The Fundamental Properties of Young Stars

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Czekala, Ian. 2016. The Fundamental Properties of Young Stars. Doctoral dissertation, Harvard University, Graduate School of Arts &amp; Sciences.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:33493279">http://nrs.harvard.edu/urn-3:HUL.InstRepos:33493279</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
The Fundamental Properties of Young Stars

A dissertation presented

by

Ian Czekala

to

The Department of Astronomy

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Astronomy and Astrophysics

Harvard University

Cambridge, Massachusetts

May 2016
The Fundamental Properties of Young Stars

Abstract

Accurate knowledge of the fundamental properties of stars—mass, temperature, and luminosity—is key to our understanding of stellar evolution. In particular, empirical measurements of stellar mass are difficult to make and are generally limited to stars that dynamically interact with a companion (e.g., eclipsing or astrometric binaries), a precious but ultimately small sample. We developed a technique that uses the rotation of the protoplanetary disk—a consequence of the star formation process still present around many pre-main sequence stars—to measure the stellar mass. To establish the absolute accuracy of this technique, in ALMA Cycle 1/2 we observed the few circumbinary disks around double-lined spectroscopic binary stars, enabling an independent confirmation of the total stellar mass. This comparison with radial-velocity results demonstrates that the disk-based dynamical mass technique can reliably achieve precise measurements of stellar mass on the order of 2-5%, clearing the way for widespread application of this technique to measure the masses of single stars. We discuss our calibration in the context of two sources, AK Sco and DQ Tau.

Second, we developed novel statistical techniques for spectroscopic inference. Young stars exhibit rich and variable spectra; although interesting phenomena in their own right, accretion veiling and star spots complicate the retrieval of accurate photospheric properties. The subtraction of an imperfect model from a continuously sampled spectrum introduces covariance between adjacent datapoints (pixels) into the residual spectrum.
For the high signal-to-noise data with large spectral range that is commonly employed in stellar astrophysics, that covariant structure can lead to dramatically underestimated parameter uncertainties (and, in some cases, biases). We construct a likelihood function that accounts for the structure of the covariance matrix, utilizing the machinery of Gaussian process kernels. This framework specifically addresses the common problem of mismatches in model spectral line strengths (with respect to data) due to intrinsic model imperfections (e.g., in the atomic/molecular databases or opacity prescriptions) by developing a novel local covariance kernel formalism that identifies and self-consistently downweights pathological spectral line “outliers.” We demonstrate some salient features of the framework by fitting the high resolution $V$-band spectrum of WASP-14, an F5 dwarf with a transiting exoplanet, and the moderate resolution $K$-band spectrum of Gliese 51, an M5 field dwarf. Direct spectroscopic inference provides one means to avoid the systematic error that results from the uncertain spectral type–effective temperature scale for low mass pre-main sequence stars when placing a star on the Hertzsprung Russell diagram.

Lastly, we discuss recent progress in measuring the masses of a large sample of single pre-main sequence stars observed with the Submillimeter Array, which will double the number of disk-based dynamical mass estimates of pre-main sequence stars. With ALMA, the disk-based technique holds enormous promise to become the primary means of stellar mass for statistically large samples of pre-main sequence stars, ushering in a new era of high precision in star and planet formation studies.
Acknowledgments

The research presented in this thesis would not have been possible without the support from numerous individuals. First and foremost, I would like to acknowledge the wonderful and steadfast mentorship of Sean Andrews, my Ph.D. advisor. I first had the pleasure of working with Sean when I was a Smithsonian Astrophysical Observatory REU student in the summer of 2009. At the time, I was an undergraduate majoring in Aerospace Engineering at the University of Virginia and new to the concept of astrophysical research as a profession. Sean’s involved and upbeat mentoring style made my first significant astrophysics research experience an incredibly positive one. Years later, as a Ph.D. advisor, Sean’s support has continued to be a primary factor in my growth as a researcher. His keen intellect has been crucial for identifying and tackling tractable scientific problems, his curiosity has been vital to exploring exciting new directions, and his frank advice has been invaluable for my professional development. It has been a true privilege to work with Sean.

I would also like to thank many members of the radio astronomy and protoplanetary disk communities. Since my arrival at Harvard, Jim Moran has provided helpful and timely advice about a wide range of subjects such as future research directions, postdoctoral fellowships, and the occasional Fourier transform. Serving as the teaching fellow for Jim’s graduate course “Noise and Data Analysis” was one of my highlights of graduate school, and it is safe to say that I rely on concepts from this course on a near-daily basis (e.g., see following chapters). I greatly appreciate the advice of David Wilner, an expert educator, radio astronomer, and collaborator, who has helped me understand the intricacies and fundamentals of (sub)mm observations of disks. I would
especially like to thank Karin Öberg, a generous and steadfast supporter whose optimism has helped me focus on the bigger picture when it would have been easiest to get bogged down in the weeds. Occasional collaboration with Karin’s exciting and welcoming astrochemistry group has been a major benefit to researching protoplanetary disks at the CfA. Lastly, I would like to thank Kelsey Johnson and Alison Peck, who showed me how exciting astrophysical research can be. By lending their time and support at a crucial moment in my scientific development, they proved to be a major influence in my decision to pursue an astrophysics Ph.D.

Next, I would like to thank several members of the stellar and exoplanet communities. Dave Latham and I fortuitously crossed paths when he helped schedule our first observations on the TRES spectrograph. Since then, he has become a valued collaborator whose extensive knowledge of pre-main sequence stars has greatly extended the impact of our research, and someone I frequently rely upon for brainstorming of new spectroscopic projects. Furthermore, I greatly appreciate the helpful spectroscopic advice from Willie Torres, whose guidance in understanding spectral lines has been vital to my education as a stellar spectroscopist. As a collaborator, he has set a high standard for thoughtfulness and dependability, one which I strive to emulate. And I would like to thank Keivan Stassun, an energetic and helpful collaborator who has challenged me to think about stellar properties and their effect on stellar evolution. Lastly, I would like to express my gratitude towards John Johnson and the Exo-lab for helpful support, discussions, and a continual engagement of interesting statistical ideas, which were crucial to the development of some of our papers.

I would like to thank Edo Berger for an exciting and fast-paced introduction to graduate school, when I joined his transients research group as an incoming student.
Edo and other members of the transients group invested a substantial amount of effort training me as a multi-wavelength observer, skills which have gone on to serve me well throughout my Ph.D. I would like to thank the many graduate students and postdocs who have helped me throughout graduate school by offering me scientific advice, sounding boards for nascent ideas, and the frequent chat to blow off some stress: Kaisey Mandel, Maxwell Moe, Greg Green, Ruth Angus, Dan Foreman-Mackey, Ben Montet, Katherine Rosenfeld, Ryan Loomis, Jane Huang, Anjali Tripathi, and Ilse Cleeves. And of course, many thanks to the department administrators who have supported my research, Peg Herlihy, Robb Scholten, Muriel Hodges, and Lisa Catella.

Above all, thank you to my family. My mother and father, Barbara and Joseph, a research librarian and a clinical psychologist, respectively, ignited the spark of learning and taught me the value of personal relationships. Without their unending love and support, I would never have made it to where I am today. Thank you to Keru Cai, for your limitless love and support in all things small and big, not the least of which being this thesis.
Chapter 1

Introduction

1.1 How stellar properties evolve with time

The star is a fundamental unit of astrophysics. It provides the dominant source of radiation in most regions of the universe, its evolution drives the chemical enrichment of a galaxy, and its gravitational field sets the dynamical landscape for planet formation. An accurate understanding of how the fundamental properties of a star change throughout its lifetime are of central importance numerous areas of astrophysical research. For our purposes, we will identify the fundamental properties as the stellar mass \( M_\ast \), age \( \tau \), effective temperature \( T_{\text{eff}} \), and luminosity \( L_\ast \). For the most part, the stellar phenomenology boils down to a single parameter: \( M_\ast \). That is to say, given an initial stellar mass, evolutionary models can predict the photospheric properties as the star
evolves, e.g., $T_{\text{eff}}(\tau, M_*)$ and $L_*(\tau, M_*)$.

When revolutions in modern astrophysics throughout the early and mid 20th century triumphantly led to the understanding of what powers the stars (e.g., Burbidge et al. 1957), questions about how stars were born and evolved closely followed. The T Tauri stars were first identified by Alfred Joy (Joy 1945), but it took several years (at least in America, see Ambartsumian 1947, 1954) before these stars were realized to be the precursors to their main sequence counterparts (Herbig 1962). What followed was an exciting rush to understand the site and processes of star formation, and how young stellar systems evolved into their more common, older cousins on the main-sequence of the Hertzsprung Russell Diagram.

It soon became clear that these young stars were born much more luminous and redder than their main sequence counterparts, implying that their radii were both larger and cooler. At the same time, detailed models of stellar evolution began to predict how stars would evolve in these young ages as they contracted to the main sequence under the force of their own gravity. For example, Louis Henyey was one of the first to make detailed calculations of how a star evolves throughout the HR diagram as it develops a radiative core (Henyey et al. 1955). Hayashi later figured out that stars would be in a fully convective state as they contracted and made their way towards the main sequence (Hayashi 1961; Hayashi et al. 1962).

