is only achieved for engulfed planetary masses of ≥50 $M_\oplus$, and last for ~900 Myr−3 Gyr. For more massive, solar-like (0.8−1.2 $M_\odot$) stars with thinner convective envelopes, observable signatures result from 10 $M_\oplus$ engulfment, and last for ~1 Myr−2 Gyr. Thus, engulfment signature detection in stars older than 2 Gyr assuming near-instantaneous nominal 10 $M_\oplus$ engulfment near ZAMS will be rare.

The lifetimes of observable engulfment signatures can increase under different accretion scenarios, or with different engulfing star parameters. Late-stage engulfment of a 10 $M_\oplus$ planet by a 1 $M_\odot$ star at 300 Myr−3 Gyr post-ZAMS produces observable enrichment on timescales of ~1.5 Gyr. For comparison, engulfment under the same conditions but occurring exactly at ZAMS results in signatures that remain observable for only ~90 Myr. Lowering the engulfing star metallicity also increases engulfment signature lifetimes; 1 $M_\odot$ models with sub-solar metallicities ($Z = 0.012$) sustain >0.05 dex levels of enrichment for ~3 Gyr following engulfment of a 10 $M_\oplus$ planet. Thus, engulfment signatures in stars older than 1.5 Gyr are more likely to be observed if the star is low metallicity, and/or if engulfment occurred late in the lifetime of the system.

Comparison to Observations

As mentioned in Section 5.1, there are several reported binaries with >0.05 dex refractory differences. All the potentially engulfing stars in these systems are in the solar-like mass regime, and those with reported ages are older than 1.5−2 Gyr. Late ages are typical for these systems; the strongest engulfment detection reported thus far is HD 240429-30 (Kronos-Krios), which is 4.0−4.28 Gyr old (Oh et al., 2018). Such refractory enrichments are unlikely to stem from ZAMS engulfment because our `MESA` models do not exhibit observable engulfment signatures on timescales longer than ~2 Gyr for solar-like stars under these conditions. Late-stage (300 Myr−3 Gyr post-ZAMS) engulfment is also unlikely as signatures will persist on timescales of only ~1.5 Gyr for 1 $M_\odot$ stars. However, four of the reported systems have sub-solar metallicities (ζ^2 Reticuli, HD 134439−40, HIP 34407−26, and HD 133131) below $Z = 0.012$, which increase signature timescales from nominal 10 $M_\oplus$ engulfment to ~3 Gyr.

The non-engulfing (Krios) and engulfing (Kronos) companions of Kronos-Krios have metallicities of [Fe/H] = 0.01 and 0.20 dex ($Z = 0.017$ and 0.027), respectively. Thus, potential explanations for the strong signature (15 $M_\oplus$ of estimated engulfed material, Oh et al. 2018) in this system could be some combination of late-stage engulfment, a large ($\gtrsim$50 $M_\oplus$) amount of mass engulfed, and chance observation of Kronos-Krios soon after its engulfment event. An alternative explanation is that the Kronos-Krios abundance differences are primordial rather than due to planet engulfment. This possibility is supported by the large projected separation of the system (11,000 ± 12 AU), which we calculated from Gaia DR3 astrometry. These lengths exceed typical turbulence scales in molecular clouds (0.05−0.2 pc, Brunt, Heyer, and Mac Low 2009 and references therein), indicating that Kronos-Krios may have formed in distinct areas of chemodynamical space within their birth cloud.

Stellar twin pairs ($\Delta T_{\rm eff}$ < 200 K, Andrews et al. 2019) are best for uncovering engulfment signatures because their abundances are free from differences due to mass-dependent evolution. However, many stellar binaries are composed of stars with different masses and stellar types. We found that higher mass stars ($M \gtrsim 1\ M_\odot$) with thinner convective zones exhibit faster refractory depletion following engulfment (Figure 5.1). Sevilla, Behmard, and Fuller (2022) found this can be true even in the absence of engulfment due to gravitational settling. Thus, binary pair stars with different masses that begin chemically homogeneous may have different surface abundances at late times, even without planet engulfment. We thus argue that engulfment signatures are most reliable for twin binaries with nearly equal mass components, and observations targeting planet engulfment should focus on such systems.

Our models indicate that engulfment signatures are longest lived in stars with $M \approx 1.1-1.2\ M_\odot$ (ZAMS temperatures of $T_{\rm eff} \approx 6100-6400$ K) and are thus more likely to be observed in stars slightly hotter than the Sun. We also predict longer-lived signatures in lower metallicity stars, which are hotter at the same mass. While the primordial metallicity (i.e., the metal content not including the engulfed planetary material) of stars is not easy to determine, we predict that stars with observable engulfment signatures will have lower primordial metallicity on average than non-enriched stars. This prediction could be tested with abundance measurements of non-refractory material (e.g., carbon or nitrogen) to estimate the pre-engulfment metallicity of the star.

Limitations

There are a few limitations in our `MESA` implementation that could be addressed in future studies. One issue is how we implemented mixing beyond thermohaline, overshoot, diffusion, and radiative levitation to account for other poorly understood mixing processes that operate within stellar interiors, e.g., turbulence and rotationally induced mixing. We used a constant `min_D_mix` coefficient to add this extra mixing in order to match observations of lithium depletion in solar-like stars (Sevilla, Behmard, and Fuller, 2022). However, this simple prescription will not reproduce the higher amount of mixing expected below the convective zones of massive stars due to, e.g., meridional circulation. Radiative levitation will also play a more prominent role in hotter and more massive stars. More accurate prescriptions for additional mixing processes would be a good addition to future studies of planet engulfment.

As discussed in the section on bulk Earth accretion, we were unable to implement instantaneous accretion representing planet engulfment because `MESA` failed to converge. Instead, we implemented engulfment via rapid accretion, resulting in an engulfment duration of slightly shorter than 10,000 years. While this period is much shorter than timescales of refractory depletion (>1 Myr), it is still much longer than instantaneous engulfment, which should occur within years. In addition, we did not consider the scenario of rapid engulfment that could plunge an engulfed planet deep within the stellar interior, and lead to its slow disintegration within the star (Jia and Spruit, 2018). This could even result in the planet plunging past the convective zone in more massive stars, significantly decreasing the strength and duration of refractory enhancements in the photosphere. These issues should be taken into consideration to more accurately model different engulfment scenarios.

5.4 Summary

We used `MESA` models to simulate planet engulfment and mixing of material throughout the star due to convection, thermohaline mixing, diffusion, gravitational settling, and radiative levitation. We examined the evolution of measurable surface abundances under a range of accretion scenarios by varying the engulfed planet mass, the properties of the engulfing star, and the engulfment time and/or duration (pre-MS disk accretion, ZAMS, late heavy bombardment, and late-stage). We found that these conditions greatly affect the strength and duration of resultant planet engulfment signatures.

Near-instantaneous engulfment occurring when the engulfing star reaches ZAMS results in different timescales for observable engulfment signatures as a function of engulfing star mass. At solar metallicity, the signature is largest and longest-lived for stars with $M \approx 1.1-1.2\ M_\odot$. We found that following planetary engulfment, thermohaline mixing quickly depletes the engulfment signature, lowering the increased surface abundances by a factor of ~2 in the first ~5−45 Myr. Afterwards, gravitational settling further depletes surface abundance enhancements on longer timescales. Observable signatures last for less than ~2 Gyr in all our models, apart from our $1.2\,M_\odot$ model where radiative levitation causes a ~5 Gyr post-engulfment increase in surface abundances of certain chemical species (e.g., ^{56}Fe, ^{27}Al, ^{28}Si). Engulfment scenarios mimicking LHB or pre-MS disk accretion result in shorter observable signature timescales compared to ZAMS engulfment. In LHB engulfment, refractory material is continuously depleted throughout longer accretion periods, resulting in almost no refractory enrichment post-engulfment. In disk accretion, refractory material is ingested by the star earlier in its lifetime, and is more heavily diluted within its larger convective envelope (relative to the size of a convective envelope of a star at ZAMS). This results in signatures weaker than those of ZAMS scenarios, but slightly stronger than those of LHB scenarios.

Observable engulfment signature lifetimes increase for late-stage engulfment (300 Myr−3 Gyr post-ZAMS) scenarios, where the signatures of $10\,M_\oplus$ engulfment can persist for ~1.5 Gyr in solar-like stars. Observable signatures are also more prominent for low-metallicity stars, lasting for ∼3 Gyr in 1 $M_\odot$ stars with Z = 0.012 and ZAMS accretion. These conditions are thus more likely explanations for engulfment signatures observed in solar-like stars with ages >1.5 Gyr.

The strong dependence of engulfment signature strength and duration on stellar type, along with fewer theoretical uncertainties when modeling equal-mass stars, both underscore that stellar twin binaries are best-suited for observational planet engulfment surveys. We conclude that twin binaries with 1.1–1.2 $M_\odot$ masses are the most promising targets for engulfment detections.

References

Adibekyan, V. et al. (June 2016). "ζ^2 Reticuli, its debris disk, and its lonely stellar companion ζ^1 Ret. Different T_c trends for different spectra". In: *Astronomy and Astrophysics* 591, A34, A34. DOI: 10.1051/0004-6361/201628453. arXiv: 1605.01918 [astro-ph.SR].

Andrews, Jeff J. et al. (Jan. 2019). "Using APOGEE Wide Binaries to Test Chemical Tagging with Dwarf Stars". In: *The Astrophysical Journal* 871.1, 42, p. 42. DOI: 10.3847/1538-4357/aaf502. arXiv: 1811.12032 [astro-ph.SR].

Asplund, Martin et al. (Sept. 2009). "The Chemical Composition of the Sun". In: *Annual Review of Astronomy and Astrophysics* 47.1, pp. 481–522. DOI: 10.1146/annurev.astro.46.060407.145222. arXiv: 0909.0948 [astro-ph.SR].

Bauer, Evan B. and Lars Bildsten (Feb. 2019). "Polluted White Dwarfs: Mixing Regions and Diffusion Timescales". In: *The Astrophysical Journal* 872.1, 96, p. 96. DOI: 10.3847/1538-4357/ab0028. arXiv: 1812.09602 [astro-ph.SR].

Biazzo, K. et al. (Nov. 2015). "The GAPS programme with HARPS-N at TNG. X. Differential abundances in the XO-2 planet-hosting binary". In: *Astronomy and Astrophysics* 583, A135, A135. DOI: 10.1051/0004-6361/201526375. arXiv: 1506.01614 [astro-ph.SR].

Bottke, William F. and Marc D. Norman (Aug. 2017). "The Late Heavy Bombardment". In: *Annual Review of Earth and Planetary Sciences* 45.1, pp. 619–647. DOI: 10.1146/annurev-earth-063016-020131.

Brassard, P. et al. (Feb. 1991). "Adiabatic Properties of Pulsating DA White Dwarfs. I. The Treatment of the Brunt-Vaeisaelae Frequency and the Region of Period Formation". In: *The Astrophysical Journal* 367, p. 601. DOI: 10.1086/169655.

Brewer, J. M. and D. A. Fischer (Aug. 2018). "Spectral Properties of Cool Stars: Extended Abundance Analysis of Kepler Objects of Interest". In: *The Astrophysical Journal Supplement Series* 237, 38, p. 38. DOI: 10.3847/1538-4365/aad501. arXiv: 1804.00673 [astro-ph.SR].

Brown, Justin M., Pascale Garaud, and Stephan Stellmach (Apr. 2013). "Chemical Transport and Spontaneous Layer Formation in Fingering Convection in Astrophysics". In: *The Astrophysical Journal* 768.1, p. 34. ISSN: 1538-4357. DOI: 10.1088/0004-637x/768/1/34. arXiv: 1212.1688. URL: http://dx.doi.org/10.1088/0004-637X/768/1/34.

Brunt, C. M., M. H. Heyer, and M. -M. Mac Low (Sept. 2009). "Turbulent driving scales in molecular clouds". In: *Astronomy and Astrophysics* 504.3, pp. 883–890. DOI: 10.1051/0004-6361/200911797. arXiv: 0910.0398 [astro-ph.GA].

Campilho, B., M. Deal, and D. Bossini (Mar. 2022). "Atomic diffusion in solar-like stars with MESA. Comparison with the Montreal/Montpellier and CESTAM stellar evolution codes". In: *Astronomy and Astrophysics* 659, A162, A162. DOI: 10.1051/0004-6361/202140821. arXiv: 2201.03439 [astro-ph.SR].

De Silva, G. M., K. C. Freeman, and J. Bland-Hawthorn (Apr. 2009). "Reconstructing Fossil Sub-structures of the Galactic Disk: Clues from Abundance Patterns of Old Open Clusters and Moving Groups". In: *Publications of the Astronomical Society of the Pacific* 26, pp. 11–16. DOI: 10.1071/AS08019. arXiv: 0810.2287.

Deal, M. et al. (Jan. 2020). "Chemical mixing in low mass stars. I. Rotation against atomic diffusion including radiative acceleration". In: *Astronomy and Astrophysics* 633, A23, A23. DOI: 10.1051/0004-6361/201936666. arXiv: 1910.14335 [astro-ph.SR].

Denissenkov, Pavel A. (Nov. 2010). "Numerical Simulations of Thermohaline Convection: Implications for Extra-mixing in Low-mass RGB Stars". In: *The Astrophysical Journal* 723.1, pp. 563–579. DOI: 10.1088/0004-637X/723/1/563. arXiv: 1006.5481 [astro-ph.SR].

Galarza, Jhon Yana et al. (Nov. 2021). "Evidence of Rocky Planet Engulfment in the Wide Binary System HIP 71726/HIP 71737". In: *The Astrophysical Journal* 922.2, p. 129. DOI: 10.3847/1538-4357/ac2362. URL: https://dx.doi.org/10.3847/1538-4357/ac2362.

Garaud, Pascale (Feb. 2011). "What happened to the other Mohicans? Realistic models of metallicity dilution by fingering convection and observational implications". In: *arXiv e-prints*, arXiv:1011.3452, arXiv:1011.3452. arXiv: 1011.3452 [astro-ph.SR].

Ginzburg, Sivan and Eugene Chiang (Oct. 2020). "Heavy-metal Jupiters by major mergers: metallicity versus mass for giant planets". In: *Monthly Notices of The Royal Astronomical Society* 498.1, pp. 680–688. DOI: 10.1093/mnras/staa2500. arXiv: 2006.12500 [astro-ph.EP].

Hartmann, Lee, Gregory Herczeg, and Nuria Calvet (Sept. 2016). "Accretion onto Pre-Main-Sequence Stars". In: *Annual Review of Astronomy and Astrophysics* 54, pp. 135–180. DOI: 10.1146/annurev-astro-081915-023347.

Hawkins, Keith et al. (Feb. 2020). “Identical or fraternal twins? The chemical homogeneity of wide binaries from Gaia DR2”. In: *Monthly Notices of The Royal Astronomical Society* 492.1, pp. 1164–1179. DOI: 10.1093/mnras/stz3132. arXiv: 1912.08895 [astro-ph.SR].

Herwig, Falk (Aug. 2000). “The evolution of AGB stars with convective overshoot”. In: *Astronomy and Astrophysics* 360.1, pp. 952–968. arXiv: 0007139 [astro-ph.SR].

Hu, Haili et al. (Nov. 2011). “Slowing down atomic diffusion in subdwarf B stars: mass loss or turbulence?” In: *Monthly Notices of the Royal Astronomical Society* 418.1, pp. 195–205. DOI: 10.1111/j.1365-2966.2011.19482.x. arXiv: 1108.1318 [astro-ph.SR].

Hui-Bon-Hoa, Alain (Aug. 2008). “The Toulouse Geneva Evolution Code (TGEC)”. In: *Astrophysics and Space Science* 316.1-4, pp. 55–60. DOI: 10.1007/s10509-007-9605-6.

Jia, Shi and H. C. Spruit (Sept. 2018). “Disruption of a Planet Spiraling into its Host Star”. In: *The Astrophysical Journal* 864.2, 169, p. 169. DOI: 10.3847/1538-4357/aad77c. arXiv: 1808.00467 [astro-ph.EP].

Jofré, Emiliano et al. (Dec. 2021). “The Peculiar Chemical Pattern of the WASP-160 Binary System: Signatures of Planetary Formation and Evolution?” In: *The Astronomical Journal* 162.6, 291, p. 291. DOI: 10.3847/1538-3881/ac25ef. arXiv: 2109.04590 [astro-ph.EP].

Kippenhahn, R., G. Ruschenplatt, and H. -C. Thomas (Nov. 1980). “The time scale of thermohaline mixing in stars”. In: *Astronomy and Astrophysics* 91.1-2, pp. 175–180.

Lodders, Katharina (July 2003). “Solar System Abundances and Condensation Temperatures of the Elements”. In: *The Astrophysical Journal* 591.2, pp. 1220–1247. DOI: 10.1086/375492.

Mack Claude E., III et al. (June 2014). “Detailed Abundances of Planet-hosting Wide Binaries. I. Did Planet Formation Imprint Chemical Signatures in the Atmospheres of HD 20782/81?” In: *The Astrophysical Journal* 787.2, 98, p. 98. DOI: 10.1088/0004-637X/787/2/98. arXiv: 1404.1967 [astro-ph.EP].

McDonough, W. F. (Dec. 2003). “Compositional Model for the Earth’s Core”. In: *Treatise on Geochemistry* 2, p. 568. DOI: 10.1016/B0-08-043751-6/02015-6.

Metzger, B. D., D. Giannios, and D. S. Spiegel (Oct. 2012). “Optical and X-ray transients from planet-star mergers”. In: *Monthly Notices of The Royal Astronomical Society* 425.4, pp. 2778–2798. DOI: 10.1111/j.1365-2966.2012.21444.x. arXiv: 1204.0796 [astro-ph.EP].

Nagar, Tushar, Lorenzo Spina, and Amanda I. Karakas (Jan. 2020). “The Chemical Signatures of Planetary Engulfment Events in Binary Systems”. In: *The Astrophysical Journal Letters* 888.1, L9, p. L9. DOI: 10.3847/2041-8213/ab5dc6.

Oh, Semyeong et al. (Feb. 2018). “Kronos and Krios: Evidence for Accretion of a Massive, Rocky Planetary System in a Comoving Pair of Solar-type Stars”. In: *The Astrophysical Journal* 854.2, 138, p. 138. DOI: 10.3847/1538-4357/aaab4d. arXiv: 1709.05344 [astro-ph.SR].

Paxton, Bill, Lars Bildsten, et al. (Jan. 2011). “Modules for Experiments in Stellar Astrophysics (MESA)”. In: *The Astrophysical Journal Supplement Series* 192.1, 3, p. 3. DOI: 10.1088/0067-0049/192/1/3. arXiv: 1009.1622 [astro-ph.SR].

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* 208.1, 4, p. 4. DOI: 10.1088/0067-0049/208/1/4. arXiv: 1301.0319 [astro-ph.SR].

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* 220.1, 15, p. 15. DOI: 10.1088/0067-0049/220/1/15. arXiv: 1506.03146 [astro-ph.SR].

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* 234.2, 34, p. 34. DOI: 10.3847/1538-4365/aaa5a8. arXiv: 1710.08424 [astro-ph.SR].

Paxton, Bill, R. Smolec, et al. (July 2019). “Modules for Experiments in Stellar Astrophysics (MESA): Pulsating Variable Stars, Rotation, Convective Boundaries, and Energy Conservation”. In: *The Astrophysical Journal Supplement Series* 243.1, 10, p. 10. DOI: 10.3847/1538-4365/ab2241. arXiv: 1903.01426 [astro-ph.SR].

Pinsonneault, M. H., D. L. DePoy, and M. Coffee (July 2001). “The Mass of the Convective Zone in FGK Main-Sequence Stars and the Effect of Accreted Planetary Material on Apparent Metallicity Determinations”. In: *The Astrophysical Journal Letters* 556.1, pp. L59–L62. DOI: 10.1086/323531. arXiv: astro-ph/0105257 [astro-ph].

Prat, V., F. Lignières, and N. Lagarde (Dec. 2015). “Numerical simulations of zero-Prandtl-number thermohaline convection”. In: *SF2A-2015: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*, pp. 419–422.

Ramírez, I., S. Khanal, P. Aleo, et al. (July 2015). “The Dissimilar Chemical Composition of the Planet-hosting Stars of the XO-2 Binary System”. In: *The Astrophysical Journal* 808, 13, p. 13. DOI: 10.1088/0004-637X/808/1/13. arXiv: 1506.01025 [astro-ph.SR].

Ramírez, I., S. Khanal, S. J. Lichon, et al. (Dec. 2019). “The chemical composition of HIP 34407/HIP 34426 and other twin-star comoving pairs”. In: *Monthly Notices of The Royal Astronomical Society* 490.2, pp. 2448–2457. DOI: 10.1093/mnras/stz2709. arXiv: 1909.07460 [astro-ph.SR].

Ramírez, I., J. Meléndez, et al. (Oct. 2011). “Elemental Abundance Differences in the 16 Cygni Binary System: A Signature of Gas Giant Planet Formation?” In: *The Astrophysical Journal* 740.2, 76, p. 76. DOI: 10.1088/0004-637X/740/2/76. arXiv: 1107.5814 [astro-ph.SR].

Richard, Olivier, Sylvie Théado, and Sylvie Vauclair (Apr. 2004). “Updated Toulouse solar models including the diffusion-circulation coupling and the effect of μ-gradients”. In: *Solar Physics* 220.2, pp. 243–259. DOI: 10.1023/B:SOLA.0000031394.68639.0b. arXiv: astro-ph/0401568 [astro-ph].

Richer, J. et al. (Jan. 1998). “Radiative Accelerations for Evolutionary Model Calculations”. In: *The Astrophysical Journal* 492.2, pp. 833–842. DOI: 10.1086/305054.

Rocha-Pinto, H. J. and W. J. Maciel (Mar. 1996). “The metallicity distribution of G dwarfs in the solar neighbourhood.” In: *Monthly Notices of the Royal Astronomical Society* 279.2, pp. 447–458. DOI: 10.1093/mnras/279.2.447. arXiv: astro-ph/9510101 [astro-ph].

Saffe, C., M. Flores, et al. (Apr. 2016). “Temperature condensation trend in the debris-disk binary system ζ^2 Reticuli”. In: *Astronomy and Astrophysics* 588, A81, A81. DOI: 10.1051/0004-6361/201528043. arXiv: 1602.01320 [astro-ph.SR].

Saffe, C., E. Jofré, et al. (July 2017). “Signatures of rocky planet engulfment in HAT-P-4. Implications for chemical tagging studies”. In: *Astronomy and Astrophysics* 604, L4, p. L4. DOI: 10.1051/0004-6361/201731430. arXiv: 1707.02180 [astro-ph.SR].

Schlesinger, Katharine J. et al. (Dec. 2012). “The Metallicity Distribution Functions of SEGUE G and K Dwarfs: Constraints for Disk Chemical Evolution and Formation”. In: *The Astrophysical Journal* 761.2, 160, p. 160. DOI: 10.1088/0004-637X/761/2/160. arXiv: 1112.2214 [astro-ph.SR].

Sestito, P. and S. Randich (Nov. 2005). “Time scales of Li evolution: a homogeneous analysis of open clusters from ZAMS to late-MS”. In: *Astronomy and Astrophysics* 442.2, pp. 615–627. DOI: 10.1051/0004-6361:20053482. arXiv: astro-ph/0507537 [astro-ph].

Sevilla, Jason, Aida Behmard, and Jim Fuller (Nov. 2022). “Long-term lithium abundance signatures following planetary engulfment”. In: *Monthly Notices of The Royal Astronomical Society* 516.3, pp. 3354–3365. DOI: 10.1093/mnras/stac2436. arXiv: 2207.13232 [astro-ph.SR].

Speagle, Joshua S. (Apr. 2020). “DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences”. In: *Monthly Notices of the Royal Astronomical Society* 493.3, pp. 3132–3158. DOI: 10.1093/mnras/staa278. arXiv: 1904.02180 [astro-ph.IM].

Teske, Johanna K., Luan Ghezzi, et al. (Mar. 2015). “Abundance Differences between Exoplanet Binary Host Stars XO-2N and XO-2S−Dependence on Stellar Parameters”. In: *The Astrophysical Journal Letters* 801.1, L10, p. L10. DOI: 10.1088/2041-8205/801/1/L10. arXiv: 1501.02167 [astro-ph.EP].

Teske, Johanna K., Sandhya Khanal, and Ivan Ramírez (Mar. 2016). “The Curious Case of Elemental Abundance Differences in the Dual Hot Jupiter Hosts WASP-94A and B”. In: *The Astrophysical Journal* 819.1, 19, p. 19. DOI: 10.3847/0004-637X/819/1/19. arXiv: 1601.01731 [astro-ph.SR].

Théado, Sylvie and Sylvie Vauclair (Jan. 2012). “Metal-rich Accretion and Thermohaline Instabilities in Exoplanet-host Stars: Consequences on the Light Elements Abundances”. In: *The Astrophysical Journal* 744.2, 123, p. 123. DOI: 10.1088/0004-637X/744/2/123. arXiv: 1109.4238 [astro-ph.SR].

Thoul, Anne A., John N. Bahcall, and Abraham Loeb (Feb. 1994). “Element Diffusion in the Solar Interior”. In: *The Astrophysical Journal* 421, p. 828. DOI: 10.1086/173695. arXiv: astro-ph/9304005 [astro-ph].

Traxler, A., P. Garaud, and S. Stellmach (Feb. 2011). “Numerically Determined Transport Laws for Fingering (“Thermohaline”) Convection in Astrophysics”. In: *The Astrophysical Journal Letters* 728.2, L29, p. L29. DOI: 10.1088/2041-8205/728/2/L29. arXiv: 1011.3461 [astro-ph.SR].

Tucci Maia, M., J. Meléndez, D. Lorenzo-Oliveira, et al. (Aug. 2019). “Revisiting the 16 Cygni planet host at unprecedented precision and exploring automated tools for precise abundances”. In: *Astronomy and Astrophysics* 628, A126, A126. DOI: 10.1051/0004-6361/201935952. arXiv: 1906.04195 [astro-ph.SR].

Tucci Maia, M., J. Meléndez, and I. Ramírez (Aug. 2014). “High Precision Abundances in the 16 Cyg Binary System: A Signature of the Rocky Core in the Giant Planet”. In: *The Astrophysical Journal Letters* 790, L25, p. L25. DOI: 10.1088/2041-8205/790/2/L25. arXiv: 1407.4132 [astro-ph.SR].

Ulrich, Roger K. (Feb. 1972). “Thermohaline Convection in Stellar Interiors.” In: *The Astrophysical Journal* 172, p. 165. DOI: 10.1086/151336.

Vauclair, Sylvie (Apr. 2004). “Metallic Fingers and Metallicity Excess in Exoplanets’ Host Stars: The Accretion Hypothesis Revisited”. In: *The Astrophysical Journal* 605.2, pp. 874–879. DOI: 10.1086/382668. arXiv: astro-ph/0309790 [astro-ph].

Chapter 6

ELEMENTAL ABUNDANCES OF *KEPLER* OBJECTS OF INTEREST IN APOGEE DR17

Behmard, Aida et al. (Apr. 2023). "Elemental Abundances of Kepler Objects of Interest in APOGEE DR17". In: *The Astronomical Journal* 165.4, 178, p. 178. DOI: 10.3847/1538-3881/acc32a.

Abstract

The elemental abundances of planet host stars can shed light on the conditions of planet forming environments. We test if individual abundances of 130 known/candidate planet hosts in APOGEE are statistically different from those of a reference doppelgänger sample. The reference set comprises objects selected with the same $T_{\rm eff}$, $\log g$, [Fe/H], and [Mg/H] as each Kepler Object of Interest (KOI). We predict twelve individual abundances (X = C, N, O, Na, Al, Si, Ca, Ti, V, Cr, Mn, Ni) for the KOIs and their doppelgängers using a local linear model of these four parameters, training on ASPCAP abundance measurements for a sample of field stars with high fidelity (SNR > 200) APOGEE observations. We compare element prediction residuals (model−measurement) for the two samples and find them to be indistinguishable, given a high quality sample selection. We report median intrinsic dispersions of ~0.038 dex and ~0.041 dex, for the KOI and doppelgänger samples, respectively, for these elements. We conclude that the individual abundances at fixed $T_{\rm eff}$, $\log g$, [Fe/H], and [Mg/H] are unremarkable for known planet hosts. Our results establish an upper limit on the abundance precision required to uncover any chemical signatures of planet formation in planet host stars.

6.1 Introduction

The elemental abundances of planet host stars bear the fingerprint of the processes governing planet formation and evolution. For example, it is well established that stars hosting giant planets often have enhanced iron abundances ([Fe/H]; Gonzalez 1997; Heiter and Luck 2003; Santos, Israelian, and Mayor 2004; Fischer and Valenti 2005). This is typically regarded as evidence for the core accretion model of planet formation (e.g., Rice and Armitage 2003; Ida and Lin 2004; Alibert, Mordasini, and Benz 2011; Mordasini et al. 2012; Maldonado, Villaver, Eiroa, and Micela

2019); the host star [Fe/H] can be considered a proxy for the solid surface density of protoplanetary disks. In this context, more solids translate to rapid growth of planetary cores that can reach a critical mass of $\sim$10 $M_\oplus$ before the disk gas dissipates. This enables accretion of a substantial gaseous envelope. The planet-[Fe/H] trend appears to weaken with decreasing planet mass/radius (Sousa, Santos, et al., 2008; Ghezzi, Cunha, et al., 2010; Ghezzi, Montet, and Johnson, 2018; Schlaufman and Laughlin, 2011; Buchhave, Latham, et al., 2012; Buchhave and Latham, 2015; Wang et al., 2015), but becomes stronger with decreasing orbital period, particularly in the $P \lesssim 10$ days regime (Mulders et al., 2016; Narang et al., 2018; Petigura et al., 2018; Wilson et al., 2018; Sousa, Adibekyan, et al., 2019; Ghezzi, Martinez, et al., 2021). Thus, the distributions of planet masses, radii, and orbital periods are sculpted by the amount of available solids and therefore the host star metallicity and planet forming environment.

The connections between [Fe/H] and planet architectures are well-studied because there are many strong iron absorption lines in the spectra of solar-like stars, making it a relatively easy abundance to constrain. High precision abundances beyond iron are more challenging to measure, but can unveil more detailed relationships between host star chemistry and planet architectures. For example, Adibekyan, Santos, et al. (2012) found that Fe-poor ($-0.1 <$ [Fe/H] < 0.2 dex) hosts of small and giant planets exhibit enhanced [X/H] ratios for Mg, Al, Si, Sc, and Ti. The authors later examined a sample of even more Fe-depleted ($-0.65 <$ [Fe/H] < -0.3 dex) stars that host small, rocky planets, and found strong enhancements in Ti (Adibekyan, Delgado Mena, et al., 2012). Similarly, Maldonado, Villaver, and Eiroa (2018) found that Fe-poor host stars of cool Jupiters tend to be enhanced in alpha-elements. These results suggest that other refractory elements can compensate for low iron content during planet building block formation (e.g., Bashi and Zucker 2019).

Abundances beyond iron can also place constraints on planet formation locations and interior compositions. For example, the stellar C/O ratio characterizes the H_2O, CO_2, and CO ice lines in protoplanetary disks, and can be used as a sensitive tracer of formation location when compared to the C/O ratio of planetary atmospheres (Öberg, Murray-Clay, and Bergin, 2011); sub-stellar and super-stellar atmospheric C/O generally indicate planet formation within and beyond the H_2O ice line, respectively. The host star C/O ratio can also dictate if planetary compositions will be dominated by carbonates or silicates, with further ratios like Mg/Si determining the types of silicates in low C/O regimes (e.g., Brewer and Fischer 2017).

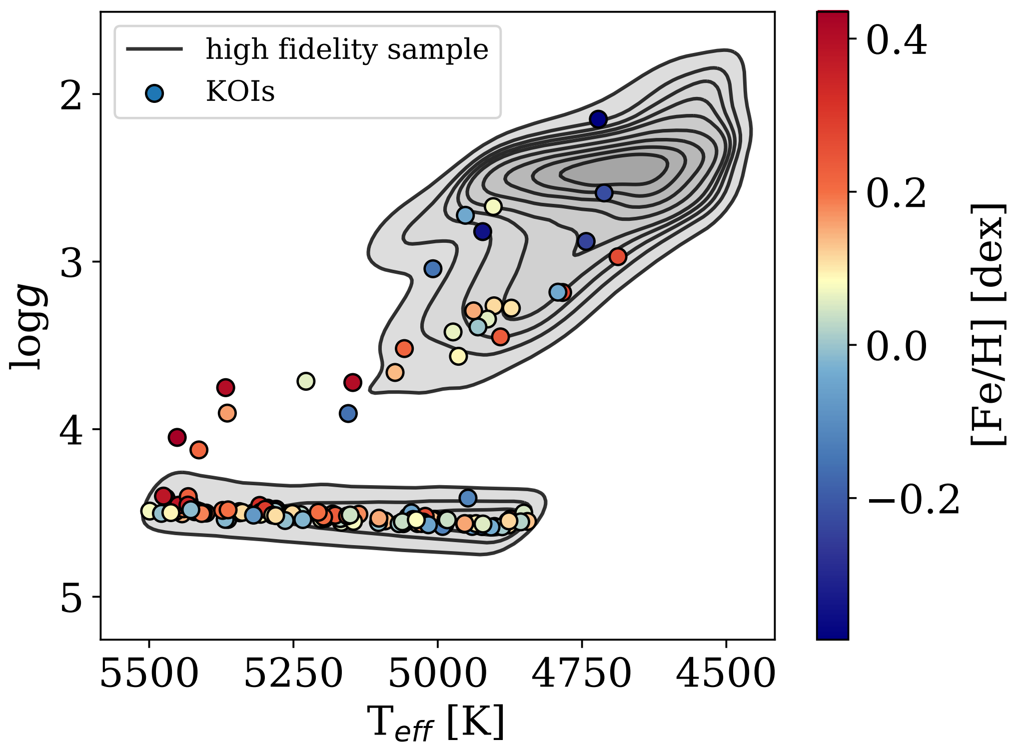


Figure 6.1: The $\log g$ vs. $T_{\rm eff}$ probability density functions for the ∼129,000 high fidelity APOGEE DR17 stars (gray). The contours represent areas encompassing 10% of the cumulative probability mass. The 130 non-'False Positive' KOIs included in the high fidelity sample with doppelgängers are shown as circles, with colors representing their [Fe/H] values.

Particular abundance patterns are also thought to be indicative of planet formation, as suggested by measured individual abundance trends with element condensation temperature (T_c). This is based on the premise that rocky planet-forming material more readily incorporates elements with high T_c (e.g., Ti, Al, Y) that reside in the solid phase throughout most of the disk. Conversely, low T_c elements (e.g., C, N, O) are more likely to remain in the gas phase. Planet compositions are thus characterized by larger abundances in order of increasing T_c. It follows that adding planetary material to host stars will create refractory enhancements in stellar photospheres and a positive abundance gradient with T_c. This could result from processes such as planet engulfment, or steady accretion of solids during Late Heavy Bombardment-like events. Depletion trends in order of T_c could likewise result from an absence of planetary material in host star photospheres. This could result from solids getting locked up in rocky planets and subsequent accretion of dust-depleted gas onto the host star (Meléndez et al., 2009), or from gaps in protoplanetary disks

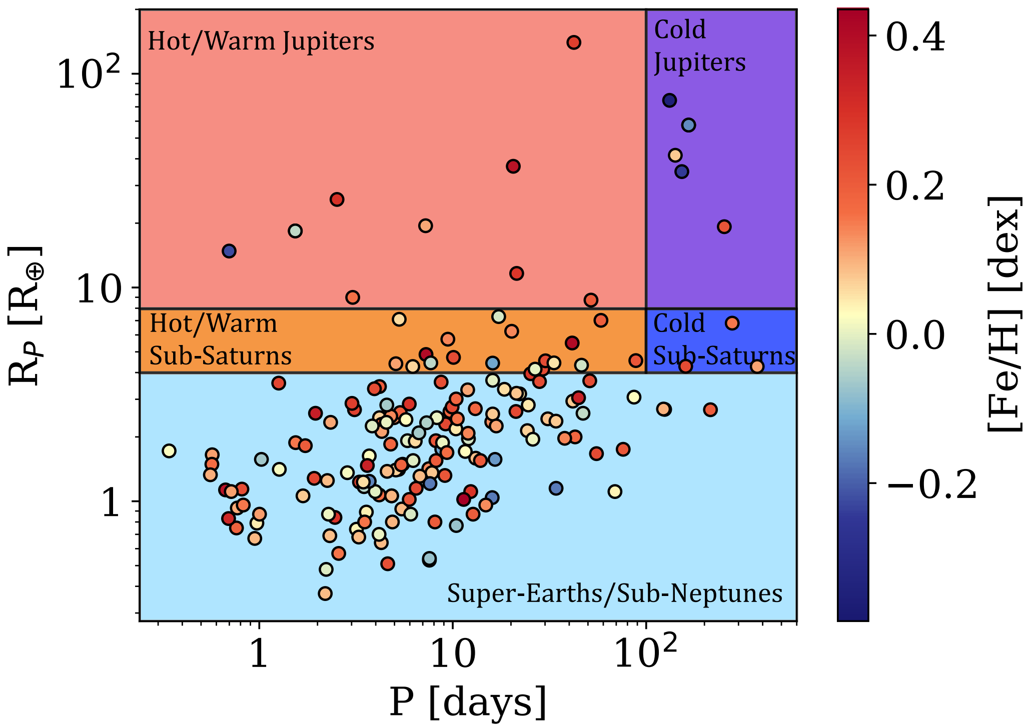


Figure 6.2: The radius vs. orbital period distribution for our sample of 130 KOI systems, with the ASPCAP-reported [Fe/H] of their host stars marked in color. Hot/warm Jupiters are defined as planets with $R > 8\ R_\oplus$ and $P < 100$ days (red region), hot/warm sub-Saturns with $4\ R_\oplus < R < 8\ R_\oplus$ and P <100 days (orange region), cold Jupiters with $R > 8\ R_\oplus$ and $P > 100$ days (purple region), cold sub-Saturns with $4\ R_\oplus < R < 8$ and $P > 100$ days (dark blue region), and super-Earths/sub-Neptunes with $R < 4\ R_\oplus$ (light blue region).

created by forming giant planets that prevent host star accretion of refractory material (Booth and Owen, 2020). Such trends with T_c have been observed in the differential abundances of several binary systems (Ramírez, Meléndez, Cornejo, et al., 2011; Mack et al., 2014; Tucci Maia, Meléndez, and Ramírez, 2014; Teske, Ghezzi, et al., 2015; Ramírez, Khanal, Aleo, et al., 2015; Biazzo et al., 2015; Saffe, Flores, et al., 2016; Teske, Khanal, and Ramírez, 2016; Adibekyan, Delgado-Mena, et al., 2016; Saffe, Jofré, et al., 2017; Oh et al., 2018; Tucci Maia, Meléndez, Lorenzo-Oliveira, et al., 2019; Ramírez, Khanal, Lichon, et al., 2019; Nagar, Spina, and Karakas, 2020; Galarza et al., 2021; Jofré et al., 2021), and in larger samples. For example, Nibauer et al. (2021) analyzed 1700 solar analogs from the Apache Point Galactic Evolution Experiment (APOGEE), and found that 70−90% of solar analogs appear depleted in refractory elements in order of T_c. Thus, there is ample evidence that abundance alteration via planet formation processes is common.