The general landscape of young stellar evolution was shortly sketched out in the 1960s. However, it took a while for the details to be worked out. A modern view of how

---

1Initial composition (i.e., metallicity) also affects stellar evolution; however, in this thesis we will focus solely on stars in nearby star-forming regions with solar abundances and therefore this effect is negligible.
Figure 1.1.—: The Dotter et al. (2008) pre-main sequence evolutionary models showing how stars evolve throughout the Hertzsprung Russell diagram. In general, as stars make their way to the main sequence they contract, dim, and become hotter. In detail, this behavior depends sensitively on the stellar mass ($M_*$).
the stellar radii and the ratio of the effective temperatures. The main limitation is that these systems are very rare, because not only do they require a chance alignment, but they also require substantial observational monitoring to detect an eclipse and thereafter extensive radial velocity monitoring to constrain the orbit.

EBs provide tight constraints on fundamental stellar parameters in the following manner. The geometry of the eclipse tightly constrains the system to be $i \approx 90^\circ$. For double-lined spectroscopic binaries, radial velocity (RV) monitoring yields the quantity $M_\ast \sin^3 i$, where $M_\ast$ is the sum of the stellar masses. In the case where there are RV measurements of an EB, one can achieve a very accurate mass determination with no information on the distance to the source. The precision of EB fundamental parameters is nearly always only limited by the precision and number of RV measurements; for extreme mass ratio binaries, it can be difficult to detect spectral features of the fainter companion and measure an RV. The *Kepler* and K2 missions have helped identify many new eclipsing binaries that will hopefully soon achieve benchmark status with repeated RV monitoring (David et al. 2016a,b).

For the purposes of calibrating early pre-main sequence evolution, it is true that a few very solid measurements (e.g., Stassun et al. 2014) can help more than many measurements that may suffer from systematic errors. So, this sample of eclipsing binaries, although small, has proven to be exceptionally useful because they provide narrow milestones that evolutionary models need to pass. However it is also apparent that EBs are not plentiful enough to fully constrain the PMS HR diagram everywhere.
1.2 The Disk Based Dynamical Mass Technique

The development of sub-mm interferometers sensitive enough to detect molecular gas emission from the protoplanetary disks of young stars created the potential for a new technique to create benchmark stellar systems. The disk-based dynamical mass technique relies on reconstructing the velocity field of the protoplanetary disk, which is assumed to be in Keplerian orbit around the star. It was recognized early on that position-velocity diagrams could be used to probe the mass of the star dynamically (Dutrey et al. 1994). However, the first disk-based dynamical mass technique is generally accepted to be Guilloteau & Dutrey (1998), who derived the mass of DM Tau using an iterative $\chi^2$ minimization technique by modeling raw visibility data from the interferometer with physically motivated models. They show that the velocity curve is consistent with Keplerian rotation, and derive a mass of $0.5 \pm 0.06 M_\odot$, a result that is consistent with our recent measurements using the Submillimeter Array (hereafter SMA).

As Guilloteau & Dutrey (1998) note, there are clear degeneracies that may crop up in the measurement of the stellar mass. The first is an uncertainty of the distance to the source, since we measure $M_*/d$. This means that uncertain distances fold into the mass determination in a linear manner, and that many typical protoplanetary disks (e.g., in Taurus) will incur an additional 15% systematic uncertainty. Second, there is a degeneracy in the determination of the inclination of the source, since we are measuring the projected velocity field. This scales like the quantity $M_* \sin^2 i$, meaning that more face-on disks will yield less-precise mass estimates. As Guilloteau & Dutrey (1998)

2The GAIA mission will soon measure parallaxes to all sources for which a dynamical mass measurement might be possible, completely eliminating this systematic uncertainty.
explain, the same 13% mass precision is obtained for a disk with a measured inclination of \( i = 70 \pm 10^\circ \) as a disk with \( i = 33 \pm 3^\circ \). Although this correlation is always present in any dynamical mass measurement, it becomes a problem when the observations do not provide sufficient spatial resolution to fully resolve the disk.

Simon et al. (2000) pioneered measurements of the first sample of disk-based dynamical masses, publishing masses for 10 pre-main sequence stars. As mentioned previously, one of the shortcomings of the disk-based dynamical mass technique is that it is not possible to directly determine the distance to the source from the disk alone, instead we measure the quantity \( M_*/d \). Likewise, luminosity estimates also are hampered by an uncertain distance converting flux to luminosity, with photometry we actually measure the flux, which is proportional to \( L_*/d^2 \). Simon et al. (2000) realized that the measured properties could be compared directly to the pre-main sequence model predictions via the distance-independent quantity \( L_*/M_*^2 \), avoiding the systematic error from the distance uncertainty. Their analysis provided constraints that disfavored some of the hotter stellar models at the time. Over the course of my Ph.D., the number of pre-main sequence benchmark systems grew from about 20 to 30 systems Stassun et al. (2014); Guilloteau et al. (2014); Kraus et al. (2015); Rizzuto et al. (2016); soon we will have more than 50 systems.

The discriminatory power of the disk-based technique was realized early on. However, lingering questions remained about whether this technique was truly accurate, since the modeling involves a number of assumptions and understanding of complicated processes. Simon et al. (2000) published the first measurements of a circumbinary disks with UZ Tau E and GG Tau, potentially enabling an independent check with radial velocity measurements. However, only UZ Tau E is a spectroscopic binary, and Simon
et al. (2000) found the derived masses to be inconsistent, implying that the disk and binary orbits might be misaligned. Rosenfeld et al. (2012) was the first to conclusively demonstrate the accuracy of the dynamical mass technique by using resolved SMA observations of V4046 Sgr with long term radial velocity monitoring to demonstrate a remarkably precise dynamical mass measurement (< 4%) that was in complete agreement with the radial-velocity derived masses (which were at higher precision). Work remained, however, to demonstrate the consistency of disk-based technique using a larger sample of spectroscopic binaries with circumbinary disks.

1.3 Confronting the Observational Systematics of Photospheric Properties

Over the past three decades, there has been tremendous progress in populating the pre-main sequence HR diagram with reliable benchmark calibrators. However, due to the variable and difficult-to-determine photospheric properties of these sources, we quickly moved from having no data to having data dominated by systematic uncertainties. Even nine years ago, Mathieu et al. (2007) noted

“Now that dynamical mass measurements are relatively abundant, and will become more so with the application of groundbased optical/infrared interferometers, the primary limitations to such tests have become systematic errors in the determination of stellar properties necessary for the comparison with evolutionary models, such as effective temperature, luminosity, and radii. Additional dynamical mass determinations between 0.5 \( M_\odot \) and 2 \( M_\odot \)
will not likely improve the constraints on evolutionary models until these systematic uncertainties in measurements of stellar properties are reduced.”

In many ways, the nature of the problem as well as the time frame closely mirror the field of Type Ia supernova cosmology. There, in just over two decades, the determination of the equation of state $w$ went from being poorly constrained to a known quantity limited by systematics in the diversity of supernova explosions. In order to better constrain cosmological parameters, the challenge now is not to observe more Type Ia supernovae, rather, it is to understand and address the sources of systematic errors, such as the nature of the dust intrinsic to the supernova host galaxy (Conley et al. 2007). Clearly, new thinking is required to achieve higher precisions in both supernovae cosmology and the calibration of pre-main sequence evolution.

It is unlikely that the authors of Mathieu et al. (2007) believe the following statements, but the quote by Mathieu et al. (2007) suggests several latent possibilities about the calibration of the pre-main sequence HR diagram so that it would be worthwhile to state them clearly and subsequently address them. First, all other properties being equal (age, metallicity), it is presupposed that there exists a one-to-one mapping between $M_*$ and the observed photospheric properties $T_{\text{eff}}, L_*$. If we could only measure the photospheric properties accurately, then pre-main sequence models would line up with the benchmark calibrators. What if this relationship actually included several other competing factors? While the evolutionary models are consistent with a one-to-one mapping within the current measurement uncertainties, there are exciting reasons to believe that this relationship might eventually break down at a higher level of precision due to varied accretion history, formation history, or chromospheric activity
CHAPTER 1. INTRODUCTION

(Stassun et al. 2012).

Second, while the Mathieu et al. (2007) statement is true to first order, there are clearly regions of the young HR diagram in need of accurate benchmark systems. I consider the calibration of the pre-main sequence evolutionary models like the process of raising a carnival tent in preparation for a circus. Sure, once you install the main support and toss over the canvas, you would, to first order, have a tent. But it would sag in places and generally provide a less-than-ideal experience for most carnival goers. After this first pass, careful and deliberate work over the following hours and days is required to bring the canvas taught and properly straighten it to provide an aesthetically pleasing arena for the circus. In much the same way, I like to think this analogy describes the current status of the pre-main sequence HR diagram. We have the general idea, but there are still interesting “wrinkles” that need to be ironed out by careful study of individual benchmark systems.

One clear source of systematic error that will not go away anytime soon is the determination of the effective temperature of the stellar photosphere. In the literature, most $T_{\text{eff}}$ measurements come from a conversion of the spectral type of the pre-main sequence star. This conversion is not necessarily unique and introduces measurement systematics through classification by a human spectroscopist. In reality, this conversion is approximate and may have some intrinsic scatter. For low-mass young stars, uncertainties in the surface gravity conversion can create systematic uncertainties as well (Luhman 1999; Kraus et al. 2015). Large luminosity uncertainties (20% up to 40%) also contribute to the difficulty of placing PMS stars on the HR diagram. For EB systems, it is possible to get around this by making dynamical mass comparisons in planes for which the tracks have the largest dynamic range. For example, Kraus et al. (2015)
analyze the young, cool eclipsing binary UScoCTIO 5 in the mass vs. radius plane to achieve maximal discriminatory power against the models and avoids a conversion into the uncertain temperature scale. Clearly, there is considerable room for improvement in the precise and accurate determination of photospheric properties of pre-main sequence stars.

Data on pre-main sequence stars (spectra) and their disks (resolved molecular gas and dust observations) has rapidly reached such a high quality, that more sophisticated models are warranted to explain all of the interesting features. However, model complexity is currently limited by computational expense. We discuss the application of Bayesian emulators, which are useful for exploring high-dimensional parameter spaces.
Chapter 2

A Disk-based Dynamical Mass Estimate for the Young Binary AK Sco

I. Czekala, S. M. Andrews, E. L. N. Jensen, K. G. Stassun, G. Torres, & D. J. Wilner


2.1 Introduction

Precise measurements of the physical properties of pre-main sequence (pre-MS) stars are fundamental to testing the theoretical predictions of stellar evolution models. Such models are in turn the basis for deriving a variety of interesting quantities in star-formation research, including the masses and ages of individual pre-MS stars from secondary properties (such as effective temperature and luminosity), the initial
CHAPTER 2. AK SCO

mass function (IMF) of star-forming regions, and the timescale for circumstellar disk evolution and planet formation, among many others. To successfully test and help refine theoretical models of pre-MS evolution, we require a sample of “benchmark” systems where the fundamental properties are known. The key constraint for such systems is a direct (i.e., dynamical) measurement of the stellar mass.

Unfortunately, few such benchmarks exist, particularly at the low end of the mass spectrum \((M_\star < 2 M_\odot)\) where the models are in greatest need of calibration (see, e.g., Hillenbrand & White 2004; Mathieu et al. 2007; Stassun et al. 2004, 2006, 2007, 2008; Gennaro et al. 2012). Eclipsing binaries (EBs) have long served as important empirical touchstones for testing stellar models. However, the recent review of pre-MS benchmarks by Stassun et al. (2014) identified only 21 low-mass EBs that have sufficiently precise measurements of their physical parameters to be suitable for testing evolutionary models. The same review also presented new evidence that many of the benchmark pre-MS EBs may have their temperatures and/or radii altered by the influence of tertiary companions which, while representing interesting physics in their own right, render them less suitable to direct tests of the evolutionary models that do not include such effects (see also Gómez Maqueo Chew et al. 2012).

Therefore, additional pre-MS stellar mass benchmarks are crucial. Ideally, these would include single stars, or binaries without the potentially complicating effects of a tertiary component. In particular, single or binary pre-MS stars with circumstellar or circumbinary disks offer the opportunity to dynamically measure the stellar masses using the Keplerian orbital motion (i.e., rotation curve) of the gas disk (e.g. Simon et al. 2000; Schaefer et al. 2009; Rosenfeld et al. 2012).
CHAPTER 2. AK SCO

To be sure, there are additional complications with the analysis of such systems, including accretion effects, isolating the individual component masses in the case of a close binary (since the dynamical mass is then the total binary mass), and a systematic (linear) dependence on the (still uncertain) distance. However, such systems are much more common than the intrinsically rare EBs, and therefore represent an important opportunity to significantly expand the sample of benchmark pre-MS stars. This is particularly compelling now, as ALMA has started to provide unprecedented sensitivity to molecular line emission from disks, and the upcoming GAIA mission is poised to give accurate distances to these systems. Just as important as the ongoing determination of empirical masses for more of these pre-MS systems is the development of robust procedures for their careful analysis, including sophisticated molecular line modeling approaches and the incorporation of state-of-the-art statistical methods for comparing the stellar properties to current theoretical pre-MS evolution models.

AK Sco represents a good case study for this purpose. AK Sco is a bright ($V \approx 8.9$), pre-MS double-lined spectroscopic binary (Andersen et al. 1989) associated with the nearby Upper Centaurus-Lupus star-forming region (Pecaut et al. 2012, $d \sim 140$ pc). It is actively accreting and has a massive circumbinary dust disk (Alencar et al. 2003). An orbital solution from long-term radial velocity monitoring of optical spectra identifies a short period ($\sim 13.6$ days), eccentric ($e = 0.47$), nearly equal-mass pair of F5 stars (Andersen et al. 1989; Alencar et al. 2003). Alencar et al. (2003) presented the most comprehensive analysis of the system to date, including an attempt to determine the stellar masses by constraining the (unknown) orbital inclination using geometric arguments (the stars are not known to eclipse) together with a simple model of the spectral energy distribution (SED). Recently, Anthonioz et al. (2015) resolved the inner
Figure 2.1.—: (a) Synthesized image of the 1.3 mm dust continuum emission. Contours start at 3σ and increase by factors of two. (b) The 0th moment map (velocity-integrated emission; contours) overlaid on the 1st moment map (intensity-weighted velocities; color scale) for the CO $J=2-1$ emission. Contour levels start at 3σ and increase in 10σ intervals. (c) and (d) The same as (b) for the $^{13}$CO and C$^{18}$O $J=2-1$ emission, respectively. Contours are spaced at 3σ intervals. The synthesized beam is shown in the lower left of each panel.

edge of the AK Sco circumbinary disk with the VLT interferometer. However, until now there has not been a dynamical measurement of the disk rotation curve with which to directly measure the binary mass.

Here we report the direct determination of the AK Sco binary mass through the dynamical measurement of its circumbinary disk rotation profile using new data from ALMA. Section 2.2 presents the ALMA observations and their calibration. Section 2.3 covers the methodology for modeling the disk gas rotation profile, and thereby the binary mass. Section 2.4 discusses these results together with an analysis of other basic stellar properties in the context of pre-MS evolutionary model predictions.
2.2 Observations and Data Reduction

AK Sco was observed using 32 ALMA antennas and the Band 6 receivers on 2014 April 10 as part of program 2012.1.00496.S (PI Andrews), with baseline lengths of 15–350 m. The ALMA correlator was configured to process four dual-polarization spectral windows. Two windows, covering the CO (230.538 GHz) and $^{13}$CO (220.399 GHz) $J=2−1$ transitions, spanned a 234 MHz bandwidth with 61 kHz (80 m s$^{-1}$) channels. A third window sampled C$^{18}$O $J=2−1$ (219.560 GHz) at twice that channel spacing (122 kHz, or 160 m s$^{-1}$). The remaining window used a coarse frequency resolution (128 channels of 15.628 MHz width) to probe the 232 GHz (1.3 mm) continuum. The observations alternated between AK Sco and the quasar J1709−3525 (3° separation) on a ∼7 minute cycle. Some additional brief observations of the bright quasar J1626−2951 and Titan were made for calibration purposes. The observing block lasted ∼1 hour, with half the time devoted to AK Sco.

The raw visibility data were calibrated using the facility software package CASA (v4.2). After applying the standard system temperature and water vapor radiometer corrections, the intrinsic passband shape was determined using the observations of J1626−2951 and removed. Complex gain variations due to instrumental and atmospheric effects were calibrated based on the regular monitoring of J1709−3525, and the overall amplitude scale was set using the observations of Titan. After a single iteration of phase-only self-calibration, the reduced visibilities were time-averaged into 30 s intervals. The local continuum level was subtracted from the spectral windows containing emission lines of interest.

Continuum and spectral line emission from all three of the targeted transitions were firmly detected. The calibrated visibilities were Fourier inverted with natural weighting,
deconvolved with the \texttt{CLEAN} algorithm, and subsequently restored with a FWHM = $0^\prime.75 \times 0^\prime.60$ synthesized beam (with P.A. = 100$^\circ$). Some representative data products are shown together in Figure 2.1.