Stellar elemental abundances beyond iron are therefore important for understanding planet formation and evolution. Drawing connections between abundances and planet architectures require sufficiently large stellar samples to establish statistically significant correlations, as well as high precision (~0.01 dex uncertainties) abundance measurements (Meléndez et al., 2009; Ramírez, Meléndez, and Asplund, 2014; Schuler, Vaz, et al., 2015). Here, we utilize APOGEE DR17, which provides high-resolution spectra ($R \approx 22{,}500$) and derived parameters for >650,000 stars (Abdurro'uf et al., 2022). This enormous sample will boost abundance pattern statistics, making it possible to compromise on individual abundance precisions. The APOGEE DR17 parameters include individual abundances for 20 species, measured with the APOGEE Stellar Parameter and Chemical Abundances Pipeline (ASPCAP) pipeline that achieves typical abundance precisions of <0.1 dex (García Pérez et al., 2016). The full second generation APOGEE sample observed at the Apache Point Observatory (APOGEE-2N) contains 2098 stars also observed by Kepler, where 824 are confirmed planet hosts. This makes APOGEE DR17 an excellent sample for exploring connections between host star chemistry and planet formation. We describe our data selection further in Section 6.2.

Our goal is to examine individual abundances in planet hosts in isolation of other parameters, such as evolutionary state and overall metallicity. We want to determine if the individual abundances are differentiable in any way from the underlying field population (where planet membership is unknown). To this end, we take the Kepler Objects of Interest (KOIs, defined as stars that host confirmed or candidate planets) observed in APOGEE, and construct a reference set of doppelgängers with identical $T_{\rm eff}$, $\log g$, [Fe/H], and [Mg/H] from the APOGEE field. This reference set corresponds to one doppelgänger per KOI. Recent work has demonstrated that (Fe, Mg) alone capture the majority of abundance dimensionality for stars more metal-rich than [Fe/H] > -1.0 dex with surprising predictive power (Weinberg, Holtzman, Hasselquist, et al., 2019; Griffith et al., 2021; Weinberg, Holtzman, Johnson, et al., 2022; Ness et al., 2022). This is because these elements are fiducial tracers of two primary production sources, specifically core collapse supernovae and low mass stellar explosions. However, small individual abundance variations at fixed (Fe, Mg) may represent (at least in part) key additional information on stellar birth and evolutionary histories (Weinberg, Holtzman, Johnson, et al., 2022; Ting and Weinberg, 2022; Ness et al., 2022). Individual abundances are inherited from birth and can be modified as a consequence of both internal (e.g., dredge up, Souto et al. 2019) and external evolution (e.g., planet engulfment, Oh et al. 2018). Therefore,

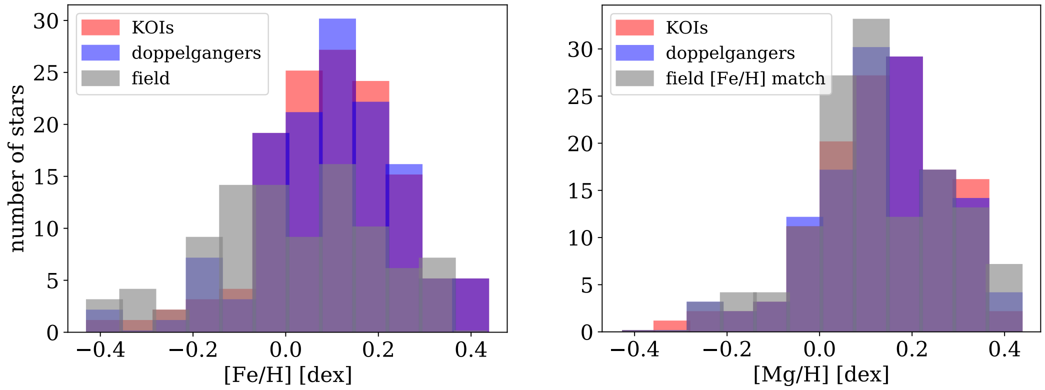


Figure 6.3: The left panel displays the [Fe/H] distributions for our KOI (red) and doppelgänger (blue) samples, as well as that of a sample of non-planet host field stars drawn from the Kepler field (gray). The right panel displays the [Mg/H] distributions for our KOI (red) and doppelgänger (blue) samples, as well as that of a sample of non-planet host field stars drawn from the Kepler field selected to have the same [Fe/H] as our KOI sample (gray).

abundance scatter in absence of (Fe, Mg) and evolutionary state contributions may encode abundance deviations from birth. Stars with planets may furthermore be born with different abundance distributions at fixed (Fe, Mg) compared to stars without.

Rather than simply comparing the individual elemental abundance distributions of our KOI and doppelgänger samples, we use a four-parameter ($T_{\rm eff}$, $\log g$, [Fe/H], [Mg/H]) model to predict the individual abundances of both the doppelgänger and KOI sets. This approach enables a quantitative exploration of the relative predictive power these four parameters hold for abundances of KOI stars compared to those of the field population. It also allows us to examine element correlations if there are clear discrepancies between the KOI and doppelgänger samples. Our model is detailed in Section 6.3. The stars we use to build our model are effectively drawn from the same underlying population as our doppelgängers in that none are confirmed/candidate planet hosts; we do not know their planet memberships. This enables us to examine how well we can predict each individual element while only considering our four predictors. We present the results of our abundance residual analysis in Section 6.4, and discuss these results in the context of potential planet host star chemistry and planet formation connections in Section 6.5.

6.2 Data

We assemble a high fidelity sample of APOGEE DR17 stars with abundance measurements for twelve elements (X = C, N, O, Na, Al, Si, Ca, Ti, V, Cr, Mn, Ni). These abundances are determined by the ASPCAP pipeline (García Pérez et al., 2016), and are reported with respect to Fe. Because we are interested in abundance patterns resulting from planet formation with respect to hydrogen rather than enhancements with respect to iron, we convert the abundances as relative to hydrogen ([X/H] = [X/Fe] + [Fe/H]). We then apply the following quality cuts which leaves a sample of ~129,000 stars (Figure 6.1):

$T_{\rm eff}$ = 4500−5500 K

$\log g > 1.8$ dex

SNR > 80

$[\mathrm{X/Fe}]_{\rm error} < 0.1$ dex

Flag `ASPCAPFLAGS` not set to `STAR_BAD`, `M_H_BAD`, `ALPHA_M_BAD`

Flag `STARFLAGS` not set to `VERY_BRIGHT_NEIGHBOR`

We then cross-match the sample with a catalog of 2098 KOIs observed by APOGEE-2N (Caleb Cañas, *private correspondence*), resulting in 220 high fidelity KOI stars from APOGEE DR17. We cross-match the resulting sample with the final Kepler planet candidate catalog data release (DR25, Coughlin, Thompson, and Kepler Team 2017) to obtain up-to-date candidate dispositions and planetary parameters. We subsequently remove all KOIs marked as 'False Positive', which indicates that the detected signals are due to events other than exoplanet transits, e.g., eclipsing binaries. This cut leaves 128 confirmed planets and 56 planet candidates hosted by 131 APOGEE DR17 stars. As expected, the KOI-APOGEE DR17 sample is dominated by "Kepler-like" architectures. That is, planets characterized as super-Earths or sub-Neptunes (e.g., Winn and Fabrycky 2015; Yang, Xie, and Zhou 2020); ~92% of the KOIs fit this category by hosting confirmed/candidate planets with orbital periods and planet radii of $P < 400$ days and $R < 4\ R_{\oplus}$, respectively (Figure 6.2).

Next, we construct a set of doppelgänger stars to our KOI-APOGEE DR17 sample. The doppelgängers are drawn from the ~129,000 high fidelity stars selected from the APOGEE field, and have unknown planet membership. We select doppelgängers by defining a similarity metric between two stars:

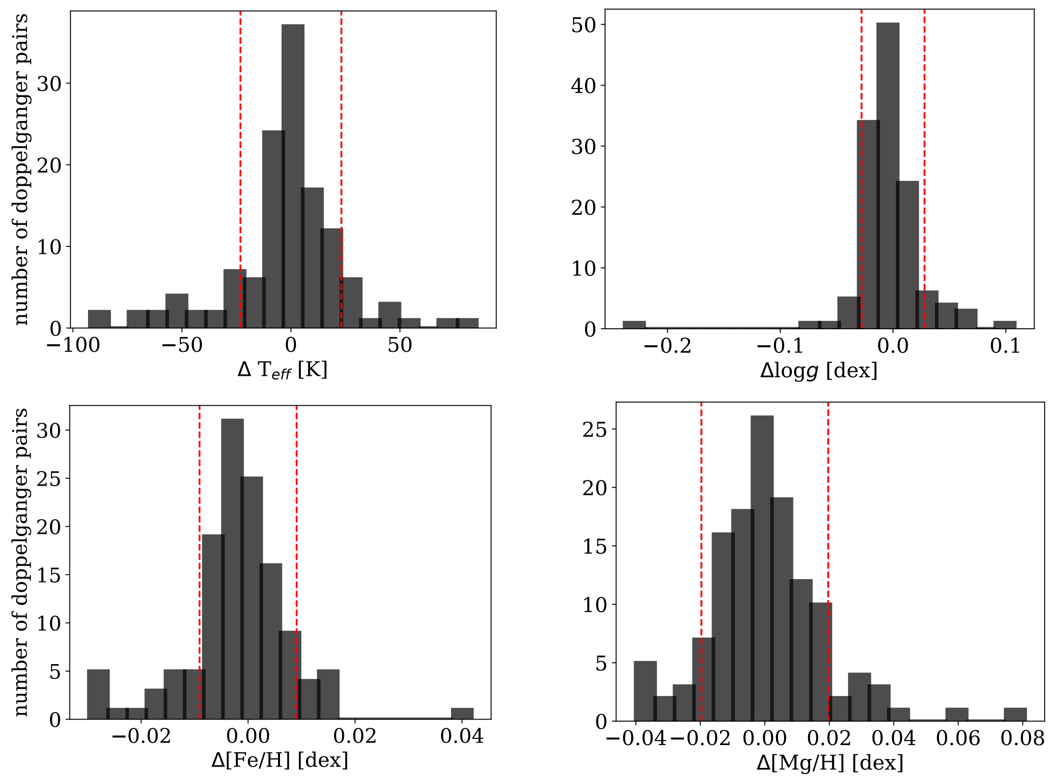


Figure 6.4: $\Delta T_{\rm eff}$ (upper left), $\Delta \log g$ (upper right), Δ[Fe/H] (lower left), and Δ[Mg/H] (lower right) distributions for each KOI-doppelgänger pair. The average associated errors added in quadrature (for KOIs and doppelgängers) for each parameter are shown in the dashed red lines, and encompass the majority of the distributions (~77%, ~85%, ~78%, and ~81% for the $\Delta T_{\rm eff}$, $\Delta \log g$, Δ[Fe/H], and Δ[Mg/H] distributions, respectively).

$$
\begin{aligned}
D^2 = \left(\frac{\Delta T_{\rm eff}}{\sqrt{\sigma^2_{T_{\rm eff},1} + \sigma^2_{T_{\rm eff},2}}}\right)^2 + \left(\frac{\Delta \log g}{\sqrt{\sigma^2_{\log g,1} + \sigma^2_{\log g,2}}}\right)^2 + \\
\left(\frac{\Delta [\rm Fe/H]}{\sqrt{\sigma^2_{[\rm Fe/H],1} + \sigma^2_{[\rm Fe/H],2}}}\right)^2 + \\
\left(\frac{\Delta [\rm Mg/H]}{\sqrt{\sigma^2_{[\rm Mg/H],1} + \sigma^2_{[\rm Mg/H],2}}}\right)^2
\end{aligned} \tag{6.1}
$$

which incorporates the relative $T_{\rm eff}$, $\log g$, [Fe/H], and [Mg/H] between the two stars, and their associated errors added in quadrature. These parameters are ideal for selecting doppelgängers because $T_{\rm eff}$ and $\log g$ describe the stellar evolutionary state, while [Fe/H] and [Mg/H] represent fiducial contributions from supernovae as mentioned earlier. Together, these parameters effectively create a four-dimensional reference frame to examine variance in individual elements. We select one doppelgänger per KOI, defined as the high fidelity non-KOI star drawn from the APOGEE field with the smallest D^2 metric value relative to that KOI, and SNR matching to within 20/pix. This SNR condition cannot be met for one KOI and any other star in the high fidelity sample, so we remove it and are left with a final KOI sample of 130 stars (Figure 6.1, colored circles).

Because we select our doppelgängers on KOI [Fe/H], there is a concern that our doppelgänger sample may have a higher rate of planet occurrence compared to a field star sample due to the known [Fe/H]-planet occurrence relation. To test this, we compare the [Fe/H] distributions of our KOI and doppelgänger samples with that of a field star sample drawn from the Kepler field that lack planets ('False Positive' candidates). The [Fe/H] distributions of these three samples all span approximately the same [Fe/H] range and roughly peak at the same [Fe/H] value (~0.1 dex) (Figure 6.3, left panel). We conclude that our doppelgänger sample is not significantly more [Fe/H]-rich that a sample of non-planet hosting field stars, and thus do not expect our doppelgängers to have a higher rate of planet occurrence compared to field stars based on [Fe/H].

Because we select our doppelgängers on KOI [Mg/H] as well, we also test if [Mg/H] correlates with planet occurrence independent of [Fe/H]. To do this, we construct another field star sample by drawing with replacement from all non-planet hosting Kepler field stars, with each draw pulling a field star with the closest [Fe/H] to that of each KOI. We then examine the [Mg/H] distributions of our KOI and doppelgänger samples, as well as that of the new KOI [Fe/H]-matching non-planet host Kepler field star sample (Figure 6.3, right panel). All three [Mg/H] distributions are quite similar, indicating that our selection on [Mg/H] does not significantly bias the doppelgänger sample relative to field stars without known planets. Thus, our doppelgänger sample does not have a higher rate of planet occurrence compared to field stars due to selection on KOI [Mg/H].

The ΔT_{eff}, $\Delta \log g$, Δ[Fe/H], and Δ[Mg/H] distributions for all KOI-doppelgänger pairs are provided in Figure 6.4. These distributions are centered on zero, which indicates that there are no systematic biases. The average associated errors added in quadrature (for KOIs and doppelgängers) for each parameter across all pairs are marked in the dashed red lines, which contain ~77%, ~85%, ~78%, and ~81% of the ΔT_{eff}, $\Delta \log g$, Δ[Fe/H], and Δ[Mg/H] distributions, respectively. Thus, the differences in parameters between KOIs and their respective doppelgängers are largely contained within their typical errors. There are 14 KOI-doppelgänger pairs that have [Fe/H] and [Mg/H] abundance differences far from the typical error boundaries (which we define as Δ[Fe/H] or Δ[Mg/H] less than −0.03 dex or greater than 0.03 dex, see Figure 6.4). We remove these outlier KOI-doppelgänger pairs and recompute the intrinsic dispersion of our abundance predictions (see Section 6.5 for a full description of our intrinsic dispersion analysis). We find that the intrinsic dispersion results do not change outside of our reported precision, and thus determine that the inclusion of these KOI-doppelgänger outlier pairs does not affect our abundance prediction results.

We constructed another doppelgänger sample also selected on K-band extinction A_k as provided by the `AK_TARG` column in APOGEE DR17. The similarity metric is modified to include the A_k term as follows:

$$\begin{aligned} D^2_{A_k} = \left(\frac{\Delta T_{\mathrm{eff}}}{\sqrt{\sigma^2_{T_{\mathrm{eff}},1} + \sigma^2_{T_{\mathrm{eff}},2}}}\right)^2 + \left(\frac{\Delta \log g}{\sqrt{\sigma^2_{\log g,1} + \sigma^2_{\log g,2}}}\right)^2 + \\ \left(\frac{\Delta[\mathrm{Fe/H}]}{\sqrt{\sigma^2_{[\mathrm{Fe/H}],1} + \sigma^2_{[\mathrm{Fe/H}],2}}}\right)^2 + \left(\frac{\Delta[\mathrm{Mg/H}]}{\sqrt{\sigma^2_{[\mathrm{Mg/H}],1} + \sigma^2_{[\mathrm{Mg/H}],2}}}\right)^2 + \\ \left(\frac{\Delta A_k}{\sqrt{\sigma^2_{A_k,1} + \sigma^2_{A_k,2}}}\right)^2. \end{aligned} \tag{6.2}$$

The K-band extinction characterizes the strength of absorption features in the optical and near-infrared wavelength range, e.g., diffuse interstellar bands (DIBs). These spectral features probe dusty regions of the interstellar medium (ISM), which is valuable from a planet formation perspective as planet occurrence is enhanced in metal-rich environments. We thus include this additional doppelgänger sample criterion for conducting a stricter test of similarity by also considering the line-of-sight ISM.

6.3 Regression Model

We construct a local linear model for each KOI to determine how well we can predict abundances from our four parameters of interest (T_{eff}, $\log g$, [Fe/H], and [Mg/H]). We note that including the evolutionary state parameters accounts for any systematic changes in abundance with T_{eff} or $\log g$. The local linear models employ simple linear regression (Hastie, Tibshirani, and Friedman, 2001), where each individual model is constructed from a training set specific to that KOI drawn from the high fidelity APOGEE DR17 sample of ∼129,000 stars. The training sets are selected by defining a region around each KOI in parameter space (e.g., Sayeed et al. 2021; Ness et al. 2022). We outline the steps of our approach below, where the parameters selected as predictors are $\vec{Y}$ = ($T_{\rm eff}$, $\log g$, [Fe/H], [Mg/H]), and the twelve predicted abundances are X = (C, N, O, Na, Al, Si, Ca, Ti, V, Cr, Mn, Ni):

1. We standardize the parameters used as predictors across the entire high fidelity sample. This is done for each star by subtracting the mean and dividing by the standard deviation: y = (Y - $\overline{Y}$)/σ_Y.

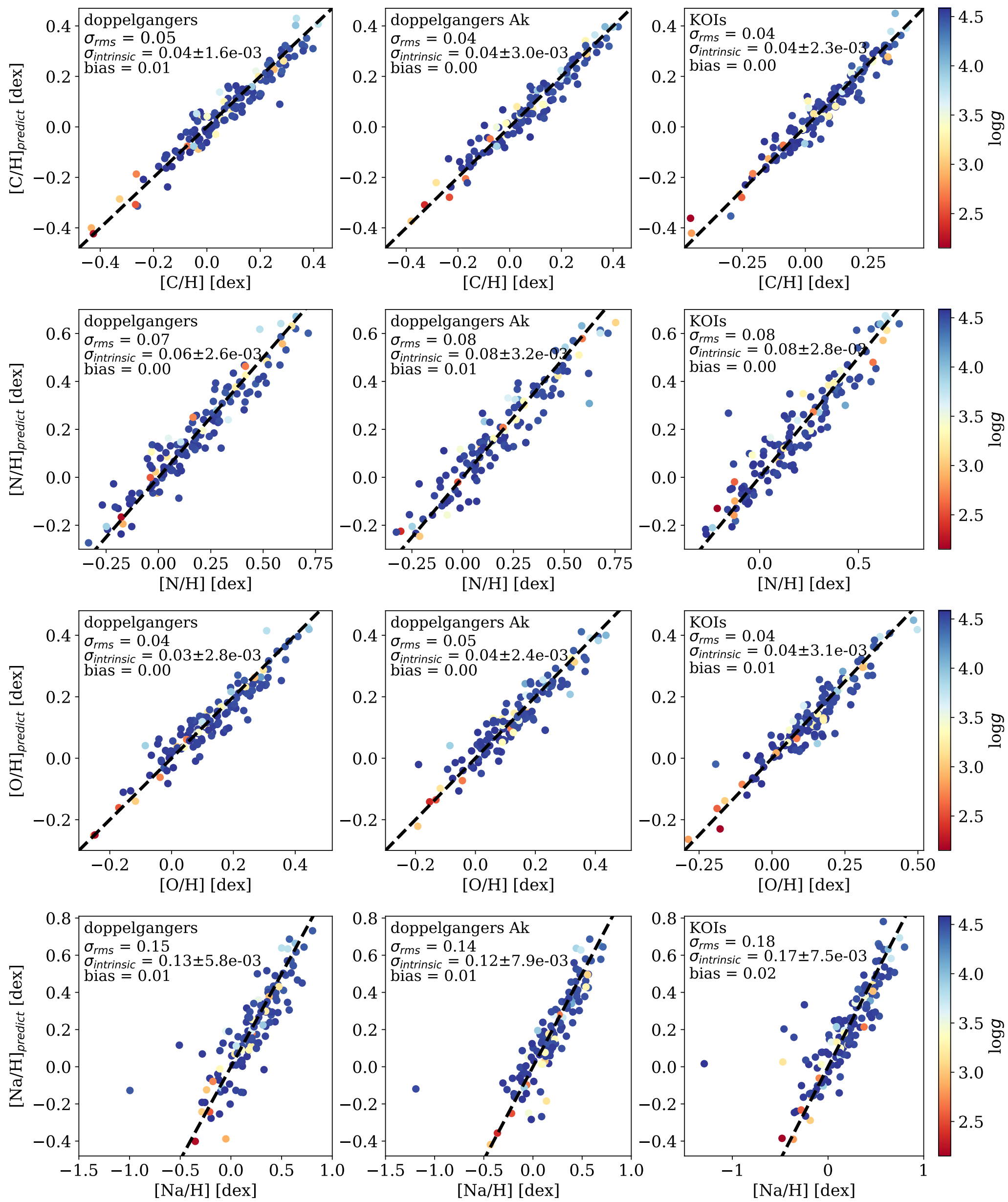
2. For each star in our high fidelity sample, we identify the nearest k neighbors in predictor parameter space according to the Euclidean distance. We carried this out with the `scikit-learn` (`sklearn`) package implemented in Python (specifically `sklearn.neighbors.KDTree`).