The synthesized map of the 1.3 mm continuum emission (Fig. 1a) has an RMS noise level of $45 \mu \text{Jy beam}^{-1}$, and shows a marginally resolved morphology: an elliptical Gaussian fit to the visibilities indicates a total flux density of $32.65 \pm 0.07 \text{mJy}$ (a peak S/N $\approx 400$; an additional systematic uncertainty of $\sim 10\%$ is imposed by the absolute accuracy of the Titan emission model), with FWHM dimensions of $0^\prime.38(\pm 0.01) \times 0^\prime.12(\pm 0.01)$ oriented at a position angle of $49 \pm 1^\circ$. The CO $J=2-1$ line is detected over a velocity span of $\sim 24 \text{ km s}^{-1}$ (about 300 channels at the native resolution). It spans a diameter of $\sim 2^\prime$, and has an integrated intensity of $2.21 \pm 0.01 \text{ Jy km s}^{-1}$. For computational simplicity, we average these data into 60 channels of 305 kHz (0.4 km s$^{-1}$) width for further analysis. The RMS noise level at that resolution is 4 mJy beam$^{-1}$; the peak brightness is 170 mJy beam$^{-1}$ (peak S/N $\approx 45$ per beam and channel). With the same averaging, the $^{13}$CO and C$^{18}$O isotopologue transitions are also detected (albeit over a smaller velocity range and at lower significance). Their integrated intensities are $0.54 \pm 0.01$ and $0.20 \pm 0.01 \text{ Jy km s}^{-1}$ (with peak S/N $\approx 12$ and 5 per beam and channel), respectively.

\section*{2.3 CO Modeling and Results}

Our primary goal is to use the observed spectral visibilities that trace the CO $J=2-1$ emission line to quantitatively characterize the velocity field of the AK Sco gas disk, and thereby to measure the total mass of the central binary host. We employed a
forward-modeling approach that uses a parametric prescription for the disk structure (densities, temperatures, and dynamics), and then simulates the observed visibilities by assuming the molecular level populations are in local thermodynamic equilibrium (LTE) and propagating synthetic photons through the model structure with the \texttt{RADMC-3D} radiative transfer code.\footnote{ita.uni-heidelberg.de/~dullemond/software/radmc-3d/} Previous work with similar intentions (Simon et al. 2000; Piétu et al. 2007; Schaefer et al. 2009; Guilloteau et al. 2014) has shown that resolved measurements of the disk rotation curve can constrain the central stellar mass ($M_*$) with high precision ($\sim$ few percent, although with a linear dependence on distance). However, in practice this approach is complicated: it involves radiative transfer modeling with a large number of (unrelated) disk structure parameters. Given that complexity, it is important to develop tests to validate its absolute accuracy. Rosenfeld et al. (2012) used the circumbinary disk around the V4046 Sgr system to demonstrate that this approach is robust, in that it provides an $M_*$ estimate consistent with the constraints from spectroscopic monitoring of the stellar radial velocities. The AK Sco system provides another rare opportunity to test the methodology.

We adopted the two-dimensional (axisymmetric) parametric disk structure model that is described in detail by Rosenfeld et al. (2012). The temperature structure is vertically isothermal, and has a power-law radial distribution with index $q$ and a normalization at 10 AU ($T_{10}$). The radial surface density profile is the standard Lynden-Bell & Pringle (1974) similarity solution for a viscous accretion disk, essentially a power-law with an exponential taper at large radii. The gradient parameter $\gamma$ is fixed to unity, so the profile is described by a characteristic radius $r_c$ and a total CO mass $M_{\text{co}}$. 

\footnote{ita.uni-heidelberg.de/~dullemond/software/radmc-3d/}
The densities are distributed vertically under the assumption of hydrostatic equilibrium. The bulk velocity field of the gas is assumed to be Keplerian, and dominated by the binary mass (assuming \(M_{\text{disk}}/M_* \ll 1\)). The line width is calculated as the quadrature sum of thermal and non-thermal (i.e., turbulent) broadening terms, with the latter denoted as a constant velocity width \(\xi\). The physical structure of the model is fully characterized by six free parameters, \(\theta_{\text{disk}} = \{T_{10}, q, M_{\text{co}}, r_c, \xi, M_*\}\).

For a given set of these structure parameters, the RADMC-3D radiative transfer code was used to calculate the corresponding molecular excitation levels (appropriately assuming LTE for this transition; see Pavlyuchenkov et al. 2007) and ray-trace synthetic model spectral images at high resolution (2.5 AU pixels). To do so, we specified an additional set of parameters. The disk inclination \(i_d\) is defined with respect to the rotation axis, such that \(i_d = 0^\circ\) corresponds to a face-on viewing geometry with the rotation axis pointing toward the observer (and the disk rotating counterclockwise), \(i_d = 90^\circ\) is an edge-on orientation, and \(i_d = 180^\circ\) is again face-on but with the rotation axis pointing away from the observer (and therefore an apparent clockwise sense of rotation). The position angle \(\varphi\) represents the projection of the angular momentum vector of the disk onto the sky (as typical, oriented E of N). The disk center is characterized by offsets \((\delta_{\alpha}, \delta_\delta)\) relative to the phase center. Along with the distance \(d\) and systemic velocity \(v_{\text{sys}}\), there are six additional free parameters, \(\theta_{\text{obs}} = \{i_d, \varphi, \delta_{\alpha}, \delta_\delta, v_{\text{sys}}, d\}\).

The end result is a set of high resolution model images that specify the sky-projected CO \(J=2-1\) surface brightness as a function of position and frequency for any set of parameters \(\theta = \{\theta_{\text{disk}}, \theta_{\text{obs}}\}\). We then employed the FFTW algorithm (Frigo & Johnson 2005) to Fourier transform these spectral images, and then performed a band-limited interpolation of the complex visibilities onto the same spatial frequencies \((u, v)\) sampled
with ALMA (using the spheroidal gridding functions advocated by Schwab 1984) to acquire a set of model visibilities $M_{u,v}(\theta)$. To quantify the model quality given the data $D_{u,v}$, we computed a simple likelihood function

$$p(D|\theta) = \prod_{u,v} w_{u,v} |D_{u,v} - M_{u,v}(\theta)|^2,$$

where $w_{u,v}$ are the visibility weights (determined from the radiometer equation). Eq. (1) is identical to a log-likelihood function described by the sum of the standard $\chi^2$ values (real and imaginary) over all the observed spatial frequencies and velocity channels. To explore the posterior distribution of the model parameters, we employed the Markov Chain Monte Carlo (MCMC) technique with a Metropolis-Hastings (M-H) jump proposal.

We assumed uniform priors in all parameters except $d$, since the ALMA data do not constrain the distance. A trigonometric parallax for AK Sco is available from Hipparcos, although the considerable optical variability is problematic: the original catalog has $d = 144^{+38}_{-25}$ pc (Perryman et al. 1997), but the van Leeuwen (2007) revision argues for a closer value $(d = 102^{+26}_{-17}$ pc). Pecaut et al. (2012) measured a kinematic parallax corresponding to $d = 144 \pm 12$ pc. Recently, Anthonioz et al. (2015) combined astrometric and radial velocity monitoring to estimate $d = 141 \pm 7$ pc. We adopted a Gaussian prior on $d$ based on the weighted average of these latter two measurements, with mean 142 pc and $\sigma = 6$ pc.
Figure 2.2.—: Channel maps at 305 kHz (0.4 km s$^{-1}$) resolution showing the CO $J=2-1$ data, the corresponding model, and the residuals (the latter two imaged in the same way as the data). Contours are drawn at 3$\sigma$ (12 mJy beam$^{-1}$) intervals, and the synthesized beam dimensions are shown in the lower left corners of each set of channel maps. The kinematic local standard of rest (LSRK) velocities are labeled in the top left of each panel. The color within each channel map corresponds to the velocity sampled in the moment maps (Figure 2.1).
CHAPTER 2. AK SCO

<table>
<thead>
<tr>
<th>$\Delta \alpha$ [&quot;']</th>
<th>-2</th>
<th>-1.5</th>
<th>-1.1</th>
<th>0</th>
<th>0.3</th>
<th>0.5</th>
<th>0.8</th>
<th>1.3</th>
<th>1.7</th>
<th>2.1</th>
<th>2.5</th>
<th>2.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6.3</td>
<td>3.5</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
<td>5.3</td>
<td>5.7</td>
<td>6.1</td>
<td>6.5</td>
<td>6.9</td>
<td>7.3</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>-5.9</td>
<td>3.7</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
<td>5.3</td>
<td>5.7</td>
<td>6.1</td>
<td>6.5</td>
<td>6.9</td>
<td>7.3</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>-5.5</td>
<td>3.5</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
<td>5.3</td>
<td>5.7</td>
<td>6.1</td>
<td>6.5</td>
<td>6.9</td>
<td>7.3</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>-5.1</td>
<td>2.9</td>
<td>3.5</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
<td>5.3</td>
<td>5.7</td>
<td>6.1</td>
<td>6.5</td>
<td>6.9</td>
<td>7.3</td>
<td>7.7</td>
</tr>
<tr>
<td>-4.7</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
<td>5.3</td>
<td>5.7</td>
<td>6.1</td>
<td>6.5</td>
<td>6.9</td>
<td>7.3</td>
</tr>
<tr>
<td>-4.3</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
<td>5.3</td>
<td>5.7</td>
<td>6.1</td>
<td>6.5</td>
<td>6.9</td>
</tr>
<tr>
<td>-3.9</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
<td>5.3</td>
<td>5.7</td>
<td>6.1</td>
<td>6.5</td>
</tr>
<tr>
<td>-3.5</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
<td>5.3</td>
<td>5.7</td>
<td>6.1</td>
</tr>
<tr>
<td>-3.1</td>
<td>0.8</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
<td>5.3</td>
<td>5.7</td>
</tr>
<tr>
<td>-2.7</td>
<td>0.3</td>
<td>0.8</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
<td>5.3</td>
</tr>
<tr>
<td>-2.3</td>
<td>0.5</td>
<td>0.8</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
<td>5.3</td>
</tr>
<tr>
<td>-1.9</td>
<td>-0.2</td>
<td>0.3</td>
<td>0.8</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
<td>4.5</td>
<td>4.9</td>
</tr>
<tr>
<td>-1.5</td>
<td>-0.7</td>
<td>-0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>-1.1</td>
<td>-0.7</td>
<td>-0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>0</td>
<td>-0.7</td>
<td>-0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>0.3</td>
<td>-0.7</td>
<td>-0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>0.5</td>
<td>-0.7</td>
<td>-0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>0.8</td>
<td>-0.7</td>
<td>-0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>1.3</td>
<td>-0.7</td>
<td>-0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>1.7</td>
<td>-0.7</td>
<td>-0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>2.1</td>
<td>-0.7</td>
<td>-0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>2.5</td>
<td>-0.7</td>
<td>-0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>2.9</td>
<td>-0.7</td>
<td>-0.3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>2.1</td>
<td>2.5</td>
<td>3.1</td>
<td>3.5</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Figure 2.2.—
After a few preliminary chains were run, we empirically tuned the covariance of the M-H jumps to match the morphology of the posterior distribution, yielding a more efficient exploration of parameter space. To determine the final posterior distributions, we ran multiple independent chains using different initializations. After a conservative period of burn-in, we computed the Gelman-Rubin statistic (Gelman et al. 2013) for each parameter over the ensemble of chains to assess convergence.\(^2\)

The parameter values inferred from modeling the observed spectral visibilities binned to 305 kHz (0.4 km s\(^{-1}\)) resolution are listed in Table 1. A trial fit of the data at the best available (quasi-independent) spectral resolution of 122 kHz (0.16 km s\(^{-1}\)) recovered these same values. The disk structure parameters are typical for similar Class II systems, although the characteristic radius and mass are on the small side. Assuming a standard CO/H\(_2\) abundance ratio (\(\sim 10^{-4}\)), the total gas mass would be \(\sim 0.007 \, M_\odot\) – roughly consistent with what would be inferred from the continuum emission (e.g., Andrews et al. 2013). The inferred total mass of the central binary is 2.49 ± 0.10 \(M_\odot\), where the uncertainty is dominated by the distance prior. At a fixed \(d\), the constraint can be framed as \(M_*/d = 0.01731 \pm 0.00015 \, M_\odot \, pc^{-1}\); i.e., the formal precision on the stellar mass is \(\sim 1\%\) for a \(\delta\)-function prior on \(d\). Figure 2.2 shows a direct comparison of the data and best-fit model in the image plane, demonstrating (through the absence of significant residuals) the fit quality. The reduced \(\chi^2\) of the best-fit model is 1.08.

\(^2\)The code used to perform the analysis described here is open source and freely available under an MIT license at https://github.com/iancze/DiskJockey.
Figure 2.3.—: The (marginal) joint posterior probability distributions for \( \{M_*, \text{inclination}\} \) and \( \{M_*, d\} \). For reference, we have overlaid the nominal parameter degeneracies (±1 \( \sigma \)) that accompany each measurement in isolation, assuming no external constraints about either parameter but precise knowledge of all other parameters.
Table 2.1: Inferred Parameters for AK Sco

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{\text{10}}$ (K)</td>
<td>92 ± 4</td>
<td>$i_d$ (°)</td>
<td>109.4 ± 0.5</td>
</tr>
<tr>
<td>$q$</td>
<td>0.51 ± 0.01</td>
<td>$\varphi$ (°)</td>
<td>141.1 ± 0.3</td>
</tr>
<tr>
<td>log $M_{\text{co}}$ ($M_\odot$)</td>
<td>-0.65 ± 0.22</td>
<td>$v_{\text{sys}}$ (km s$^{-1}$)</td>
<td>+5.49 ± 0.06</td>
</tr>
<tr>
<td>$r_c$ (AU)</td>
<td>14.3 ± 1.2</td>
<td>$\delta_\alpha$ (&quot;)</td>
<td>0.053 ± 0.002</td>
</tr>
<tr>
<td>$M_*$ ($M_\odot$)</td>
<td>2.49 ± 0.10</td>
<td>$\delta_\delta$ (&quot;)</td>
<td>0.045 ± 0.002</td>
</tr>
<tr>
<td>$\xi$ (km s$^{-1}$)</td>
<td>0.31 ± 0.02</td>
<td>$d$ (pc)</td>
<td>143.6 ± 5.7</td>
</tr>
</tbody>
</table>

The quoted values represent the “best-fit”, the peaks of the marginal posterior distributions. The uncertainties correspond to the 68.3% ($\sim 1\sigma$) confidence intervals. $v_{\text{sys}}$ is reported in the LSRK frame, for the standard radio definition. The corresponding heliocentric value is $-1.92 \pm 0.06$, consistent with the Alencar et al. (2003) value derived from optical spectroscopy.
2.4 Discussion

We have presented ALMA Cycle 1 observations around a wavelength of 1.3 mm that resolve the dust and molecular line emission from the disk that orbits the double-lined spectroscopic binary AK Sco. To our knowledge, these are the first observations that confirm the presence of a substantial molecular gas reservoir in this system. The disk itself is relatively small (a characteristic radius of $\sim 14$ AU), but still contains a modest mass ($\sim 5$–$10 M_{\text{Jup}}$ in total, for standard assumptions about opacities and abundances).

We have focused on a detailed modeling analysis of the spatially and spectrally resolved CO $J=2$–$1$ line emission, to map out the disk velocity field and make a dynamical estimate of the total mass of the host binary. Adopting a well-motivated prior on the distance ($142 \pm 6$ pc; see Sect. 3), we inferred a precise (4%) combined stellar mass estimate given the ALMA spectral visibilities, $M_* = 2.49 \pm 0.10 M_\odot$.

2.4.1 Comparison to Previous Results

There are two independent dynamical constraints on the AK Sco binary mass available in the literature. The first comes from an optical spectroscopic campaign to monitor the radial velocity (RV) variations of the double-lined system, performed by Alencar et al. (2003). They measured a precise orbital period ($\sim 13.6$ d) and mass ratio ($0.987 \pm 0.005$), a substantial eccentricity ($e_* = 0.471 \pm 0.002$) and a total mass constraint $M_* \sin^3 i_* = 2.114 \pm 0.010 M_\odot$ where $i_*$ is the (unknown) inclination of the binary orbit.

These measurements do not depend on $d$. If we assume the gas disk and binary orbits are co-planar ($i_* = i_d$), then our constraint on $i_d$ would convert the RV constraint on the binary mass to $2.52 \pm 0.03 M_\odot$, identical (within 1%) to our disk-based dynamical mass estimate.
estimate. These two dynamical mass constraints remain consistent (at the 1σ level) so long as the binary and disk orbital planes are aligned within ~2°, compatible with the statistical constraints on the mutual inclination distributions found for exoplanet systems (Tremaine & Dong 2012; Figueira et al. 2012; Fabrycky et al. 2014). Additionally, by combining the distance-independent measurements of $M_*/d$ and $M_* \sin^3 i_*$, we derive a dynamical distance to the system of $d = 145.5 \pm 2.0$ pc, which is consistent with our initial assumptions in the form of the distance prior.

This consistency between mass constraints, along with the similar results for the V4046 Sgr binary (Rosenfeld et al. 2012), effectively validates the absolute accuracy of the disk-based dynamical mass measurement technique. Whereas the V4046 Sgr measurement involved a large circumstellar disk ($r_c = 45$ AU), we have demonstrated that—given the sensitivity and resolution of ALMA—the dynamical mass technique also works for much smaller disks like AK Sco ($r_c = 14$ AU). In the coming age of ALMA, this should be the workhorse approach for determining masses for large, statistically relevant, samples of young stars. Most importantly, it is the only method capable of performing that task for single stars.

The second mass constraint comes from the recent work by Anthonioz et al. (2015), which combined multi-epoch $H$-band interferometric observations with the Alencar et al. (2003) RV data. In the context of a simple model that includes the binary and a narrow dust ring, these astrometric constraints suggest inclinations of $i_* = 115 \pm 3°$ and $i_{\text{ring}} = 121 \pm 8°$, respectively, which are ~1.8 and 1.4σ larger than the $i_d$ we infer from the ALMA data. In conjunction with the RV constraints, this larger orbital inclination suggests a higher binary mass, $M_* = 2.80 \pm 0.11 M_\odot$, a 2σ discrepancy with respect to the disk-based mass estimate. For reference, Figure 2.3 shows the mass constraints as a
function of both inclination and distance.