3. We construct a local model for each KOI from the k nearest non-KOI neighbors. We select k = 100 but note that the model appears insensitive to the choice of k; k = 50−300 produces comparable results (Ness et al., 2022). The 100 nearest neighbor training sets for each KOI include the KOI's corresponding doppelgänger, but not the KOI itself.

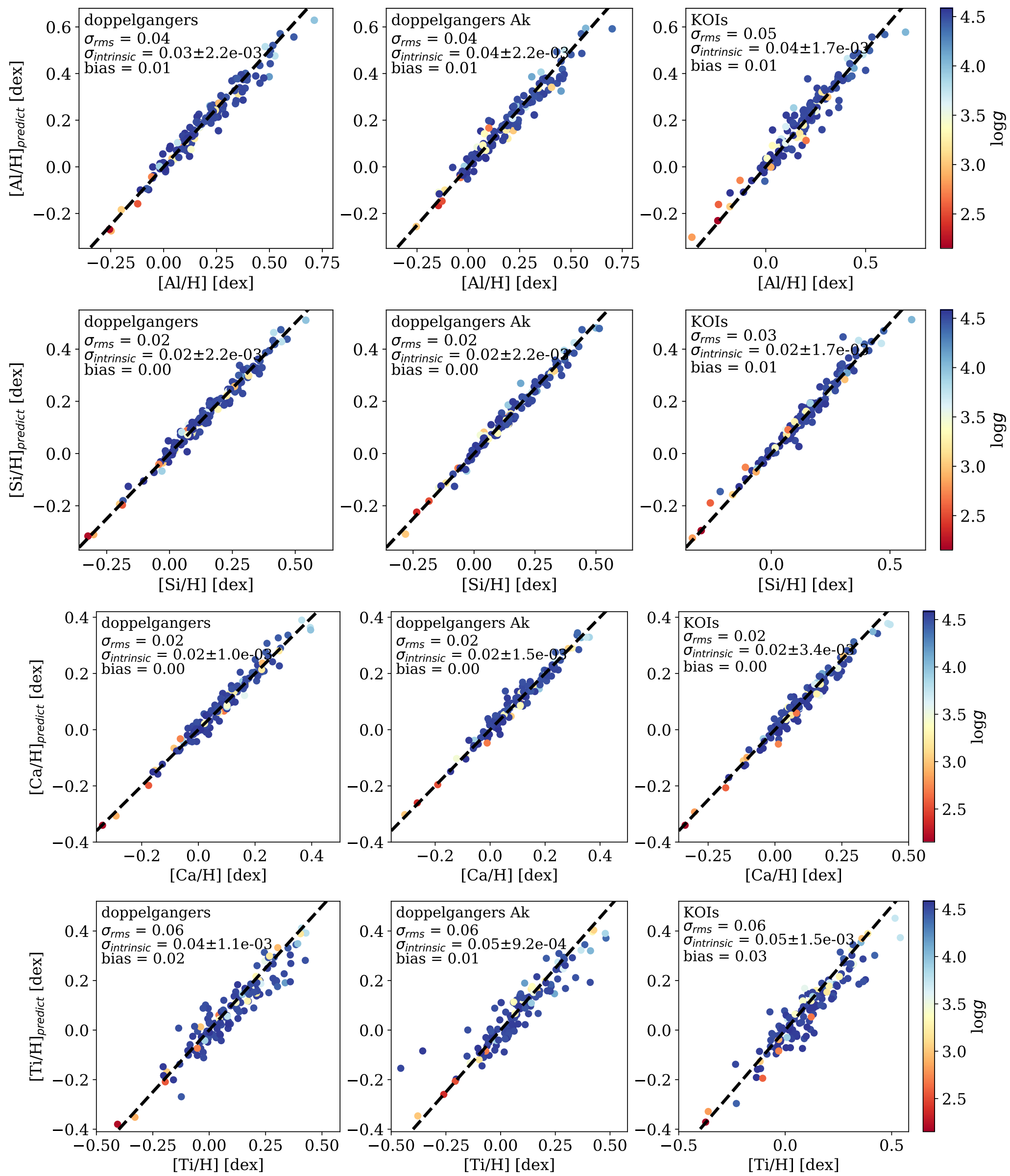
4. We use linear regression, again applied via the `sklearn` package, to train each local linear model. This modeling step elucidates the relationships between the predictor parameters $\vec{Y}$ = ($T_{\rm eff}$, $\log g$, [Fe/H], [Mg/H]), and each of the twelve abundances [X/H], separately. Each local model includes five coefficients constrained from linear regression, corresponding to the intercept and one for each predictor parameter.

5. We predict a new set of twelve abundances for each KOI from their individual local linear models. The predicted [X/H] can be compared with the measured [X/H] from ASPCAP.

6. We carry out this procedure for (i) our KOI-APOGEE DR17 sample and (ii) our corresponding doppelgänger samples. The 100 nearest neighbor training sets for each doppelgänger do not include the doppelgänger itself. The result is a set of local linear models with one model per star for every star in each sample. We subsequently use these models for abundance prediction.



The average parameter space size spanned by the 100 nearest neighbors of all KOIs and doppelgängers can be likened to the average difference in each parameter between the KOIs/doppelgängers and their nearest neighbors. The average $\Delta T_{\rm eff}$, $\Delta\log g$, Δ[Mg/H], and Δ[Fe/H] for all KOIs and their neighbors range from 11.5−190 K, 0.011−0.358 dex, 0.007−0.107 dex, and 0.013−0.111 dex, respectively. The scatter in average $\Delta T_{\rm eff}$, $\Delta\log g$, Δ[Mg/H], and Δ[Fe/H] across all KOIs are 46.4 K, 0.067 dex, 0.025 dex, and 0.023 dex, respectively. For the doppelgängers, the average differences range from 13.9−212 K, 0.012−0.348 dex, 0.007−0.128 dex, and 0.0146−0.133 dex. The scatter in average differences are 46.4 K, 0.063 dex,



0.028 dex, and 0.024 dex.

6.4 Results

Local Linear Model Predictions

Our local linear model-predicted abundances are plotted against ASPCAP abundances for the twelve considered elements in Figure 6.5. The doppelgänger sample, doppelgänger sample also selected on A_k, and KOI sample are shown in the panels from left to right. We calculated the intrinsic dispersion of the abundance predictions from the root mean square (rms) difference between the model−measurement abundances, and the average ASPCAP abundance uncertainty (which can be assumed as the same for each star): $\sigma_{\rm intrinsic} = \sqrt{\sigma^2_{\rm rms} - \sigma^2_{\rm measurement}}$. For the KOIs, the intrinsic dispersion values across all elements range from $\sigma_{\rm intrinsic}$ = 0.019−0.167 dex, with Na and Ca exhibiting the highest and lowest values, respectively. In other words, the measured Na abundances tend to deviate most significantly from the model and the measured Ca abundances the least. If we group the abundances into light (C, N, O, Na, Al, V), alpha (Si, Ca), and iron-peak (Ti, Cr, Mn, Ni) element groups, the median intrinsic dispersion values are 0.060 dex, 0.021 dex, and 0.040 dex, respectively. For the doppelgängers, the intrinsic dispersion values range from 0.017−0.128 dex, with Na also exhibiting the highest value, but Si exhibiting the lowest value. The light, alpha, and iron-peak element groups have median intrinsic dispersion values of 0.053 dex, 0.017 dex, and 0.036 dex, respectively. While the majority of our sample is dominated by "Kepler-like" super-Earth/sub-Neptune architectures, we examine if other planet architectures result in different intrinsic dispersions. We divide the KOI sample into four architecture categories, namely hot/warm Jupiters ($R > 8\ R_\oplus$ and $P < 100$ days), hot/warm sub-Saturns ($4\ R_\oplus < R < 8\ R_\oplus$ and $P < 100$ days), cold Jupiters ($R > 8\ R_\oplus$ and $P > 100$ days), and super-Earths/sub-Neptunes ($R < 4\ R_\oplus$). There are eight hot Jupiters, 17 sub-Saturns, and five cold Jupiters among the KOI sample, with the remaining classified as super-Earths/sub-Neptunes. The average intrinsic dispersions are 0.033 dex, 0.056 dex, 0.037 dex, and 0.046 dex for the hot/warm Jupiters, hot/warm sub-Saturns, cold Jupiters, and super-Earths/sub-Neptunes, respectively. Thus, the intrinsic dispersion values do not change much as a function of planet architecture, and are well-within the intrinsic dispersion range of the entire KOI sample. We conclude that architecture differences do not change the deviations between predicted and ASPCAP abundances, though larger samples of non-"Kepler-like" systems are needed to further investigate this.

We calculate the error on intrinsic dispersion $\sigma_{\rm intrinsic}$ by sampling all abundances from their error distributions 20 times, running the local linear models, then taking the scatter of the resulting $\sigma_{\rm intrinsic}$ values as the error on $\sigma_{\rm intrinsic}$. We also calculate the abundance prediction bias as the difference of the average predicted and ASPCAP abundances for each element. The bias is approximately zero for the doppelgänger and KOI samples across all twelve abundances, which indicates that our predicted abundances are unbiased. The rms difference and intrinsic dispersion measurements are approximately equal across all abundances except for Na; for the KOIs, $\sigma_{\rm intrinsic} \approx$ 0.17 dex, while for the doppelgänger and A_k doppelgänger samples, $\sigma_{\rm intrinsic} \approx 0.13$ dex and 0.12 dex, respectively. We plot the intrinsic dispersion measurements for all abundances and samples in order of their condensation temperature T_c in Figure 6.6. There is no apparent T_c trend, and the largest difference between the intrinsic dispersion values of the KOI and doppelgänger samples for [Na/H] is apparent.

The large intrinsic dispersion value for predicted [Na/H] exhibited by the KOI sample appears to be driven by five outlier stars that lie anomalously far from the 1-to-1 trend (Figure 6.5, [Na/H] KOI panel). We examine the spectra of these outlier stars near the Na spectral features used to derive [Na/H] in the ASPCAP pipeline (window centered on 16378.276 Å, Feeney, Wandelt, and Ness 2021), shown in Figure 6.7. There are no obvious differences between the Na features of the KOIs and their corresponding doppelgängers. Thus, we propose that the differences in intrinsic dispersion values between the KOIs and doppelgängers are due to poorly measured Na values rather than any real astrophysical differences between the samples. We calculate the average abundance error of our twelve considered elements for the sample of high fidelity (SNR > 200) field stars, and find that [Na/Fe] exhibits an average error of 0.060 dex, the second-to-largest among these elements after [V/Fe] (0.073 dex). For comparison, the typical average error of these twelve abundances is ~0.03 dex. We conclude that Na is generally poorly measured by the ASPCAP pipeline.

We also examine the heliocentric velocity V_{helio} as Ness et al. (2022) found abundance residual trends with V_{helio} that indicate contamination from ISM features. Our [Na/H] residuals do not display such trends, so we conclude that the intrinsic dispersion differences between KOIs and doppelgängers are not due to ISM contamination.

Higher Quality Sample

We examine if the KOI [Na/H] intrinsic dispersion remains much larger than those of the doppelgänger samples with more stringent cuts on $[X/Fe]_{\text{error}}$. To investigate this, we construct a higher quality sample with cuts on $[X/Fe]_{\text{error}} < 0.07$ dex rather than 0.1 dex. This results in a substantial sample size decrease, from 130 to 27 KOIs. The five outlier stars in predicted and ASPCAP [Na/H] abundance space are removed (Figure 6.8). The new [Na/H] intrinsic dispersion measurement for the KOIs is $\sigma_{\text{intrinsic}} \approx 0.062$ dex, compared to those of the doppelganger and A_k doppelganger samples, $\sigma_{\text{intrinsic}} \approx 0.085$ dex and 0.054 dex, respectively. Median intrinsic dispersions across all abundances are ~0.038 dex and ~0.041 dex for the KOI and doppelgänger samples, respectively. This constitutes further evidence that there is likely no systematic difference between the KOI and doppelgänger samples, and the initial differences in [Na/H] intrinsic dispersion were driven by targets with large abundance uncertainties.

Sensitivity to Training Set

We are interested in testing how sensitive our abundance predictions are to the local linear model training sets. Instead of selecting the nearest $k = 100$ neighbors in predictor parameter $\vec{Y} = (T_{\rm eff}, \log g, [{\rm Fe/H}], [{\rm Mg/H}])$ space as training sets, we took a more generous approach by using a random selection of 100 stars from our high fidelity APOGEE DR17 sample. We did this because we suspect that using even a random selection of stars to construct the local linear models will do not change the results significantly; we assume that the relationships between each element we predict for and the four parameters of our models ([Fe/H], [Mg/H], $T_{\rm eff}$, $\log g$) are fairly similar from star to star.

After using a random selection of 100 high fidelity APOGEE DR17 stars for our local linear model construction, we find that the intrinsic dispersion results change only slightly overall. The average intrinsic dispersion across all elements for the KOIs increase from 0.055 dex to 0.080 dex, at most increasing by a factor of ~2.2 for the most precisely predicted elemental abundances. The intrinsic dispersion range for the KOIs shift from $\sigma_{\rm intrinsic}$ = 0.019−0.167 dex to 0.036−0.195 dex, and from $\sigma_{\rm intrinsic}$ = 0.017−0.128 dex to 0.032−0.173 dex for the doppelgängers. The intrinsic dispersion errors also do not change within our reported precision. We conclude that local linear models provide better fits to the data, but linear models whose training sets span the full range of our parameter space also well describe how abundance labels vary with our four predictive labels. The good performance of the global model is presumably because [X/Fe] varies reasonably linearly, conditioned on the labels, over the full range of our sample. This is true for the galactic disk, but not similarly true for the galactic halo (i.e., Ness et al. 2022, Weinberg et al. *in prep.*). Thus, if there are interesting/significantly non-linear variations in the individual abundances at fixed (Fe, Mg), they must be below the APOGEE abundance measurement precision. We conclude that our results will not meaningfully change if we adopt the global model, which is reassuring. The local linear model is however preferable and generalizes well to larger parameter spaces, where the global model would presumably deteriorate as the data parameter space increases.

Condensation Temperature Trends

We further explore possible T_c patterns by fitting linear trends to the T_c-ordered abundances of each KOI and doppelgänger star, respectively. We carry this out for the model-predicted and ASPCAP-provided abundances. The distributions of the linear trend slopes are shown in Figure 6.9. Both distributions are centered on approximately zero, which indicates that there is no excess of T_c-dependent enrichment or depletion trends.

We construct similar distributions of linear trend slopes resulting from fits to T_c-ordered abundances, but only considering elements with T_c > 1000 K (Mn, Cr, Si, Ni, V, Ca, Ti, Al). Elements are refractory rather than volatile above this temperature, and populate the steepest regions of abundance versus T_c trends in patterns exhibiting refractory enhancement or depletion. (Meléndez et al., 2009; Ramírez, Meléndez, and Asplund, 2009; Bedell, Bean, et al., 2018). This is similar to the analysis presented in Nibauer et al. (2021), which assessed T_c trends of elements with T_c > 900 K. The resulting slope distributions for the predicted and ASPCAP abundances are shown in Figure 6.10. Both the ASPCAP and predicted abundance distributions for the KOI and doppelgänger samples exhibit a tail of slopes towards the right of the distribution center that appears to peak at $\sim 4\times 10^{-4}$ dex K^{-1}. A similar secondary peak was found by Nibauer et al. (2021) in their T_c slope distribution for >900 K elements from APOGEE DR16 data. This indicates that our data reproduces the T_c patterns in field stars, which is reassuring. However, we find fewer stars in the secondary peak at positive gradients compared to Nibauer et al. (2021). This is potentially explained by our different stellar samples that span different evolutionary states; Nibauer et al. (2021) examined stars across a narrow range of the main sequence whereas we study stars across the main sequence and red giant branch.

Another interesting feature in our ASPCAP abundance distributions is a small tail towards negative slopes that stretches beyond -2×10^{-4} dex K^{-1} (Figure 6.10, left panel). This tail of negative slope values is not apparent in our predicted abundance distributions (Figure 6.10, right panel), or the Nibauer et al. (2021) results. We calculate that ~10% and ~7.7% of the KOI and doppelgänger ASPCAP distributions are below -2×10^{-4} dex K^{-1}, respectively, and therefore compose the negative slope tail. The presence of this tail in the ASPCAP abundance distribution but not the predicted abundance distribution may indicate that there is abundance information not fully captured by (Fe, Mg) alone that may alter dex vs. T_c trends at the most negative slope regions. The absence of this tail in the Nibauer et al. (2021) abundances, which are comparable to ASPCAP abundances, may again be due to evolutionary state differences in our respective stellar samples.

T_c patterns can also be examined by splitting abundances into volatile and refractory groups, and fitting individual linear trends to both sets. This was done by Bedell, Bean, et al. (2018) for a sample of solar twins (see their Figure 4). Stars with enrichment trends will exhibit steeper linear fits to abundances with $T_c > 1000$ K compared to abundances with $T_c < 1000$ K, while the opposite will be true for depletion trends. We carry out this analysis for our KOIs and their doppelgängers, and provide examples of our linear trend fits in Figure 6.11. Because strong enrichment results in steeper refractory trends, the linear fits will have lower intercept values. Thus, enrichment pattern strength can be likened to the difference in volatile and refractory linear fit intercepts. We plot the distributions of these intercept differences in Figure 6.12. The distribution corresponding to ASPCAP abundances exhibits a tail towards higher intercept differences that is not present in the distribution derived from predicted abundances. We examine the ASPCAP T_c trends for the KOIs with the top five largest intercept differences, and find that they have anomalously low measured [Na/H] (ranging from -0.23 dex to -1.29 dex) that are $\gtrsim$0.2 dex below the other measured abundances. The associated errors on measured [Na/H] are large (0.074$-$0.94 dex). In addition, four out of the five KOIs with largest intercept differences overlap with the five KOIs that are outliers in predicted and ASPCAP [Na/H] space (Figure 6.5, [Na/H] KOI panel). This is further evidence that the [Na/H] intrinsic dispersion differences in the initial sample selected on $[\mathrm{X/Fe}]_{\mathrm{error}} < 0.1$ dex are the result of large abundance uncertainties. We conclude that if there are underlying differences in the individual abundance T_c trends for the KOI and doppelgänger samples at fixed evolutionary state, [Fe/H], and [Mg/H], they are marginal. To be detected, these differences must exceed the sensitivity of our predicted abundances, which is typically $\sigma_{\mathrm{intrinsic}} \approx 0.038$ dex and 0.041 dex for the KOIs and doppelgängers, respectively.