There are various scenarios that might alleviate this (modest) tension. The difference between $i_{\text{ring}}$ and $i_d$ might be explained with a warp in the disk structure, which would perturb the line-of-sight projection of the velocity field measured in the CO line and thereby accommodate a higher $M_*$ estimate from the ALMA data (e.g., Rosenfeld et al. 2012, 2014; Marino et al. 2015). Given the eccentricity of the binary, one might also expect that the inclination estimates could be biased due to the assumptions of circular disk and ring structures. For the Anthonioz et al. results, an imperfect ring model can affect the estimate of $i_*$ since the binary itself is not well-resolved (the semi-major axis of the apparent orbit is only 1 mas, about one third of the resolution). We also explored an eccentric disk model, where the structure is built up from a series of infinitesimal apse-aligned elliptical rings: the ALMA data rule out a significant (mean) eccentricity in the gas disk, with $e_d < 0.04$ (at 99.7% confidence). In the end, it may just be that the (admittedly) simplistic ring model adopted in the preliminary analysis of Anthonioz et al. (2015) can be improved (the quoted reduced $\chi^2$ of their fit is $\sim 2$), and such modifications would bring the results into agreement. Ultimately, it would be interesting to combine all the measurements in a joint analysis.

2.4.2 Comparison to Pre-MS Evolution Models

To compare the disk-based dynamical mass with predictions from pre-MS evolution models, we performed the standard analysis of estimating stellar parameters from the Hertzsprung-Russell (H-R) diagram. The $UBVRI$ photometry compilation of Jensen & Mathieu (1997) was used to construct the AK Sco SED. The near-infrared is
CHAPTER 2. AK SCO

contaminated by the dust disk, and so was excluded from our analysis. A base parametric SED model was constructed from synthetic photometry in the BT-SETTL catalog of stellar models (Allard et al. 2003), interpolated for any given effective temperature \(T_{\text{eff}}\) and surface gravity \((\log g)\) assuming a fixed solar metallicity. This base model was adjusted for extinction using the Fitzpatrick (1999) reddening curve with \(R_V = 4.3\) (as determined by Manset et al. (2005)) and then scaled by the squared ratio of the radius \((R_*)\) and distance \((d)\). We assumed the components of the AK Sco binary are identical.

This SED model has five free parameters, \(\theta_{\text{sed}} = \{T_{\text{eff}}, \log g, A_V, R_*, d\}\), but the photometric data alone cannot uniquely constrain all of them. To aid in the inference of the stellar properties, we imposed simple Gaussian priors on \(d\) (as in Sect. 3) and \(T_{\text{eff}}\). The latter was based on the F5 spectral classification and de-reddened color indices (Andersen et al. 1989), which we associate with \(T_{\text{eff}} = 6450 \pm 150\) K (e.g., Bessell 1979; Popper 1980; Gray 2005; Casagrande et al. 2010; Pecaut & Mamajek 2013). AK Sco is highly variable, with erratic changes much larger than the uncertainties on individual measurements (e.g., Andersen et al. 1989; Alencar et al. 2003). To deal with that variability, we incorporated a nuisance “jitter” parameter \((\sigma)\) at each band that characterizes the additional dispersion (assuming a Gaussian distribution). With this parametric model setup, we explored the posterior distribution of \(\theta = \{\theta_{\text{sed}}, \sigma_U, \sigma_B, \sigma_V, \sigma_R, \sigma_I\}\) conditioned on the SED data using MCMC with the ensemble sampler emcee (Foreman-Mackey et al. 2013).

The modeling results are shown in Figure 2.4. We found \(T_{\text{eff}} = 6365 \pm 155\) K, \(\log g = 3.5 \pm 0.5\), \(R_* = 1.43 \pm 0.07\) \(R_\odot\), and \(A_V = 0.70 \pm 0.1\) mag for each star (and \(d = 142 \pm 6\) pc, as expected given the prior). The corresponding luminosity of each component is \(L_* = 3.0 \pm 0.5\) \(L_\odot\). The variability dispersion terms, in terms of a flux
Figure 2.4.—: left: The SED compared with the best-fit model. The residuals are shown overlaid with the inferred “jitter” terms in each band to account for the systematic variance from variability. right top: The $R_*$, $T_{\text{eff}}$ version of the H-R diagram, with the marginal posteriors inferred from the SED modeling shown as 1 and 2 $\sigma$ contours. The Dotter et al. (2008) pre-MS model mass tracks are overlaid, with color scales indicating different ages (note that these values refer to each individual star in the AK Sco binary). right bottom: The joint mass and age constraints from various pre-MS models, shown as 1 $\sigma$ contours and marginalized distributions. The gray band marks the disk-based constraint on $M_*$ (1 $\sigma$). The Baraffe et al. (2015) models do not explore the parameter space $M_* > 1.4M_\odot$. 
density fraction, range from 0.1 \((I\text{-band})\) to 0.2 \((U\text{-band})\).

Models of pre-MS evolution predict the joint behavior of \(T_{\text{eff}}\) and \(R_*\) (or equivalently \(L_*\)) as a function of \(M_*\) and age \((\tau_*)\). Following the Bayesian formalism of Jørgensen \\& Lindegren (2005), we mapped the posterior constraints from the SED modeling into an inference on \(\{M_*, \tau_*\}\) for several evolutionary model grids (for some practical examples, see Rosenfeld et al. 2012; Andrews et al. 2013). This approach is also illustrated in Figure 2.4. The AK Sco binary mass and age inferred from these model grids (Siess et al. 2000; Dotter et al. 2008; Tognelli et al. 2011; Baraffe et al. 2015) are listed in Table 2. Their weighted means and standard deviations are \(M_* = 2.68 \pm 0.03 \, M_\odot\) and \(\tau_* = 18 \pm 1\) Myr. These model-predicted masses are consistent, but are all \(\sim 1.5–2\sigma\) higher than the dynamical mass inferred from the ALMA data. The ages are also consistent with the age of the Upper Centaurus-Lupus association \((\langle \tau_* \rangle = 16 \pm 1\) Myr; Pecaut et al. 2012).

The 7–10\% discrepancy between the pre-MS model masses and the disk-based dynamical mass is slightly larger than the typical level of disagreement noted for young EBs in this mass range (Stassun et al. 2014). We suspect this modest mismatch might be attributed to the complexity of the AK Sco binary environment. The AK Sco stars are separated by only \(\sim 11\, R_*\) at periastron (every \(\sim 2\) weeks), which leads to accretion bursts onto the stellar surface (Gómez de Castro et al. 2013). Perhaps related, the AK Sco stars have unusually broad ultraviolet lines that modulate with the binary period and are indicative of perturbed, hot \(\sim 60,000\, \text{K}\) “atmospheres” that extend out to 5 \(R_*\) (Gómez de Castro 2009). This behavior is compounded by the erratic variability noted in broadband photometry and the Balmer emission lines (Alencar et al. 2003), suggesting additional complexity in the accretion/outflow behavior. This dynamic, complicated
environment likely impacts the physical structures and evolution of the stars, and is
(obviously) not included in the pre-MS models. Alternatively, and perhaps more likely,
the standard means of estimating parameters like $T_{\text{eff}}$ in such a situation should also be
affected: a small shift ($< 100$ K cooler) would bring the measurements into agreement.

2.5 Summary and Conclusions

We have analyzed ALMA observations of the CO $J=2–1$ transition from the AK Sco
circumbinary disk. The main conclusions of this work include:

- A relatively compact disk in orbit around the AK Sco binary, with $\sim 5–10$ $M_{\text{Jup}}$ of
gas and dust, is detected in the 1.3 mm continuum and main isotopologues of CO (the
$^{12}$CO, $^{13}$CO, and $^{18}$O $J=2–1$ transitions). This suggests an unexpectedly long-lived
($\sim 18$ Myr) disk of primordial origin, as opposed to a second-generation debris disk.

- We employed a parametric disk structure model front-end fed into the radiative
transfer code RADMC-3D to generate synthetic spectral visibilities to compare with the
observations in the MCMC framework. The related software is provided as an open source
resource to the community. The results offer a high-quality dynamical measurement
of the binary mass, $M_*= 2.49\pm 0.10$ $M_{\odot}$, that is independent of pre-MS evolution models.

- This disk-based dynamical mass estimate is in good agreement with the constraints
from radial velocity monitoring of the binary, so long as the disk and binary orbital
planes are aligned within $\sim 2^\circ$. There is minor tension with a recent combined astrometric + radial velocity analysis, although we expect the comparison could be improved with a consistent joint analysis.

• With the standard approach of estimating stellar parameters from the H-R diagram, we make comparisons between pre-MS evolutionary model predictions and the dynamical mass estimated here. These models suggest a slightly higher (by 7–10%) stellar mass for AK Sco, in modest disagreement (at 1.5–2 $\sigma$) with our results: this discrepancy may be attributed to the complicated accretion and interaction environment of the binary. The model-dependent ages ($18 \pm 1$ Myr) are consistent with the proposed AK Sco membership in the Upper Centaurus Lupus association.

• The overall consistency between $M_*$ estimates for AK Sco validates the absolute accuracy of the disk-based dynamical mass technique. This method has great promise in the ALMA era, since it is uniquely capable of providing precise (few %) masses of statistically large samples of single pre-MS stars that can be used to test and calibrate models of early stellar evolution.
Table 2.2. Evolutionary Model Predictions

<table>
<thead>
<tr>
<th>Model Grid</th>
<th>$M_\ast [M_\odot]$</th>
<th>$\tau_\ast$ [Myr]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siess et al. (2000)</td>
<td>2.70 $\pm$ 0.06</td>
<td>19 $\pm$ 3</td>
</tr>
<tr>
<td>Tognelli et al. (2011)</td>
<td>2.71 $\pm$ 0.07</td>
<td>18 $\pm$ 2</td>
</tr>
<tr>
<td>Dotter et al. (2008)</td>
<td>2.60 $\pm$ 0.07</td>
<td>17 $\pm$ 2</td>
</tr>
<tr>
<td>Baraffe et al. (2015)</td>
<td>2.68 $\pm$ 0.05</td>
<td>19 $\pm$ 2</td>
</tr>
</tbody>
</table>

Note. — The quoted uncertainties correspond to the 68.3% ($\sim 1 \sigma$) confidence intervals.
Chapter 3

A Disk-based Dynamical Mass Estimate for the Young Binary DQ Tau


3.1 Introduction

Theoretical models for pre-main sequence (pre-MS) stellar evolution are fundamental tools for learning about star and planet formation. But the accuracy of such models—especially at young ages—is unclear, due to our limited understanding of some complex physical effects like accretion (e.g., Baraffe et al. 2009) or magnetic fields and convection
(e.g., Feiden & Chaboyer 2013). While such issues are being explored theoretically, robust observational constraints on key stellar parameters can be used to help guide improvements to the models. Most important are dynamical constraints on stellar masses, $M_*$ (e.g., see Hillenbrand & White 2004). Usually these are determined from the orbital motions of binary systems (Stassun et al. 2014), but they could increasingly be measured for single stars based on the rotation of their associated gas disks (e.g., Simon et al. 2000).

We are interested in comparing the constraints from these two approaches, to illuminate and quantify any associated systematic problems in the inference of $M_*$. To do that, we have targeted the few roughly equal-mass double-lined spectroscopic binaries that host circum-binary disks, including V4046 Sgr (Rosenfeld et al. 2012) and AK Sco (Czekala et al. 2015). In both cases, excellent agreement (to within $\sim 1\%$) is found between the estimates of $M_*$ from radial velocity monitoring of the stars and the tomographic reconstruction of the CO gas velocity field in the disk. The confluence of these measurements also indicates that the disk and binary orbital planes are well-aligned (within 1–2$^\circ$). Moreover, the predictions of theoretical pre-MS models faithfully reproduce these results for these two particular examples. However, these model successes are perhaps not surprising, since both V4046 Sgr and AK Sco are relatively old (10 and 18 Myr) and massive (1.8 and 2.5 $M_\odot$ total mass in each system) and the models should be more robust in that range of age and mass. An important supplementary test would employ a cooler and younger binary.

In those respects, DQ Tau is an exemplary target. DQ Tau is a roughly equal-mass double-lined spectroscopic binary with a period of $\sim 16$ days and a notably eccentric orbit (Mathieu et al. 1997). It has a composite spectral type of $\sim$M0–M1 (e.g., Herbig 1977;
CHAPTER 3. DQ TAU

Herczeg & Hillenbrand 2014) and is located in the nearby and relatively young Taurus clouds. DQ Tau exhibits enhancements of various tracers of accretion and activity — optical brightening (Mathieu et al. 1997), emission line variations (Basri et al. 1997), and mm/radio emission (Salter et al. 2010) — that have been associated with both pulsed accretion and reconnection events from colliding magnetospheres near peri-astron. The DQ Tau binary hosts a circumbinary disk with substantial millimeter continuum emission from dust (Beckwith et al. 1990; Andrews & Williams 2005; Guilloteau et al. 2011). There is recent evidence for molecular gas in rotation around the central binary host (Williams & Best 2014), although there is non-negligible contamination from the local molecular cloud (Guilloteau et al. 2013).

We present new observations of molecular gas in the DQ Tau circumbinary disk made with the Atacama Large Millimeter/Submillimeter Array (ALMA), and use them to place a dynamical constraint on the total mass of the DQ Tau binary. We also provide an updated orbital solution for the DQ Tau binary based on long-term radial velocity monitoring. Section 3.2 presents the data and its calibration. Section 3.3 describes our modeling of the gas disk velocity field, provides an update of the original Mathieu et al. (1997) orbital solution, and highlights the key results. Section 3.4 discusses these results together and assesses the predictions of pre-MS evolution models. And Section 3.5 provides a summary in the contexts of other young circumbinary disk systems and the utility of the disk-based dynamical mass technique.
3.2 Observations and Data Reduction

3.2.1 Millimeter Interferometry

ALMA observed the DQ Tau system on 2015 May 24, using 34 of its 12 m antennas with separations ranging from 21 to 540 m. The observations were configured with the same spectral setup as in Czekala et al. (2015), employing the Band 6 receivers to cover the CO, $^{13}$CO, and C$^{18}$O J=2−1 transitions in 61, 61, and 122 kHz channels, respectively, as well as the adjacent continuum (at 232 GHz, or 1.3 mm). The nearby quasar J0510+1800 (6° separation) was observed regularly to monitor variations in the complex gain response of the interferometer. The bright quasar J0423−0120 was also observed to calibrate the bandpass behavior and absolute flux levels. The total on-source integration time was \( \sim \) 28 minutes. The visibilities were calibrated using standard techniques with the CASA software package (v4.3). After a phase and amplitude self-calibration based on the bright continuum, the spectral line visibilities were time-averaged (to 30 s intervals) and continuum-subtracted.

Images of the continuum and spectral line data were created by Fourier inverting the calibrated visibilities (assuming a Briggs robust weighting parameter of 0.5, to balance S/N and resolution), deconvolving with the CLEAN algorithm, and restoring with a synthesized beam with FWHM = 0′′8 × 0′′6 (at P.A. = 145°). The continuum image shows a bright, marginally resolved source with a peak intensity of \( \sim 68 \) mJy beam\(^{-1}\) and integrated flux density of 79 mJy. The RMS noise level is \( 70 \) µJy beam\(^{-1}\) (the peak S/N is \( \sim 1000 \); the map sensitivity is clearly limited by dynamic range). All of this emission is expected to be from dust; the peri-apse continuum enhancement noted by Salter et al.
is not present at the observed orbital phase ($\phi = 0.71$). But the focus here is on the emission from the CO spectral lines.

The $^{12}$CO (hereafter CO) and $^{13}$CO $J=2-1$ transitions were imaged in 0.1 km s$^{-1}$-wide channels, and reach an RMS noise level of 8 mJy beam$^{-1}$ in each channel. Line emission from these transitions is detected over a $\sim$7 km s$^{-1}$-wide velocity range, exhibiting the classical morphological pattern of Keplerian rotation. The peak S/N is 35 for CO, but only 7 for $^{13}$CO. Both of these transitions show considerable contamination from the local molecular cloud material, affecting a 2 km s$^{-1}$-wide span slightly blueshifted from the systemic velocity. The $^{18}$O $J=2-1$ transition was imaged in 0.2 km s$^{-1}$-wide channels, with an RMS of $\sim$4 mJy beam$^{-1}$ in each, but the line is only marginally detected (S/N$\sim$3) in a few of these channels. Given the line intensities, our focus will be on an interpretation of the CO emission. The CO channel maps are shown in the top portion of Figure 3.1.

3.2.2 Optical Spectroscopy

Three sets of optical spectroscopic observations, including material also used by Mathieu et al. (1997), were used to re-examine the orbital solution of the DQ Tau binary. The first set consists of 30 spectra obtained at the Harvard-Smithsonian Center for Astrophysics (CfA) between 1984 and 2005 with two similar instruments equipped with intensified photon-counting Reticon detectors, as described in more detail by the above authors. These spectrographs are no longer in operation. A subset of 23 of these spectra was used by Mathieu et al. (1997); we have reanalyzed all 30 of them here with improved techniques. These single-order spectra (45Å centered around the Mg I b triplet near
5190 Å) have relatively low S/N, ranging from 6 to 16 per 8.5 km s\(^{-1}\) resolution element. A second set of 22 observations consists of radial velocity differences measured from spectra also described by Mathieu et al. (1997) and collected with instruments at the Lick Observatory, the Keck Observatory, and the McDonald Observatory. Finally, more recently (2013 October to December) we obtained three additional spectra of DQ Tau at the CfA for a different purpose, with the 1.5 m Tillinghast reflector at the Fred L. Whipple Observatory on Mount Hopkins (AZ). For this we used the bench-mounted TRES instrument (Fürész 2008) that delivers a resolving power of \( R \approx 44,000 \) in 51 echelle orders spanning the wavelength range 3900–9100 Å. These three spectra have signal-to-noise ratios in the Mg I\( b \) region of 16, 26, and 23 per 6.8 km s\(^{-1}\) resolution element.