6.5 Discussion

The planet-metallicity correlation remains the only proven connection between host star chemistry and planet properties. We demonstrate that after removing the effects of evolutionary state and metallicity from two primary sources (Fe, Mg), the individual abundances of confirmed/candidate planet hosts (KOIs) and a reference doppelgänger set with unknown planet membership are indistinguishable. More specifically, we compute model-measurement abundance residuals from ASPCAP and predicted abundances using a four-parameter model ($T_{\rm eff}$, $\log g$, [Fe/H], [Mg/H]), and find that there are no differences in residual structure between the KOI and doppelgänger samples. We calculate the median intrinsic dispersion across all analyzed elements other than (Fe, Mg) to be $\sigma_{\rm intrinsic} \approx 0.038$ dex and 0.041 dex for the KOI and doppelgänger samples, respectively, which can be taken as the minimum abundance precision required for discerning individual abundance signatures related to planet formation.

Because we do not know the planet membership of our doppelgänger sample, some doppelgänger stars may be planet hosts. This is plausible because large planet discovery surveys such as the Kepler and TESS missions have revealed that planets are common. Using Kepler DR25, Hsu et al. (2019) recently calculated an upper limit occurrence rate of 0.27 planets per star for $0.5-16$ $R_{\oplus}$ planets around FGK dwarfs. Breaking occurrence rates by planet architectures reveals that the majority of these planets are small (R = $1-4$ $R_{\oplus}$) and generally classified as super-Earths and sub-Neptunes (e.g., Burke et al. 2015; Zhu et al. 2018; Bryson et al. 2021). If a significant fraction of our doppelgänger set consists of planet hosts, it makes sense that the abundance distributions of our KOI and doppelgänger samples are indistinguishable at fixed (Fe, Mg) and evolutionary state.

To reliably examine abundance differences between planet hosts and reference doppelgänger stars drawn from the field, none of the reference stars should host planets. Unfortunately, constructing a sample of doppelgänger stars that we know lack planets is difficult. This would require extensive monitoring of targets with Doppler planet search surveys to ensure that there are no RV signals indicative of planets. Carrying out such observations for an entire reference set of stars would be time and resource intensive. However, certain planet populations can be ruled out with minimal telescope time; close-in giant planets are more easily detected in RV and transit data without long cadence compared to smaller planets on longer orbits. In addition, close-in giants are intrinsically rare. RV surveys produce hot Jupiter ($P <$ 10 days) occurrence rates of $\sim$0.8$-$1.2% around solar-like stars (e.g., Mayor et al. 2011; Wright et al. 2012; Wittenmyer et al. 2020), and transit surveys yield even smaller occurrence rates of $\sim$0.4$-$0.6% (Howard et al., 2012; Fressin et al., 2013; Petigura et al., 2018; Kunimoto and Matthews, 2020). These rates are still small for warm Jupiters ($P <$ 50 days), with estimates of $\sim$1.3%. They remain small for hot and warm sub-Saturns (R = $4-8$ $R_{\oplus}$) as well, which have occurrence rate estimates of ~0.4% and ~2.3%, respectively (Howard et al., 2012). Thus, constructing a reference sample without close-in giant hosts is feasible. We hope to examine close-in giants in future studies, but this will require another planet host sample as only 18 of our KOIs host confirmed/candidate hot/warm sub-Saturn to Jupiter-sized planets according to the standard definition ($R > 4\ R_\oplus$ and $P < 100$ days).

Previous studies have found interesting abundance differences between stars that host and do not host close-in giants. For example, Meléndez et al. (2009) determined that the Sun exhibits a refractory depletion trend with T_c relative to eleven solar twins from the *Hipparcos* catalog, as well as four solar analogs with close-in giant planets. However, six other solar analogs lacking close-in giants as verified by RV monitoring show the solar depletion trend 50−70% of the time. One potential explanation for the solar pattern is sequestration of rocky material in the terrestrial planets, and late (10−25 Myr) accretion of dust-depleted gas once the solar convective zone began shrinking to its current mass fraction (~2%, Hughes, Rosner, and Weiss 2007). Another explanation is that all solar twins and most solar analogs lacking close-in giants engulfed planetary material at late times (>25 Myr), once their convective zones were thin. This scenario would produce refractory enrichment in stellar photospheres. However, it assumes that most solar-like stars are depleted in refractories (at least in the absence of events like planet engulfment), and more recent abundance studies of larger Sun-like samples show that this is not the case (e.g., Bedell, Bean, et al. 2018). Either way, the findings of Meléndez et al. (2009) suggest that close-in giant planets play a role in altering host star abundances. While their results defy a clear explanation, a larger sample of close-in giant hosts and reference stars lacking close-in giants could be leveraged to examine these trends more closely.

The KOI and doppelgänger median abundance prediction intrinsic dispersions are ~0.038 dex and ~0.041 dex, respectively. These values can be considered the upper limit of abundance precision needed to discern planet formation signatures in the elemental abundance patterns of host stars. Planet formation processes can exceed these levels in rare cases, such as the reported planet engulfment detection in the HD 240429-30 system (~0.2 dex, Oh et al. 2018). Planet hosts may also be born with different abundances compared to stars without planets. The planet-metallicity correlation indicates that this is true for at least [Fe/H]. Such primordial abundance deviations must also exceed our intrinsic dispersion levels to be detectable.

Our KOI and doppelgänger residual abundance distributions are indistinguishable, which yields two possibilities: (1) our reference doppelgänger set includes too many planet hosts, or (2) primordial or post-birth abundance patterns related to planet formation in our samples are below detectable levels. We can tackle the first possibility by focusing on more easily detectable planet architectures, namely close-in giants as discussed earlier. The second possibility could be addressed with higher-precision abundances from advances in spectral synthesis pipelines and/or line lists (e.g., Schuler, Flateau, et al. 2011; Liu et al. 2018; Bedell, Meléndez, et al. 2014), or from spectrographs with higher resolving power (e.g., Adibekyan, Sousa, et al. 2020). Many stars in our KOI and doppelgänger samples have abundance uncertainties that exceed our intrinsic dispersion values. Large uncertainties are the root cause of the particularly poorly measured Na abundances for the five outlier stars in our initial sample selected on $[X/H]_{err} < 0.1$ dex. Upgrades to the ASPCAP pipeline, such as improved line lists and advances to the spectral synthesis pipeline, may improve APOGEE abundance precisions in the years to come.

References

Abdurro'uf et al. (Apr. 2022). "The Seventeenth Data Release of the Sloan Digital Sky Surveys: Complete Release of MaNGA, MaStar, and APOGEE-2 Data". In: *The Astrophysical Journal Supplement Series* 259.2, 35, p. 35. DOI: 10.3847/1538-4365/ac4414. arXiv: 2112.02026 [astro-ph.GA].

Adibekyan, V., E. Delgado-Mena, et al. (June 2016). "ζ^2 Reticuli, its debris disk, and its lonely stellar companion ζ^1 Ret. Different T_c trends for different spectra". In: *Astronomy and Astrophysics* 591, A34, A34. DOI: 10.1051/0004-6361/201628453. arXiv: 1605.01918 [astro-ph.SR].

Adibekyan, V., S. G. Sousa, et al. (Oct. 2020). "Benchmark stars, benchmark spectrographs. Detailed spectroscopic comparison of ESPRESSO, PEPSI, and HARPS data for Gaia benchmark stars". In: *Astronomy and Astrophysics* 642, A182, A182. DOI: 10.1051/0004-6361/202038793. arXiv: 2008.08371 [astro-ph.SR].

Adibekyan, V. Zh., E. Delgado Mena, et al. (Nov. 2012). "Exploring the α-enhancement of metal-poor planet-hosting stars. The Kepler and HARPS samples". In: *Astronomy and Astrophysics* 547, A36, A36. DOI: 10.1051/0004-6361/201220167. arXiv: 1209.6272 [astro-ph.EP].

Adibekyan, V. Zh., N. C. Santos, et al. (July 2012). "Overabundance of α-elements in exoplanet-hosting stars". In: *Astronomy and Astrophysics* 543, A89, A89. DOI: 10.1051/0004-6361/201219564. arXiv: 1205.6670 [astro-ph.EP].

Alibert, Y., C. Mordasini, and W. Benz (Feb. 2011). "Extrasolar planet population synthesis. III. Formation of planets around stars of different masses". In: *Astronomy and Astrophysics* 526, A63, A63. DOI: 10.1051/0004-6361/201014760. arXiv: 1101.0513 [astro-ph.EP].

Bashi, Dolev and Shay Zucker (Aug. 2019). “Small Planets in the Galactic Context: Host Star Kinematics, Iron, and Alpha-element Enhancement”. In: *The Astronomical Journal* 158.2, 61, p. 61. DOI: 10.3847/1538-3881/ab27c9. arXiv: 1906.01982 [astro-ph.EP].

Bedell, Megan, Jacob L. Bean, et al. (Sept. 2018). “The Chemical Homogeneity of Sun-like Stars in the Solar Neighborhood”. In: *The Astrophysical Journal* 865.1, 68, p. 68. DOI: 10.3847/1538-4357/aad908. arXiv: 1802.02576 [astro-ph.SR].

Bedell, Megan, Jorge Meléndez, et al. (Nov. 2014). “Stellar Chemical Abundances: In Pursuit of the Highest Achievable Precision”. In: *The Astrophysical Journal* 795.1, 23, p. 23. DOI: 10.1088/0004-637X/795/1/23. arXiv: 1409.1230 [astro-ph.SR].

Biazzo, K. et al. (Nov. 2015). “The GAPS programme with HARPS-N at TNG. X. Differential abundances in the XO-2 planet-hosting binary”. In: *Astronomy and Astrophysics* 583, A135, A135. DOI: 10.1051/0004-6361/201526375. arXiv: 1506.01614 [astro-ph.SR].

Booth, Richard A. and James E. Owen (Apr. 2020). “Fingerprints of giant planets in the composition of solar twins”. In: *Monthly Notices of The Royal Astronomical Society* 493.4, pp. 5079–5088. DOI: 10.1093/mnras/staa578. arXiv: 2002.11135 [astro-ph.EP].

Brewer, John M. and Debra A. Fischer (May 2017). “Erratum: “C/O and Mg/Si Ratios of Stars in the Solar Neighborhood” (<A href=“http://doi.org/10.3847/0004-637x/831/1/20”>2016, ApJ, 831, 20</A>)”. In: *The Astrophysical Journal* 840.2, 121, p. 121. DOI: 10.3847/1538-4357/aa6d53. arXiv: 1608.06286 [astro-ph.SR].

Bryson, Steve et al. (Jan. 2021). “The Occurrence of Rocky Habitable-zone Planets around Solar-like Stars from Kepler Data”. In: *The Astronomical Journal* 161.1, 36, p. 36. DOI: 10.3847/1538-3881/abc418. arXiv: 2010.14812 [astro-ph.EP].

Buchhave, Lars A. and David W. Latham (Aug. 2015). “The Metallicities of Stars with and without Transiting Planets”. In: *The Astrophysical Journal* 808.2, 187, p. 187. DOI: 10.1088/0004-637X/808/2/187. arXiv: 1507.03557 [astro-ph.EP].

Buchhave, Lars A., David W. Latham, et al. (June 2012). “An abundance of small exoplanets around stars with a wide range of metallicities”. In: *Nature* 486.7403, pp. 375–377. DOI: 10.1038/nature11121.

Burke, Christopher J. et al. (Aug. 2015). “Terrestrial Planet Occurrence Rates for the Kepler GK Dwarf Sample”. In: *The Astrophysical Journal* 809.1, 8, p. 8. DOI: 10.1088/0004-637X/809/1/8. arXiv: 1506.04175 [astro-ph.EP].

Coughlin, Jeffrey, Susan E. Thompson, and Kepler Team (June 2017). “The Final Kepler Planet Candidate Catalog (DR25)”. In: *American Astronomical Society Meeting Abstracts #230*. Vol. 230. American Astronomical Society Meeting Abstracts, 102.04, p. 102.04.

Feeney, Stephen M., Benjamin D. Wandelt, and Melissa K. Ness (Mar. 2021). “SSSpaNG! stellar spectra as sparse, data-driven, non-Gaussian processes”. In: *Monthly Notices of the Royal Astronomical Society* 501.3, pp. 3258–3271. DOI: 10.1093/mnras/staa3586. arXiv: 1912.09498 [astro-ph.SR].

Fischer, Debra A. and Jeff Valenti (Apr. 2005). “The Planet-Metallicity Correlation”. In: *The Astrophysical Journal* 622.2, pp. 1102–1117. DOI: 10.1086/428383. URL: https://doi.org/10.1086/428383.

Fressin, François et al. (Apr. 2013). “The False Positive Rate of Kepler and the Occurrence of Planets”. In: *The Astrophysical Journal* 766.2, 81, p. 81. DOI: 10.1088/0004-637X/766/2/81. arXiv: 1301.0842 [astro-ph.EP].

Galarza, Jhon Yana et al. (Nov. 2021). “Evidence of Rocky Planet Engulfment in the Wide Binary System HIP 71726/HIP 71737”. In: *The Astrophysical Journal* 922.2, p. 129. DOI: 10.3847/1538-4357/ac2362. URL: https://dx.doi.org/10.3847/1538-4357/ac2362.

García Pérez, Ana E. et al. (2016). “ASPCAP: The APOGEE Stellar Parameter and Chemical Abundances Pipeline”. In: *The Astronomical Journal*.

Ghezzi, L., K. Cunha, et al. (Sept. 2010). “Stellar Parameters and Metallicities of Stars Hosting Jovian and Neptunian Mass Planets: A Possible Dependence of Planetary Mass on Metallicity”. In: *The Astrophysical Journal* 720.2, pp. 1290–1302. DOI: 10.1088/0004-637X/720/2/1290. arXiv: 1007.2681 [astro-ph.SR].

Ghezzi, Luan, Cintia F. Martinez, et al. (Oct. 2021). “A Spectroscopic Analysis of the California-Kepler Survey Sample. II. Correlations of Stellar Metallicities with Planetary Architectures”. In: *The Astrophysical Journal* 920.1, 19, p. 19. DOI: 10.3847/1538-4357/ac14c3. arXiv: 2107.04153 [astro-ph.EP].

Ghezzi, Luan, Benjamin T. Montet, and John Asher Johnson (June 2018). “Retired A Stars Revisited: An Updated Giant Planet Occurrence Rate as a Function of Stellar Metallicity and Mass”. In: *The Astrophysical Journal* 860.2, 109, p. 109. DOI: 10.3847/1538-4357/aac37c. arXiv: 1804.09082 [astro-ph.SR].

Gonzalez, Guillermo (Feb. 1997). “The stellar metallicity-giant planet connection”. In: *Monthly Notices of The royal* 285.2, pp. 403–412. DOI: 10.1093/mnras/285.2.403.

Griffith, Emily et al. (Mar. 2021). “The Similarity of Abundance Ratio Trends and Nucleosynthetic Patterns in the Milky Way Disk and Bulge”. In: *The Astrophysical Journal* 909.1, 77, p. 77. DOI: 10.3847/1538-4357/abd6be. arXiv: 2009.05063 [astro-ph.SR].

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (2001). *The Elements of Statistical Learning*. Springer Series in Statistics. New York, NY, USA: Springer New York Inc.

Heiter, U. and R. E. Luck (Oct. 2003). "Abundance Analysis of Planetary Host Stars. I. Differential Iron Abundances". In: *The Astronomical Journal* 126.4, pp. 2015–2036. DOI: 10.1086/378366. arXiv: astro-ph/0307321 [astro-ph].

Howard, Andrew W. et al. (Aug. 2012). "Planet Occurrence within 0.25 AU of Solar-type Stars from Kepler". In: *The Astrophysical Journal Supplement Series* 201.2, 15, p. 15. DOI: 10.1088/0067-0049/201/2/15. arXiv: 1103.2541 [astro-ph.EP].

Hsu, Danley C. et al. (Sept. 2019). "Occurrence Rates of Planets Orbiting FGK Stars: Combining Kepler DR25, Gaia DR2, and Bayesian Inference". In: *The Astronomical Journal* 158.3, 109, p. 109. DOI: 10.3847/1538-3881/ab31ab. arXiv: 1902.01417 [astro-ph.EP].

Hughes, D. W., R. Rosner, and N. O. Weiss (2007). *The Solar Tachocline*.

Ida, Shigeru and D. N. C. Lin (Nov. 2004). "Toward a Deterministic Model of Planetary Formation. II. The Formation and Retention of Gas Giant Planets around Stars with a Range of Metallicities". In: *The Astrophysical Journal* 616.1, pp. 567–572. DOI: 10.1086/424830. URL: https://doi.org/10.1086/424830.

Jofré, Emiliano et al. (Dec. 2021). "The Peculiar Chemical Pattern of the WASP-160 Binary System: Signatures of Planetary Formation and Evolution?" In: *The Astronomical Journal* 162.6, 291, p. 291. DOI: 10.3847/1538-3881/ac25ef. arXiv: 2109.04590 [astro-ph.EP].

Kunimoto, Michelle and Jaymie M. Matthews (June 2020). "Searching the Entirety of Kepler Data. II. Occurrence Rate Estimates for FGK Stars". In: *The Astronomical Journal* 159.6, 248, p. 248. DOI: 10.3847/1538-3881/ab88b0. arXiv: 2004.05296 [astro-ph.EP].

Liu, F. et al. (July 2018). "Detailed chemical compositions of the wide binary HD 80606/80607: revised stellar properties and constraints on planet formation". In: *Astronomy and Astrophysics* 614, A138, A138. DOI: 10.1051/0004-6361/201832701. arXiv: 1802.09306 [astro-ph.SR].

Mack Claude E., III et al. (June 2014). "Detailed Abundances of Planet-hosting Wide Binaries. I. Did Planet Formation Imprint Chemical Signatures in the Atmospheres of HD 20782/81?" In: *The Astrophysical Journal* 787.2, 98, p. 98. DOI: 10.1088/0004-637X/787/2/98. arXiv: 1404.1967 [astro-ph.EP].

Maldonado, J., E. Villaver, and C. Eiroa (May 2018). "Chemical fingerprints of hot Jupiter planet formation". In: *Astronomy and Astrophysics* 612, A93, A93. DOI: 10.1051/0004-6361/201732001. arXiv: 1712.01035 [astro-ph.EP].

Maldonado, J., E. Villaver, C. Eiroa, and G. Micela (Apr. 2019). “Connecting substellar and stellar formation: the role of the host star’s metallicity”. In: *Astronomy and Astrophysics* 624, A94, A94. DOI: 10.1051/0004-6361/201833827. arXiv: 1903.01141 [astro-ph.SR].

Mayor, M. et al. (Sept. 2011). “The HARPS search for southern extra-solar planets XXXIV. Occurrence, mass distribution and orbital properties of super-Earths and Neptune-mass planets”. In: *arXiv e-prints*, arXiv:1109.2497, arXiv:1109.2497. arXiv: 1109.2497 [astro-ph.EP].

Meléndez, J. et al. (Oct. 2009). “The Peculiar Solar Composition and Its Possible Relation to Planet Formation”. In: *The Astrophysical Journal Letters* 704.1, pp. L66–L70. DOI: 10.1088/0004-637X/704/1/L66. arXiv: 0909.2299 [astro-ph.SR].

Mordasini, C. et al. (Nov. 2012). “Characterization of exoplanets from their formation. II. The planetary mass-radius relationship”. In: *Astronomy and Astrophysics* 547, A112, A112. DOI: 10.1051/0004-6361/201118464. arXiv: 1206.3303 [astro-ph.EP].

Mulders, Gijs D. et al. (Dec. 2016). “A Super-solar Metallicity for Stars with Hot Rocky Exoplanets”. In: *The Astronomical Journal* 152.6, 187, p. 187. DOI: 10.3847/0004-6256/152/6/187. arXiv: 1609.05898 [astro-ph.EP].

Nagar, Tushar, Lorenzo Spina, and Amanda I. Karakas (Jan. 2020). “The Chemical Signatures of Planetary Engulfment Events in Binary Systems”. In: *The Astrophysical Journal Letters* 888.1, L9, p. L9. DOI: 10.3847/2041-8213/ab5dc6.

Narang, Mayank et al. (Nov. 2018). “Properties and Occurrence Rates for Kepler Exoplanet Candidates as a Function of Host Star Metallicity from the DR25 Catalog”. In: *The Astronomical Journal* 156.5, 221, p. 221. DOI: 10.3847/1538-3881/aae391. arXiv: 1809.08385 [astro-ph.EP].

Ness, Melissa K. et al. (Feb. 2022). “The Homogeneity of the Star-forming Environment of the Milky Way Disk over Time”. In: *The Astrophysical Journal* 926.2, 144, p. 144. DOI: 10.3847/1538-4357/ac4754. arXiv: 2109.05722 [astro-ph.GA].

Nibauer, Jacob et al. (Feb. 2021). “Statistics of the Chemical Composition of Solar Analog Stars and Links to Planet Formation”. In: *The Astrophysical Journal* 907.2, 116, p. 116. DOI: 10.3847/1538-4357/abd0f1. arXiv: 2010.07241 [astro-ph.SR].

Öberg, Karin I., Ruth Murray-Clay, and Edwin A. Bergin (Dec. 2011). “The Effects of Snowlines on C/O in Planetary Atmospheres”. In: *The Astrophysical Journal Letters* 743.1, L16, p. L16. DOI: 10.1088/2041-8205/743/1/L16. arXiv: 1110.5567 [astro-ph.GA].

Oh, Semyeong et al. (Feb. 2018). “Kronos and Krios: Evidence for Accretion of a Massive, Rocky Planetary System in a Comoving Pair of Solar-type Stars”. In: *The Astrophysical Journal* 854.2, 138, p. 138. DOI: 10.3847/1538-4357/aaab4d. arXiv: 1709.05344 [astro-ph.SR].

Petigura, Erik A. et al. (Feb. 2018). “The California-Kepler Survey. IV. Metal-rich Stars Host a Greater Diversity of Planets”. In: *The Astronomical Journal* 155.2, 89, p. 89. DOI: 10.3847/1538-3881/aaa54c. arXiv: 1712.04042 [astro-ph.EP].

Ramírez, I., S. Khanal, P. Aleo, et al. (July 2015). “The Dissimilar Chemical Composition of the Planet-hosting Stars of the XO-2 Binary System”. In: *The Astrophysical Journal* 808, 13, p. 13. DOI: 10.1088/0004-637X/808/1/13. arXiv: 1506.01025 [astro-ph.SR].

Ramírez, I., S. Khanal, S. J. Lichon, et al. (Dec. 2019). “The chemical composition of HIP 34407/HIP 34426 and other twin-star comoving pairs”. In: *Monthly Notices of The Royal Astronomical Society* 490.2, pp. 2448–2457. DOI: 10.1093/mnras/stz2709. arXiv: 1909.07460 [astro-ph.SR].

Ramírez, I., J. Meléndez, and M. Asplund (Dec. 2009). “Accurate abundance patterns of solar twins and analogs. Does the anomalous solar chemical composition come from planet formation?” In: *Astronomy and Astrophysics* 508.1, pp. L17–L20. DOI: 10.1051/0004-6361/200913038.

– (Jan. 2014). “Chemical signatures of planets: beyond solar-twins”. In: *Astronomy and Astrophysics* 561, A7, A7. DOI: 10.1051/0004-6361/201322558. arXiv: 1310.8581 [astro-ph.SR].

Ramírez, I., J. Meléndez, D. Cornejo, et al. (Oct. 2011). “Elemental Abundance Differences in the 16 Cygni Binary System: A Signature of Gas Giant Planet Formation?” In: *The Astrophysical Journal* 740.2, 76, p. 76. DOI: 10.1088/0004-637X/740/2/76. arXiv: 1107.5814 [astro-ph.SR].

Rice, W. K. M. and Philip J. Armitage (Oct. 2003). “On the Formation Timescale and Core Masses of Gas Giant Planets”. In: *The Astrophysical Journal* 598.1, pp. L55–L58. DOI: 10.1086/380390. URL: https://doi.org/10.1086/380390.

Saffe, C., M. Flores, et al. (Apr. 2016). “Temperature condensation trend in the debris-disk binary system ζ^2 Reticuli”. In: *Astronomy and Astrophysics* 588, A81, A81. DOI: 10.1051/0004-6361/201528043. arXiv: 1602.01320 [astro-ph.SR].

Saffe, C., E. Jofré, et al. (July 2017). “Signatures of rocky planet engulfment in HAT-P-4. Implications for chemical tagging studies”. In: *Astronomy and Astrophysics* 604, L4, p. L4. DOI: 10.1051/0004-6361/201731430. arXiv: 1707.02180 [astro-ph.SR].

Santos, N. C., G. Israelian, and M. Mayor (Mar. 2004). "Spectroscopic [Fe/H] for 98 extra-solar planet-host stars. Exploring the probability of planet formation". In: *Astronomy and Astrophysics* 415, pp. 1153–1166. doi: 10.1051/0004-6361:20034469. arXiv: astro-ph/0311541 [astro-ph].

Sayeed, Maryum et al. (Apr. 2021). "The Swan: Data-driven Inference of Stellar Surface Gravities for Cool Stars from Photometric Light Curves". In: *The Astronomical Journal* 161.4, 170, p. 170. doi: 10.3847/1538-3881/abdf4c. arXiv: 2011.10062 [astro-ph.SR].

Schlaufman, Kevin C. and Gregory Laughlin (Sept. 2011). "Kepler Exoplanet Candidate Host Stars Are Preferentially Metal Rich". In: *The Astrophysical Journal* 738.2, 177, p. 177. doi: 10.1088/0004-637X/738/2/177. arXiv: 1106.6043 [astro-ph.EP].

Schuler, Simon C., Davin Flateau, et al. (May 2011). "Abundances of Stars with Planets: Trends with Condensation Temperature". In: *The Astrophysical Journal* 732.1, 55, p. 55. doi: 10.1088/0004-637X/732/1/55. arXiv: 1103.0757 [astro-ph.SR].

Schuler, Simon C., Zachary A. Vaz, et al. (Dec. 2015). "Detailed Abundances of Stars with Small Planets Discovered by Kepler. I. The First Sample". In: *The Astrophysical Journal* 815.1, 5, p. 5. doi: 10.1088/0004-637X/815/1/5. arXiv: 1511.00934 [astro-ph.SR].

Sousa, S. G., N. C. Santos, et al. (Aug. 2008). "Spectroscopic parameters for 451 stars in the HARPS GTO planet search program. Stellar [Fe/H] and the frequency of exo-Neptunes". In: *Astronomy and Astrophysics* 487.1, pp. 373–381. doi: 10.1051/0004-6361:200809698. arXiv: 0805.4826 [astro-ph].

Sousa, Sérgio G., Vardan Adibekyan, et al. (May 2019). "The metallicity-period-mass diagram of low-mass exoplanets". In: *Monthly Notices of the Royal Astronomical Society* 485.3, pp. 3981–3990. doi: 10.1093/mnras/stz664. arXiv: 1903.04937 [astro-ph.EP].

Souto, Diogo et al. (Mar. 2019). "Chemical Abundances of Main-sequence, Turnoff, Subgiant, and Red Giant Stars from APOGEE Spectra. II. Atomic Diffusion in M67 Stars". In: *The Astrophysical Journal* 874.1, 97, p. 97. doi: 10.3847/1538-4357/ab0b43. arXiv: 1902.10199 [astro-ph.SR].

Teske, Johanna K., Luan Ghezzi, et al. (Mar. 2015). "Abundance Differences between Exoplanet Binary Host Stars XO-2N and XO-2S−Dependence on Stellar Parameters". In: *The Astrophysical Journal Letters* 801.1, L10, p. L10. doi: 10.1088/2041-8205/801/1/L10. arXiv: 1501.02167 [astro-ph.EP].

Teske, Johanna K., Sandhya Khanal, and Ivan Ramírez (Mar. 2016). "The Curious Case of Elemental Abundance Differences in the Dual Hot Jupiter Hosts WASP-94A and B". In: *The Astrophysical Journal* 819.1, 19, p. 19. doi: 10.3847/0004-637X/819/1/19. arXiv: 1601.01731 [astro-ph.SR].

Ting, Yuan-Sen and David H. Weinberg (Mar. 2022). “How Many Elements Matter?” In: *The Astrophysical Journal* 927.2, 209, p. 209. DOI: 10.3847/1538-4357/ac5023. arXiv: 2102.04992 [astro-ph.GA].

Tucci Maia, M., J. Meléndez, D. Lorenzo-Oliveira, et al. (Aug. 2019). “Revisiting the 16 Cygni planet host at unprecedented precision and exploring automated tools for precise abundances”. In: *Astronomy and Astrophysics* 628, A126, A126. DOI: 10.1051/0004-6361/201935952. arXiv: 1906.04195 [astro-ph.SR].

Tucci Maia, M., J. Meléndez, and I. Ramírez (Aug. 2014). “High Precision Abundances in the 16 Cyg Binary System: A Signature of the Rocky Core in the Giant Planet”. In: *The Astrophysical Journal Letters* 790, L25, p. L25. DOI: 10.1088/2041-8205/790/2/L25. arXiv: 1407.4132 [astro-ph.SR].

Wang, Ji et al. (June 2015). “Influence of Stellar Multiplicity On Planet Formation. III. Adaptive Optics Imaging of Kepler Stars With Gas Giant Planets”. In: *The Astrophysical Journal* 806.2, 248, p. 248. DOI: 10.1088/0004-637X/806/2/248. arXiv: 1505.05363 [astro-ph.EP].

Weinberg, David H., Jon A. Holtzman, Sten Hasselquist, et al. (Mar. 2019). “Chemical Cartography with APOGEE: Multi-element Abundance Ratios”. In: *The Astrophysical Journal* 874.1, 102, p. 102. DOI: 10.3847/1538-4357/ab07c7. arXiv: 1810.12325 [astro-ph.GA].

Weinberg, David H., Jon A. Holtzman, Jennifer A. Johnson, et al. (June 2022). “Chemical Cartography with APOGEE: Mapping Disk Populations with a 2-process Model and Residual Abundances”. In: *The Astrophysical Journal Supplement Series* 260.2, 32, p. 32. DOI: 10.3847/1538-4365/ac6028. arXiv: 2108.08860 [astro-ph.GA].

Wilson, Robert F. et al. (Feb. 2018). “Elemental Abundances of Kepler Objects of Interest in APOGEE. I. Two Distinct Orbital Period Regimes Inferred from Host Star Iron Abundances”. In: *The Astronomical Journal* 155.2, 68, p. 68. DOI: 10.3847/1538-3881/aa9f27. arXiv: 1712.01198 [astro-ph.EP].

Winn, Joshua N. and Daniel C. Fabrycky (Aug. 2015). “The Occurrence and Architecture of Exoplanetary Systems”. In: *Annual Review of Astronomy and Astrophysics* 53, pp. 409–447. DOI: 10.1146/annurev-astro-082214-122246. arXiv: 1410.4199 [astro-ph.EP].

Wittenmyer, Robert A. et al. (Feb. 2020). “Cool Jupiters greatly outnumber their toasty siblings: occurrence rates from the Anglo-Australian Planet Search”. In: *Monthly Notices of the Royal Astronomical Society* 492.1, pp. 377–383. DOI: 10.1093/mnras/stz3436. arXiv: 1912.01821 [astro-ph.EP].

Wright, J. T. et al. (July 2012). “The Frequency of Hot Jupiters Orbiting nearby Solar-type Stars”. In: *The Astrophysical Journal* 753.2, 160, p. 160. DOI: 10.1088/0004-637X/753/2/160. arXiv: 1205.2273 [astro-ph.EP].

Yang, Jia-Yi, Ji-Wei Xie, and Ji-Lin Zhou (Apr. 2020). "Occurrence and Architecture of Kepler Planetary Systems as Functions of Stellar Mass and Effective Temperature". In: *The Astronomical Journal* 159.4, 164, p. 164. DOI: 10.3847/1538-3881/ab7373. arXiv: 2002.02840 [astro-ph.EP].

Zhu, Wei et al. (June 2018). "About 30% of Sun-like Stars Have Kepler-like Planetary Systems: A Study of Their Intrinsic Architecture". In: *The Astrophysical Journal* 860.2, 101, p. 101. DOI: 10.3847/1538-4357/aac6d5. arXiv: 1802.09526 [astro-ph.EP].

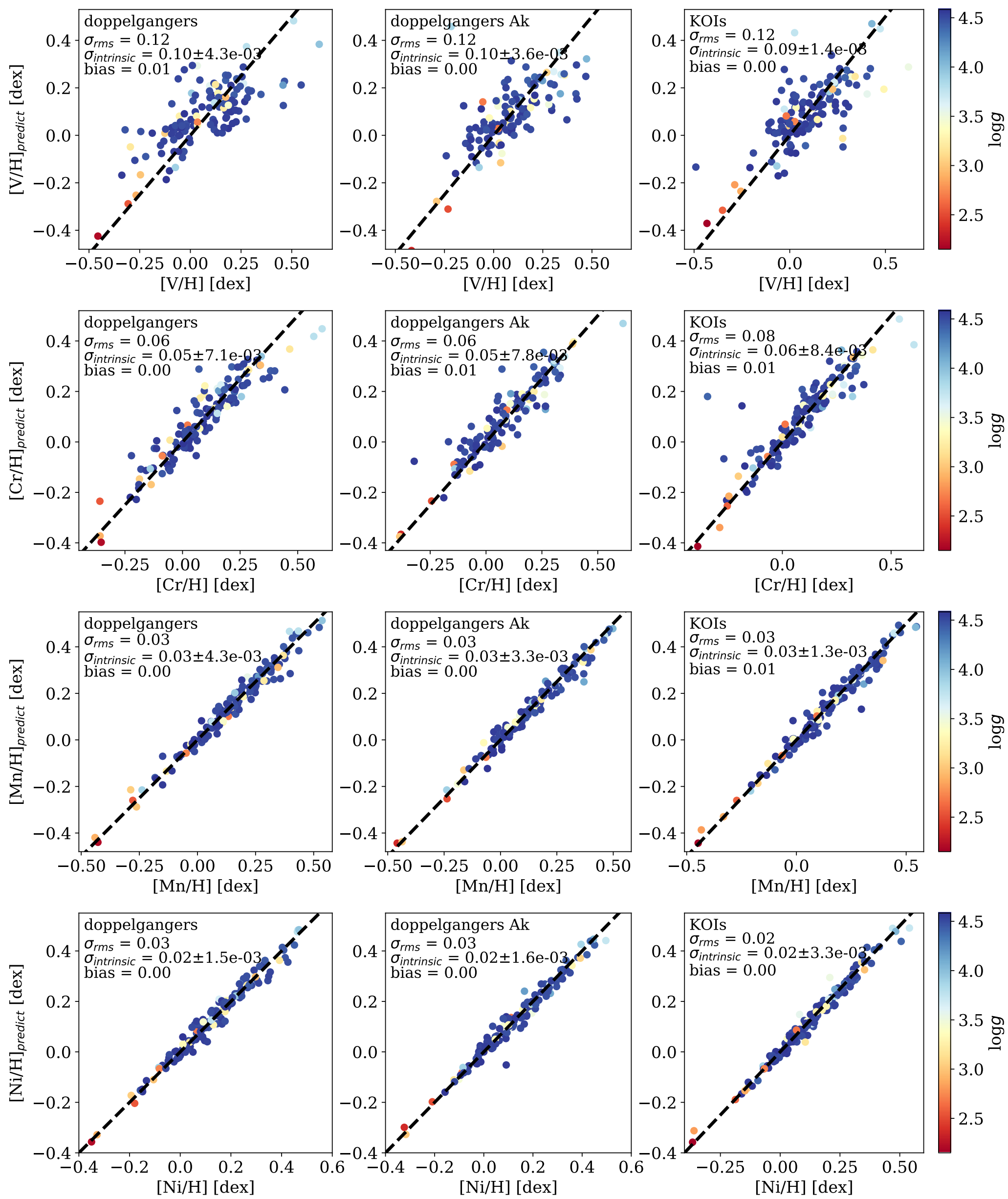


Figure 6.5: Local linear model-predicted vs. ASPCAP abundances of the doppelgänger (left), doppelgänger A_k (middle), and KOI (right) samples for the twelve elements considered. The points are colored by log$g$. The rms difference between the ASPCAP and predicted abundances, intrinsic dispersions, and bias measurements are provided in the upper left corners of each panel. A dashed 1-to-1 line is plotted in all panels for comparison.

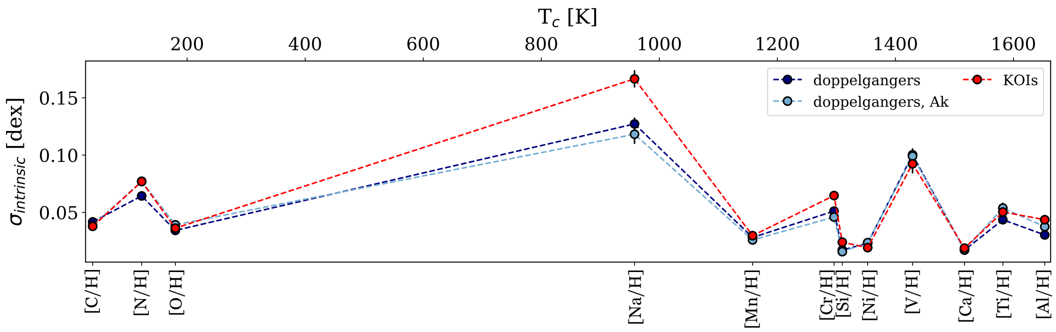


Figure 6.6: Intrinsic dispersion measurements $\sigma_{\text{intrinsic}}$ for all T_c ordered abundances, for the KOI sample (red), doppelgänger sample (navy), and A_k doppelgänger sample (light blue). There are no apparent trends with T_c, but the KOI [Na/H] prediction intrinsic dispersion is noticeably larger than those of the doppelgänger samples.

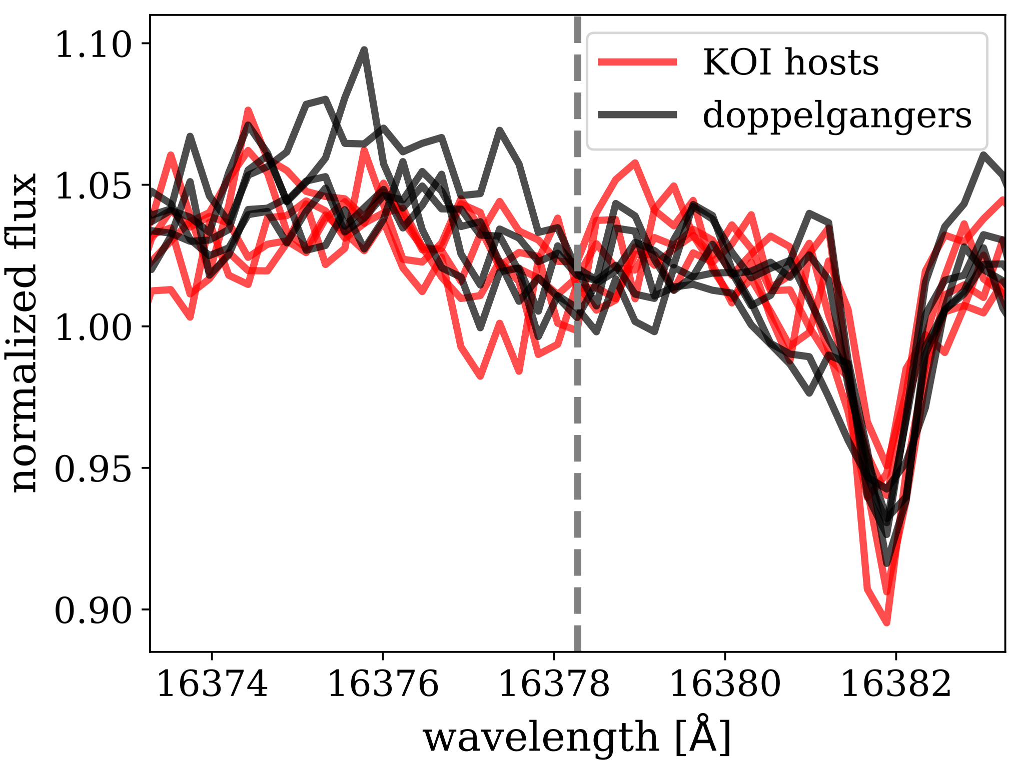


Figure 6.7: Spectra of the five KOIs that appear to be outliers in predicted and ASPCAP [Na/H] space (red), zoomed into the region with Na spectral features used to derive [Na/H] abundances with the ASPCAP pipeline (window center at 16378.276 Å, Feeney, Wandelt, and Ness 2021, marked by the dashed line). The spectra of the corresponding outlier doppelgängers are shown in black.

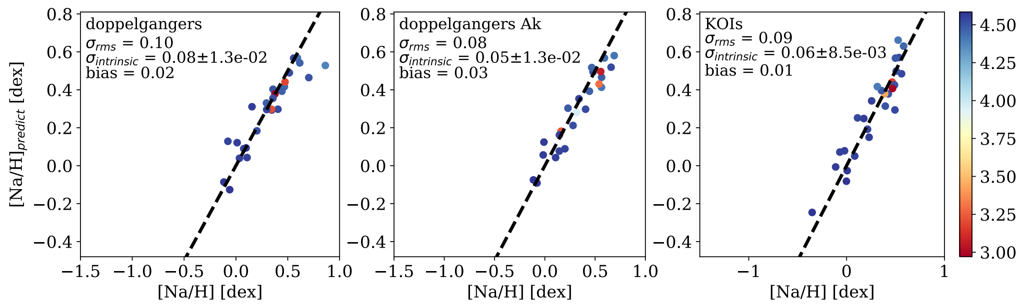


Figure 6.8: Local linear model-predicted vs. ASPCAP abundances of the higher quality ([X/H]$_{err}$ < 0.07 dex) doppelgänger (left), doppelgänger A_k (middle), and KOI (right) samples for [Na/H]. The points are colored by $\log g$. The rms difference between the ASPCAP and predicted abundances, intrinsic dispersion, and bias measurements are provided in the upper left corners of each panel. A dashed 1-to-1 line is plotted in all panels for comparison.

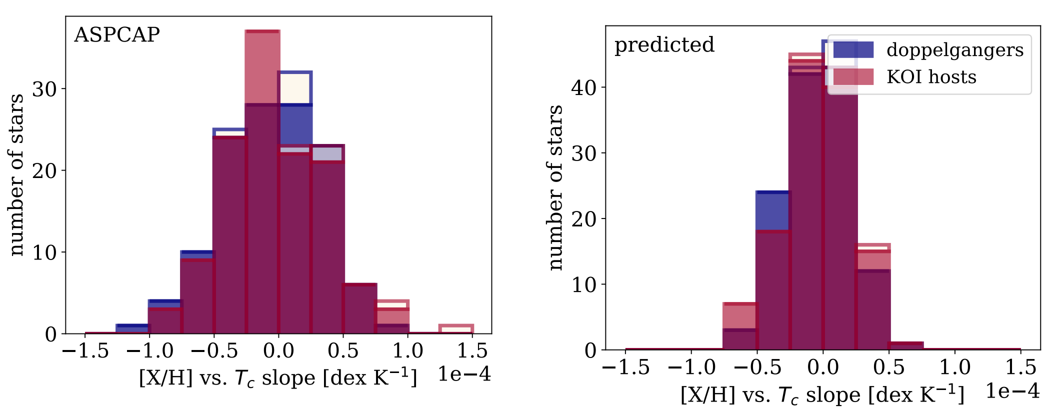


Figure 6.9: Linear trend slope distributions for the T_c-ordered abundances of individual stars in the KOI (red) and doppelgänger (blue) samples. The distributions for the predicted and ASPCAP abundances are provided in the right and left panels, respectively. The slope contributions from our five Na outlier stars are represented by the transparent regions of the distribution.

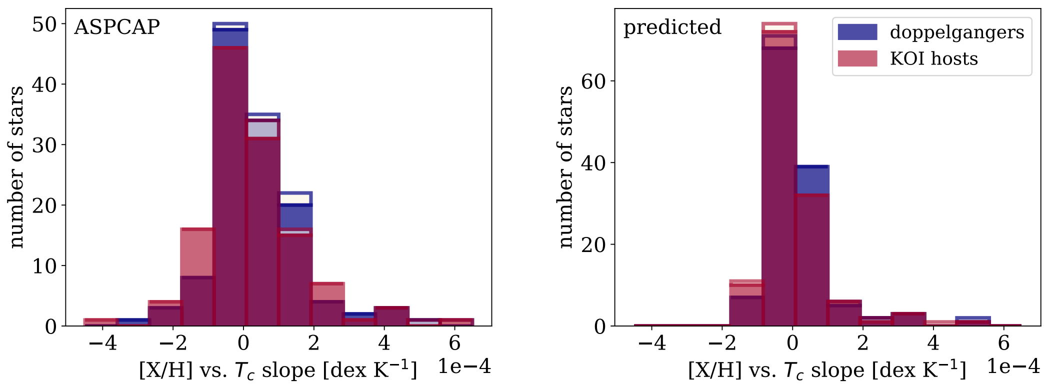


Figure 6.10: Linear trend slope distributions for the T_c-ordered abundances of individual stars in the KOI (red) and doppelgänger (blue) samples, for only refractory elements ($T_c > 1000$ K). The distributions for the predicted and ASPCAP abundances are provided in the right and left panels, respectively. The slope contributions from our five Na outlier stars are represented by the transparent regions of the distribution.

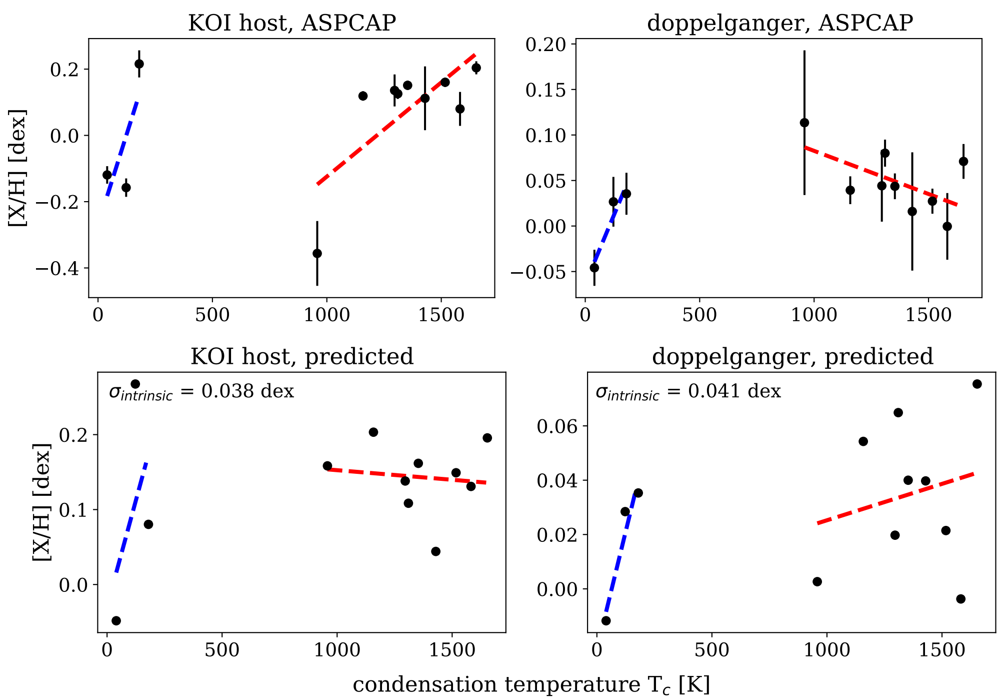


Figure 6.11: Examples of separate linear trends fitted to volatile (blue) and refractory (red) abundances ordered by T_c for our KOIs and corresponding doppelgängers. The boundary between volatile and refractory elements is set to T_c = 1000 K. Linear fits to the ASPCAP and predicted abundances are provided in the upper and lower panels, respectively. There are no individual errors associated with the predicted abundances, but the median intrinsic dispersion values for the KOI and doppelgänger samples across all abundances are shown in the upper left corners of the predicted abundance panels.

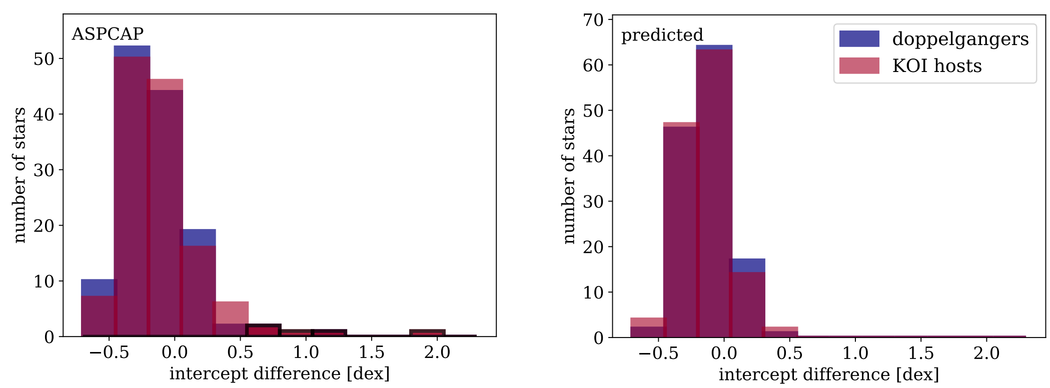


Figure 6.12: Intercept difference (refractory−volatile) distributions from linear trend fits to the T_c-ordered refractory and volatile abundances of individual stars in the KOI (red) and doppelgänger (blue) samples. The distributions for the predicted and ASPCAP abundances are provided in the right and left panels, respectively. The contributions corresponding to the KOIs with the top five largest intercept differences are outlined in black in the KOI ASPCAP distribution. Four out of these five KOIs overlap with the five KOIs that are outliers in predicted and ASPCAP [Na/H] space.

Chapter 7

CONCLUSION

7.1 Summary

This thesis examines the intricate connections between the properties of stars and those of their hosted planets. The projects therein employ a combination of datasets from single Keck-HIRES observing programs and large stellar surveys (e.g., *Gaia*, APOGEE), as well as Bayesian probabilistic models and stellar evolution codes (i.e., `MESA`).

In Chapter 2, we adapted a machine learning (ML) framework (*The Cannon*, Ness, Hogg, et al. 2015) for application to cool (T_{eff} < 4700 K) stars. ML methods such as *The Cannon* are valuable alternatives for characterizing stars outside the realm of solar because they are not dependent on physical models (e.g., in Behmard, Petigura, and Howard (2019), the training set abundances were measured with empirical methods). Instead, *The Cannon* leverages training sets of stellar spectra to construct data-driven models that can be applied to stars of the same type. We tested our implementation with a sample of K and M dwarfs that have well-measured iron abundances from empirical methods (e.g., analysis of individual spectral features, Mann, Brewer, et al. 2013; Mann, Deacon, et al. 2014), and achieved iron abundance precisions of 0.08 dex (Behmard, Petigura, and Howard, 2019). Iron is a proxy for the amount of protoplanetary disk solids that can contribute to planet cores (Gonzalez, 1997), so well-constrained iron abundances are essential for vetting planet formation theories. Thus, our adaptation of *The Cannon* is a valuable tool for examining planet formation in cool star systems.

Approximately half of planet host stars harbor stellar companions (e.g., Howell et al. 2021) Thus, characterizing planetary systems with companions is essential for fully understanding planet formation. We addressed this in Chapter 3 by compiling a sample of companions to TESS planet host stars. We identified these companions with a Bayesian probabilistic model (Oh et al., 2017) that leverages Gaia astrometric data to pinpoint comoving stellar pairs. We then constrained the angle of alignment between the companion and host star orbits, and uncovered an excess of misaligned close-in giant planet systems (Behmard, Dai, and Howard, 2022). This indicates that the dynamical histories of many close-in giants may involve axial tilt ("obliquity") excitation mechanisms via e.g., planet-planet or planet-disk interactions (Millholland and Batygin, 2019).

In Chapter 4, we led an observing program to obtain Keck-HIRES spectroscopy for 36 solar-like planet host binary systems, and measured abundances across 15 elements. Binaries are excellent laboratories for probing planet formation because binary companions share the same molecular birth cloud, and are thus born chemically homogeneous to within 0.03−0.05 dex (e.g., Bovy 2016; Ness, Rix, et al. 2018). Planet-related processes (e.g., planet engulfment) can alter the birth compositions of each star in different ways, so binary differential abundances may bear the fingerprint of planet formation. To complement this work, in Chapter 5 we used the stellar evolution code `MESA` to model how long rocky element enhancements following planet engulfment will last. We found that stellar interior mixing weakens such engulfment signatures on short timescales ($\lesssim$1 Gyr) in solar-like stars (Behmard, Sevilla, and Fuller, 2023), which explains the lack of signatures in our binary sample (Behmard, Dai, Brewer, et al., 2023). We also found a tentative trend of increasing abundance differences between companions as a function of binary separation. This new result provides important context for investigations of binary star chemistry for unveiling the processes underlying planet formation.

In Chapter 6, we extended our abundance methodology to APOGEE DR17, which provides abundances for >650,000 stars that span the entirety of our galaxy. We used the APOGEE DR17 sample to quantify the abundance precision needed for discerning large-scale patterns related to planet formation ($\lesssim$0.04 dex, Behmard, Ness, et al. 2023)−valuable information for investigating planet formation and evolution on a galactic scale.

7.2 Future Directions

The results presented in this thesis motivate new and exciting research pathways, which I discuss below.

New Methods for Constraining Planet Host Star Compositions

Though there have been significant advances in measuring elemental abundances for planet host stars, stellar models still struggle with stars that deviate from solar. Cool ($\lesssim$5000 K) stars host molecular species in their atmospheres, which create complex spectral features that stellar models struggle to reproduce (e.g., Rojas-Ayala et al., 2010; Mann, Brewer, et al., 2013). Establishing reliable models for cool stars, e.g., M dwarfs, is of great importance because they are common−M dwarfs comprise ~70% of the stars in our galaxy (Bochanski et al., 2010). Cool star models can be improved with better knowledge of relevant atomic and molecular transition quantum properties, such as associated opacities. Several groups have undertaken laboratory studies to measure such opacities, resulting in catalogs such as the Opacity Project (Seaton et al., 1994), the Laurence Livermore National Labs OPAL opacity tables (Iglesias and Rogers, 1996), and the Los Alamos National Labs OPLIB opacity tables (Colgan et al., 2016), to name a few. Further laboratory studies can constrain more opacities relevant to cool star atmospheres and improve stellar model line lists (Allard, Scholz, and Wehrse, 1992). This is necessary as current stellar models are not reliable for stars with temperatures below ~4700 K (e.g., Brewer and Fischer 2018), that is, mid-K to M-dwarfs.

ML methods such as *The Cannon* are valuable alternatives because they do not rely on stellar models. As outlined in Chapter 2, we trained *The Cannon* on a sample of cool stars with precise iron abundance measurements (Behmard, Petigura, and Howard, 2019). Assembling this training set was possible because iron is a relatively easy element to measure, thanks to numerous strong iron absorption lines in stellar spectra. In the future, ML methods like *The Cannon* could be used to construct automated models for measuring the abundances of elements beyond iron in stars that deviate from solar. These models would be an ideal tool for measuring the elemental abundances of large samples of cool (e.g., M dwarf) planet hosts. However, in order to construct such models, we must first assemble training sets of cool stars with precisely measured abundances for a wide array of elements. Such training sets could be sourced from cool star/solar-like stellar binaries that by definition shared the same birth cloud, and were thus born chemically homogeneous (e.g., Bovy, 2016; Ness, Rix, et al., 2018). It would be straightforward to measure the elemental abundances of the solar-like binary companions with stellar models, then assign those same abundances to their cool star companions. This approach of calibrating cool star compositions with those of solar-like companions is considered the most reliable empirical method for measuring M dwarf abundances (e.g., Mann, Deacon, et al., 2014). Cool star/solar-like binaries could be identified with astrometric data and statistical models that verify gravitationally bound stellar pairs, such as the companion search model employed in Behmard, Dai, and Howard (2022) (Chapter 3). Alternatively, training sets could be constructed from M dwarfs with widely separated solar-like companions that share the same cluster/association. This route is potentially quite valuable because cluster/association member stars tend to have better characterized properties (e.g., compositions, ages) than field stars (Chandar et al., 2016). Stellar ages in particular are valuable for probing planet formation signatures; our work presented in Chapters 4 and 5 indicate that planet engulfment signatures will not remain observable for longer than ~1 Gyr in solar-like stars, so searching for engulfment signatures among young cluster/association members may yield more engulfment detections.

Large-scale spectroscopic surveys can also be used to create training sets of M dwarfs with precisely measured abundances for a wide array of elements. For example, APOGEE provides publicly available abundances for 15 elements (C, N, O, Na, Mg, Al, Si, S, K, Ca, Ti, V, Mn, Fe, and Ni). The APOGEE pipeline utilizes physical models, making it optimal for solar-like stars. Thus, the abundances of solar-like binary or cluster/association companions, and by extension those of the training set M dwarfs, will be very precise. The latest APOGEE data release (DR18) is expected to include >400,000 more stars than DR17, and a 100 pc volume-limited sample containing ~300,000 M dwarfs (Kollmeier *et al.* 2019, M. Ness, priv. comm.). Thus, DR18 will provide a significantly larger M dwarf/solar-like companion training set capable of measuring abundances to even higher precisions. Thus, there is a path forward for constructing ML models that constrain the detailed chemistry of large cool star samples. This will enable new lines of investigation for exploring planet formation around cool star hosts.

Connections Between Host Star and Planetary Atmosphere Compositions

Many studies of planetary atmospheres lack constraints on host star chemical compositions, and typically assume solar-like abundances across different elements. However, planet formation can only be investigated when atmospheric abundances are interpreted relative to accurate host star abundances, which reflect protoplanetary disk compositions. For example, sub-stellar and super-stellar atmospheric C/O can indicate planet formation within and beyond the ice line, respectively (Öberg, Murray-Clay, and Bergin, 2011, Reggiani et al., 2022). Thus, precisely measured host star abundances are key for correctly interpreting planetary atmosphere compositions and illuminating planet formation histories, e.g., initial formation locations and subsequent migration (e.g., Mollière, Molyarova, et al., 2022, Pacetti et al., 2022). In the future, it will be particularly important to develop new cool star models that can easily and quickly measure the abundances of a wide range of elements; planets orbiting cool stars are among the most favorable types of targets for atmospheric characterization with JWST. There are already ~15 M dwarfs with super-Earth to Jupiter-sized planets scheduled for atmospheric observations during

JWST Cycle 1. It is likely that this sample will continue to grow in the coming years.

As we enter the era of detailed exoplanet atmosphere characterization, it is now possible to constrain the abundances of isotopologue species using high resolution spectroscopy. Isotopologue ratios are sensitive tracers of planet formation conditions and planetary system evolution. For example, D/H ratios can probe accretion of D-rich ices from beyond the water ice line and/or atmospheric loss (Drake, 2005; Feuchtgruber et al., 2013). Carbon isotope ratios also probe planet formation locations; atmospheric $^{12}C/^{13}C$ values lower than those of the protoplanetary disk indicate formation beyond the CO ice line. In the absence of direct measurements of host star isotopologue ratios, published studies compared the isotopologue ratios in planetary atmospheres to average local interstellar medium (ISM) values (Mollière and Snellen, 2019; Zhang et al., 2021). Current studies that seek to measure isotopic species in stars require extremely high resolution spectroscopy (e.g., IRTF/iSHELL spectrograph, $R \approx 80{,}000$, Crossfield et al. 2019), boutique stellar models, and careful individual measurements of isotopologue rovibrational features. Thus, these studies are intensely time-consuming and have been carried out for only a handful of stars (e.g., Gay and Lambert 2000; Crossfield et al. 2019). In the future, it may be possible to use ML models for stellar isotopologue abundance measurements if we can curate appropriate training sets. This may require streamlining isotopologue measurement pipelines for easier-to-model stars, such as those with solar-like types, then finding bound companions with different stellar types in order to diversify training sets for application to a wide set of planet host stars. This is a useful potential route for measuring isotopologue abundances in stars whose spectra exhibit weak isotopologue features (e.g., hotter stars), or are rich in other atomic and molecular features that may crowd isotopologue lines (e.g., cool stars such as M dwarfs).

Earth-Like Planet Formation

Automated models for measuring stellar abundances of small, cool stars such as M dwarfs would be ideal for investigating the formation of small, rocky exoplanets. M dwarfs are excellent targets for small exoplanet discoveries−in addition to being common, M dwarfs have small sizes and low masses that result in strong RV and transit signals, thereby increasing the probability of detecting small planets. It is also more likely to find a planet in the habitable zone (HZ) of an M dwarf star compared to that of a solar-like star; M dwarfs are cooler, so their HZs are relatively close-in. Mann *et al.* (2013) estimated that a planet in the HZ of an M dwarf will have ∼3 times the RV signal, ∼4 times the transit depth, and ∼2 times the probability of transiting compared to a HZ planet orbiting a solar-like star. Thus, M dwarfs are ideal for detecting and characterizing exoplanets that are truly similar to Earth.

M dwarfs are especially attractive targets for the ongoing TESS mission, whose primary goal is the discovery of small planets orbiting bright, nearby stars (Ricker et al., 2015). This goal is motivated by our current census of exoplanets, which is dominated by planets with sizes between those of Earth and Neptune (e.g., Fressin et al., 2013; Burke et al., 2015). Because small planets are common, there is great interest in understanding where and how these planets formed. Their interior compositions can provide answers, but must be constrained with host star abundances of refractory (e.g., Fe, Mg, Si) and volatile (e.g., C, O) elements that dictate their amounts in protoplanetary disks. This initial inventory is needed to break degeneracies between different possible planet compositions with equivalent bulk densities (e.g., Dorn et al., 2015; Bitsch and Battistini, 2020). For example, Agol et al., 2021 determined that the seven Earth-like TRAPPIST-1 planets have lower densities than those of the solar system terrestrial planets, which points to lighter interiors (e.g., lower iron content), or volatile (e.g., water) enrichment. The latter case would be a signpost of planet formation beyond the ice line. However, we cannot distinguish between these scenarios without knowledge of TRAPPIST-1's abundances for elements other than iron, which have never been measured for this star.

An automated M dwarf model would enable measurement of detailed elemental abundances for M dwarf hosts like TRAPPIST-1, creating new pathways for investigating Earth-like planet formation with host star chemistry. TESS has detected 135 Earth-like ($1-2\ R_\oplus$) planet candidates and 39 Earth-like confirmed planets around M dwarfs so far, which provides a large sample for this purpose. Follow-up RV observations for TESS M dwarf hosts of Earth-like planets will constrain planet masses, and focusing on M dwarfs with the most anomalous abundances (e.g., atypical C/O, Fe/Mg, etc.) would make it possible to probe edge cases of Earth-like planet formation. Coupled with publicly available TESS planet radii measurements, planet masses will provide planet bulk densities that probe Earth-like planet interior compositions.

References

Agol, Eric et al. (Feb. 2021). "Refining the Transit-timing and Photometric Analysis of TRAPPIST-1: Masses, Radii, Densities, Dynamics, and Ephemerides". In: *The*

Planetary Science Journal 2.1, 1, p. 1. DOI: 10.3847/PSJ/abd022. arXiv: 2010.01074 [astro-ph.EP].

Allard, F., M. Scholz, and R. Wehrse (Mar. 1992). “Molecular opacities in M-star atmospheres.” In: *Revista Mexicana de Astronomia y Astrofisica* 23, pp. 203–215.

Behmard, Aida, Fei Dai, John M. Brewer, et al. (May 2023). “Planet engulfment detections are rare according to observations and stellar modelling”. In: *Monthly Notices of the Royal Astronomical Society* 521.2, pp. 2969–2987. DOI: 10.1093/mnras/stad745. arXiv: 2210.12121 [astro-ph.EP].

Behmard, Aida, Fei Dai, and Andrew W. Howard (Apr. 2022). “Stellar Companions to TESS Objects of Interest: A Test of Planet-Companion Alignment”. In: *The Astronomical Journal* 163.4, 160, p. 160. DOI: 10.3847/1538-3881/ac53a7. arXiv: 2202.01798 [astro-ph.EP].

Behmard, Aida, Melissa K. Ness, et al. (Apr. 2023). “Elemental Abundances of Kepler Objects of Interest in APOGEE DR17”. In: *The Astronomical Journal* 165.4, 178, p. 178. DOI: 10.3847/1538-3881/acc32a.

Behmard, Aida, Erik A. Petigura, and Andrew W. Howard (May 2019). “Data-driven Spectroscopy of Cool Stars at High Spectral Resolution”. In: *The Astrophysical Journal* 876.1, 68, p. 68. DOI: 10.3847/1538-4357/ab14e0. arXiv: 1904.00094 [astro-ph.EP].

Behmard, Aida, Jason Sevilla, and Jim Fuller (Feb. 2023). “Planet engulfment signatures in twin stars”. In: *Monthly Notices of the Royal Astronomical Society* 518.4, pp. 5465–5474. DOI: 10.1093/mnras/stac3435. arXiv: 2210.11679 [astro-ph.SR].

Bitsch, Bertram and Chiara Battistini (Jan. 2020). “Influence of sub- and super-solar metallicities on the composition of solid planetary building blocks”. In: *Astronomy & Astrophysics* 633, A10, A10. DOI: 10.1051/0004-6361/201936463. arXiv: 1911.09725 [astro-ph.EP].

Bochanski, John J. et al. (June 2010). “The Luminosity and Mass Functions of Low-mass Stars in the Galactic Disk. II. The Field”. In: *The Astronomical Journal* 139.6, pp. 2679–2699. DOI: 10.1088/0004-6256/139/6/2679. arXiv: 1004.4002 [astro-ph.SR].

Bovy, Jo (Jan. 2016). “The Chemical Homogeneity of Open Clusters”. In: *The Astrophysical Journal* 817.1, 49, p. 49. DOI: 10.3847/0004-637X/817/1/49. arXiv: 1510.06745 [astro-ph.GA].

Brewer, J. M. and D. A. Fischer (Aug. 2018). “Spectral Properties of Cool Stars: Extended Abundance Analysis of Kepler Objects of Interest”. In: *The Astrophysical Journal Supplement Series* 237, 38, p. 38. DOI: 10.3847/1538-4365/aad501. arXiv: 1804.00673 [astro-ph.SR].

Burke, Christopher J. et al. (Aug. 2015). "Terrestrial Planet Occurrence Rates for the Kepler GK Dwarf Sample". In: *The Astrophysical Journal* 809.1, 8, p. 8. doi: 10.1088/0004-637X/809/1/8. arXiv: 1506.04175 [astro-ph.EP].

Chandar, Rupali et al. (June 2016). "The Age, Mass, and Size Distributions of Star Clusters in M51". In: *The Astrophysical Journal* 824.2, 71, p. 71. doi: 10.3847/0004-637X/824/2/71.

Colgan, J. et al. (Feb. 2016). "A New Generation of Los Alamos Opacity Tables". In: *The Astrophysical Journal* 817.2, 116, p. 116. doi: 10.3847/0004-637X/817/2/116. arXiv: 1601.01005 [astro-ph.SR].

Crossfield, I. J. M. et al. (Jan. 2019). "Unusual Isotopic Abundances in a Fully Convective Stellar Binary". In: *The Astrophysical Journal Letters* 871.1, L3, p. L3. doi: 10.3847/2041-8213/aaf9b6. arXiv: 1901.02607 [astro-ph.SR].

Dorn, Caroline et al. (May 2015). "Can we constrain the interior structure of rocky exoplanets from mass and radius measurements?" In: *Astronomy & Astrophysics* 577, A83, A83. doi: 10.1051/0004-6361/201424915. arXiv: 1502.03605 [astro-ph.EP].

Drake, M. J. (Apr. 2005). "The Leonard Medal Address: Origin of water in the terrestrial planets". In: *Meteoritics & Planetary Science* 40, p. 519. doi: 10.1111/j.1945-5100.2005.tb00960.x.

Feuchtgruber, H. et al. (Mar. 2013). "The D/H ratio in the atmospheres of Uranus and Neptune from Herschel-PACS observations". In: *Astronomy and Astrophysics* 551, A126, A126. doi: 10.1051/0004-6361/201220857. arXiv: 1301.5781 [astro-ph.EP].

Fressin, François et al. (Apr. 2013). "The False Positive Rate of Kepler and the Occurrence of Planets". In: *The Astrophysical Journal* 766.2, 81, p. 81. doi: 10.1088/0004-637X/766/2/81. arXiv: 1301.0842 [astro-ph.EP].

Gay, Pamela L. and David L. Lambert (Apr. 2000). "The Isotopic Abundances of Magnesium in Stars". In: *The Astrophysical Journal* 533.1, pp. 260–270. doi: 10.1086/308653. arXiv: astro-ph/9911217 [astro-ph].

Gonzalez, Guillermo (Feb. 1997). "The stellar metallicity-giant planet connection". In: *Monthly Notices of The royal* 285.2, pp. 403–412. doi: 10.1093/mnras/285.2.403.

Howell, Steve B. et al. (Apr. 2021). "Speckle Observations of TESS Exoplanet Host Stars: Understanding the Binary Exoplanet Host Star Orbital Period Distribution". In: *The Astronomical Journal* 161.4, 164, p. 164. doi: 10.3847/1538-3881/abdec6. arXiv: 2101.08671 [astro-ph.EP].

Iglesias, Carlos A. and Forrest J. Rogers (June 1996). "Updated Opal Opacities". In: *The Astrophysical Journal* 464, p. 943. doi: 10.1086/177381.

Kollmeier, Juna et al. (Sept. 2019). "SDSS-V Pioneering Panoptic Spectroscopy". In: *Bulletin of the American Astronomical Society*. Vol. 51, 274, p. 274.

Mann, Andrew W., John M. Brewer, et al. (Feb. 2013). "Prospecting in Late-type Dwarfs: A Calibration of Infrared and Visible Spectroscopic Metallicities of Late K and M Dwarfs Spanning 1.5 dex". In: *The Astronomical Journal* 145.2, 52, p. 52. DOI: 10.1088/0004-6256/145/2/52. arXiv: 1211.4630 [astro-ph.SR].

Mann, Andrew W., Niall R. Deacon, et al. (June 2014). "Prospecting in Ultracool Dwarfs: Measuring the Metallicities of Mid- and Late-M Dwarfs". In: *The Astronomical Journal* 147.6, 160, p. 160. DOI: 10.1088/0004-6256/147/6/160. arXiv: 1403.5560 [astro-ph.SR].

Mann, Andrew W., Eric Gaidos, and Megan Ansdell (Dec. 2013). "Spectro-thermometry of M Dwarfs and Their Candidate Planets: Too Hot, Too Cool, or Just Right?" In: *The Astrophysical Journal* 779.2, 188, p. 188. DOI: 10.1088/0004-637X/779/2/188. arXiv: 1311.0003 [astro-ph.EP].

Millholland, Sarah and Konstantin Batygin (May 2019). "Excitation of Planetary Obliquities through Planet-Disk Interactions". In: *The Astrophysical Journal* 876.2, 119, p. 119. DOI: 10.3847/1538-4357/ab19be. arXiv: 1904.07338 [astro-ph.EP].

Mollière, P. and I. A. G. Snellen (Feb. 2019). "Detecting isotopologues in exoplanet atmospheres using ground-based high-dispersion spectroscopy". In: *Astronomy and Astrophysics* 622, A139, A139. DOI: 10.1051/0004-6361/201834169. arXiv: 1809.01156 [astro-ph.EP].

Mollière, Paul, Tamara Molyarova, et al. (July 2022). "Interpreting the Atmospheric Composition of Exoplanets: Sensitivity to Planet Formation Assumptions". In: *The Astrophysical Journal* 934.1, 74, p. 74. DOI: 10.3847/1538-4357/ac6a56. arXiv: 2204.13714 [astro-ph.EP].

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, 16, p. 16. DOI: 10.1088/0004-637X/808/1/16. arXiv: 1501.07604 [astro-ph.SR].

Ness, M., H. -W. Rix, et al. (Feb. 2018). "Galactic Doppelgängers: The Chemical Similarity Among Field Stars and Among Stars with a Common Birth Origin". In: *The Astrophysical Journal* 853.2, 198, p. 198. DOI: 10.3847/1538-4357/aa9d8e. arXiv: 1701.07829 [astro-ph.SR].

Öberg, Karin I., Ruth Murray-Clay, and Edwin A. Bergin (Dec. 2011). "The Effects of Snowlines on C/O in Planetary Atmospheres". In: *The Astrophysical Journal Letters* 743.1, L16, p. L16. DOI: 10.1088/2041-8205/743/1/L16. arXiv: 1110.5567 [astro-ph.GA].

Oh, Semyeong et al. (June 2017). "Comoving Stars in Gaia DR1: An Abundance of Very Wide Separation Comoving Pairs". In: *The Astronomical Journal* 153.6, 257, p. 257. DOI: 10.3847/1538-3881/aa6ffd. arXiv: 1612.02440 [astro-ph.SR].

Pacetti, Elenia et al. (Sept. 2022). "Chemical Diversity in Protoplanetary Disks and Its Impact on the Formation History of Giant Planets". In: *The Astrophysical Journal* 937.1, 36, p. 36. DOI: 10.3847/1538-4357/ac8b11. arXiv: 2206.14685 [astro-ph.EP].

Reggiani, Henrique et al. (Apr. 2022). "Evidence that the Hot Jupiter WASP-77 A b Formed Beyond Its Parent Protoplanetary Disk's H2O Ice Line". In: *The Astronomical Journal* 163.4, 159, p. 159. DOI: 10.3847/1538-3881/ac4d9f. arXiv: 2201.08508 [astro-ph.EP].

Ricker, George R. et al. (Jan. 2015). "Transiting Exoplanet Survey Satellite (TESS)". In: *Journal of Astronomical Telescopes, Instruments, and Systems* 1, 014003, p. 014003. DOI: 10.1117/1.JATIS.1.1.014003.

Rojas-Ayala, Bárbara et al. (Sept. 2010). "Metal-rich M-Dwarf Planet Hosts: Metallicities with K-band Spectra". In: *The Astrophysical Journal Letters* 720.1, pp. L113–L118. DOI: 10.1088/2041-8205/720/1/L113. arXiv: 1007.4593 [astro-ph.SR].

Seaton, M. J. et al. (Feb. 1994). "Opacities for stellar envelopes". In: *Monthly Notices of the Royal Astronomical Society* 266.4, pp. 805–828. ISSN: 0035-8711. DOI: 10.1093/mnras/266.4.805. eprint: https://academic.oup.com/mnras/article-pdf/266/4/805/3050789/mnras266-0805.pdf. URL: https://doi.org/10.1093/mnras/266.4.805.

Zhang, Yapeng et al. (July 2021). "The ^{13}CO-rich atmosphere of a young accreting super-Jupiter". In: *Nature* 595.7867, pp. 370–372. DOI: 10.1038/s41586-021-03616-x. arXiv: 2107.06297 [astro-ph.EP].

BIBLIOGRAPHY

Akeson, R. L. et al. (Aug. 2013). "The NASA Exoplanet Archive: Data and Tools for Exoplanet Research". In: *Publications of the Astronomical Society of the Pacific* 125.930, p. 989. DOI: 10.1086/672273. arXiv: 1307.2944 [astro-ph.IM].

Andrews, Jeff J. et al. (Jan. 2019). "Using APOGEE Wide Binaries to Test Chemical Tagging with Dwarf Stars". In: *The Astrophysical Journal* 871.1, 42, p. 42. DOI: 10.3847/1538-4357/aaf502. arXiv: 1811.12032 [astro-ph.SR].

Bedell, Megan et al. (Sept. 2018). "The Chemical Homogeneity of Sun-like Stars in the Solar Neighborhood". In: *The Astrophysical Journal* 865.1, 68, p. 68. DOI: 10.3847/1538-4357/aad908. arXiv: 1802.02576 [astro-ph.SR].

Brewer, J. M. and D. A. Fischer (Aug. 2018). "Spectral Properties of Cool Stars: Extended Abundance Analysis of Kepler Objects of Interest". In: *The Astrophysical Journal Supplement Series* 237, 38, p. 38. DOI: 10.3847/1538-4365/aad501. arXiv: 1804.00673 [astro-ph.SR].

Feeney, Stephen M., Benjamin D. Wandelt, and Melissa K. Ness (Mar. 2021). "SSSpaNG! stellar spectra as sparse, data-driven, non-Gaussian processes". In: *Monthly Notices of the Royal Astronomical Society* 501.3, pp. 3258–3271. DOI: 10.1093/mnras/staa3586. arXiv: 1912.09498 [astro-ph.SR].

Hawkins, Keith et al. (Feb. 2020). "Identical or fraternal twins? The chemical homogeneity of wide binaries from Gaia DR2". In: *Monthly Notices of The Royal Astronomical Society* 492.1, pp. 1164–1179. DOI: 10.1093/mnras/stz3132. arXiv: 1912.08895 [astro-ph.SR].

Hwang, Hsiang-Chih, Yuan-Sen Ting, and Nadia L. Zakamska (May 2022). "The eccentricity distribution of wide binaries and their individual measurements". In: *Monthly Notices of the Royal Astronomical Society* 512.3, pp. 3383–3399. DOI: 10.1093/mnras/stac675. arXiv: 2111.01789 [astro-ph.SR].

Johnson, John Asher et al. (Aug. 2010). "Giant Planet Occurrence in the Stellar Mass-Metallicity Plane". In: *Publications of the Astronomical Society of the Pacific* 122.894, p. 905. DOI: 10.1086/655775. arXiv: 1005.3084 [astro-ph.EP].

Lodders, Katharina (July 2003). "Solar System Abundances and Condensation Temperatures of the Elements". In: *The Astrophysical Journal* 591.2, pp. 1220–1247. DOI: 10.1086/375492.

Ness, M. et al. (July 2015). "The Cannon: A data-driven approach to Stellar Label Determination". In: *The Astrophysical Journal* 808.1, 16, p. 16. DOI: 10.1088/0004-637X/808/1/16. arXiv: 1501.07604 [astro-ph.SR].

Oh, Semyeong et al. (Feb. 2018). “Kronos and Krios: Evidence for Accretion of a Massive, Rocky Planetary System in a Comoving Pair of Solar-type Stars”. In: *The Astrophysical Journal* 854.2, 138, p. 138. DOI: 10.3847/1538-4357/aaab4d. arXiv: 1709.05344 [astro-ph.SR].

Raghavan, Deepak et al. (Sept. 2010). “A Survey of Stellar Families: Multiplicity of Solar-type Stars”. In: *The Astrophysical Journal Supplement Series* 190.1, pp. 1–42. DOI: 10.1088/0067-0049/190/1/1. arXiv: 1007.0414 [astro-ph.SR].

Spina, Lorenzo et al. (Aug. 2021). “Chemical evidence for planetary ingestion in a quarter of Sun-like stars”. In: *Nature Astronomy*. DOI: 10.1038/s41550-021-01451-8. arXiv: 2108.12040 [astro-ph.SR].

Tokovinin, A. and O. Kiyaeva (Feb. 2016). “Eccentricity distribution of wide binaries”. In: *Monthly Notices of the Royal Astronomical Society* 456.2, pp. 2070–2079. DOI: 10.1093/mnras/stv2825. arXiv: 1512.00278 [astro-ph.SR].

Yee, S. W., E. A. Petigura, and K. von Braun (Feb. 2017). “Precision Stellar Characterization of FGKM Stars using an Empirical Spectral Library”. In: *The Astrophysical Journal* 836, 77, p. 77. DOI: 10.3847/1538-4357/836/1/77. arXiv: 1701.00922 [astro-ph.SR].