Radial velocities (RVs) for each component of DQ Tau were measured from all of the CfA spectra using the two-dimensional cross-correlation technique TODCOR (Zucker & Mazeh 1994), with templates taken from a library of synthetic spectra based on PHOENIX model atmospheres (see Husser et al. 2013) computed for the appropriate instrumental resolution. Based on indications from the work of Mathieu et al. (1997) that the mass ratio is close to unity, we assumed the stars have the same temperature. Adopting solar metallicity, we experimented with templates of fixed surface gravities from \( \log g = 3.5 \) to 4.5. The best matches to the DQ Tau spectra were found for temperatures of \( \sim 4000 \) K (although a relatively wide range of \( \pm 300 \) K around that value is permissible) and \( v \sin i \) values of 14 and 11 km s\(^{-1}\) for the primary and secondary, respectively. These latter values are similar to the measurements of Nguyen et al. (2012), who obtained 14.7 ± 1.6 and 11.3 ± 0.7 km s\(^{-1}\).
Figure 3.1.—: CO $J=2-1$ channel maps for the DQ Tau data (top), the best-fit model (middle), and the imaged residuals (bottom) at 0.1 km s$^{-1}$ velocity resolution. Contours are drawn at intervals of 3× the RMS noise level (9.5 mJy beam$^{-1}$). The synthesized beam is drawn in the lower left corner of each set of channel maps, and the LSRK velocities are labeled in each panel. We do not model the channels with $v = 7.9 - 10.1$ km s$^{-1}$ (inclusive) because of cloud contamination; these channel maps are marked with an X in the lower panel.
Figure 3.1.—
### CHAPTER 3. DQ TAU

Table 3.1: Heliocentric radial velocity measurements of DQ Tau from CfA.

<table>
<thead>
<tr>
<th>HJD (2,400,000+)</th>
<th>Orbital phase</th>
<th>(RV_1) ((\text{km s}^{-1}))</th>
<th>(\sigma_1) ((\text{km s}^{-1}))</th>
<th>((O - C)_1) ((\text{km s}^{-1}))</th>
<th>(RV_2) ((\text{km s}^{-1}))</th>
<th>(\sigma_2) ((\text{km s}^{-1}))</th>
<th>((O - C)_2) ((\text{km s}^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>45982.0293</td>
<td>0.1435</td>
<td>34.76</td>
<td>3.93</td>
<td>−2.71</td>
<td>12.44</td>
<td>2.53</td>
<td>+1.75</td>
</tr>
<tr>
<td>46389.8446</td>
<td>0.9520</td>
<td>−2.93</td>
<td>3.61</td>
<td>−0.57</td>
<td>56.13</td>
<td>2.32</td>
<td>+2.89</td>
</tr>
<tr>
<td>46745.8122</td>
<td>0.4794</td>
<td>37.11</td>
<td>7.72</td>
<td>+6.50</td>
<td>13.90</td>
<td>4.97</td>
<td>−4.12</td>
</tr>
<tr>
<td>47073.9364</td>
<td>0.2446</td>
<td>37.38</td>
<td>3.56</td>
<td>+0.49</td>
<td>15.01</td>
<td>2.29</td>
<td>+3.69</td>
</tr>
<tr>
<td>47075.9910</td>
<td>0.3747</td>
<td>34.73</td>
<td>3.35</td>
<td>+0.97</td>
<td>17.06</td>
<td>2.16</td>
<td>+2.41</td>
</tr>
<tr>
<td>47078.0084</td>
<td>0.5023</td>
<td>31.64</td>
<td>3.23</td>
<td>+1.79</td>
<td>20.70</td>
<td>2.08</td>
<td>+1.87</td>
</tr>
<tr>
<td>47127.8232</td>
<td>0.6549</td>
<td>26.07</td>
<td>4.56</td>
<td>+2.22</td>
<td>24.40</td>
<td>2.94</td>
<td>−0.84</td>
</tr>
<tr>
<td>47128.8081</td>
<td>0.7172</td>
<td>23.51</td>
<td>5.11</td>
<td>+2.88</td>
<td>27.30</td>
<td>3.29</td>
<td>−1.38</td>
</tr>
<tr>
<td>47159.6132</td>
<td>0.6667</td>
<td>20.10</td>
<td>4.96</td>
<td>−3.19</td>
<td>25.73</td>
<td>3.19</td>
<td>−0.11</td>
</tr>
<tr>
<td>47159.6318</td>
<td>0.6679</td>
<td>17.31</td>
<td>4.46</td>
<td>−5.92</td>
<td>27.86</td>
<td>2.87</td>
<td>+1.96</td>
</tr>
<tr>
<td>47427.9265</td>
<td>0.6468</td>
<td>27.28</td>
<td>3.19</td>
<td>+3.06</td>
<td>23.67</td>
<td>2.06</td>
<td>−1.17</td>
</tr>
<tr>
<td>47546.6122</td>
<td>0.1579</td>
<td>36.99</td>
<td>5.66</td>
<td>−0.66</td>
<td>7.64</td>
<td>3.64</td>
<td>−2.86</td>
</tr>
<tr>
<td>47546.6272</td>
<td>0.1588</td>
<td>33.17</td>
<td>5.28</td>
<td>−4.48</td>
<td>10.94</td>
<td>3.40</td>
<td>+0.44</td>
</tr>
<tr>
<td>47546.6272</td>
<td>0.1588</td>
<td>35.11</td>
<td>3.15</td>
<td>−2.54</td>
<td>9.87</td>
<td>2.02</td>
<td>−0.63</td>
</tr>
<tr>
<td>47546.6423</td>
<td>0.1598</td>
<td>35.24</td>
<td>6.15</td>
<td>−2.42</td>
<td>10.05</td>
<td>3.96</td>
<td>−0.44</td>
</tr>
<tr>
<td>47789.9412</td>
<td>0.5569</td>
<td>28.81</td>
<td>2.92</td>
<td>+0.89</td>
<td>19.93</td>
<td>1.88</td>
<td>−0.96</td>
</tr>
<tr>
<td>47840.8048</td>
<td>0.7758</td>
<td>12.43</td>
<td>3.01</td>
<td>−4.45</td>
<td>33.26</td>
<td>1.94</td>
<td>+0.57</td>
</tr>
<tr>
<td>47845.8709</td>
<td>0.0964</td>
<td>40.09</td>
<td>2.92</td>
<td>+4.81</td>
<td>13.83</td>
<td>1.88</td>
<td>+0.79</td>
</tr>
<tr>
<td>47896.7312</td>
<td>0.3151</td>
<td>35.14</td>
<td>3.12</td>
<td>−0.19</td>
<td>11.47</td>
<td>2.00</td>
<td>−1.51</td>
</tr>
<tr>
<td>47898.7686</td>
<td>0.4440</td>
<td>35.89</td>
<td>3.31</td>
<td>+4.16</td>
<td>16.98</td>
<td>2.13</td>
<td>+0.16</td>
</tr>
<tr>
<td>47902.6961</td>
<td>0.6926</td>
<td>19.12</td>
<td>3.35</td>
<td>−2.86</td>
<td>26.86</td>
<td>2.16</td>
<td>−0.38</td>
</tr>
<tr>
<td>48525.9580</td>
<td>0.1356</td>
<td>37.08</td>
<td>3.27</td>
<td>−0.22</td>
<td>10.20</td>
<td>2.10</td>
<td>−0.67</td>
</tr>
<tr>
<td>48670.6445</td>
<td>0.2920</td>
<td>36.83</td>
<td>3.41</td>
<td>+0.95</td>
<td>14.36</td>
<td>2.19</td>
<td>+1.97</td>
</tr>
<tr>
<td>48871.9572</td>
<td>0.0321</td>
<td>12.75</td>
<td>3.41</td>
<td>−7.74</td>
<td>23.86</td>
<td>2.19</td>
<td>−4.97</td>
</tr>
<tr>
<td>49344.8584</td>
<td>0.9596</td>
<td>−6.27</td>
<td>4.17</td>
<td>−3.48</td>
<td>50.94</td>
<td>2.69</td>
<td>−2.76</td>
</tr>
<tr>
<td>49373.6908</td>
<td>0.7842</td>
<td>14.82</td>
<td>4.56</td>
<td>−1.44</td>
<td>37.61</td>
<td>2.94</td>
<td>+4.25</td>
</tr>
<tr>
<td>49410.6311</td>
<td>0.1220</td>
<td>34.84</td>
<td>4.56</td>
<td>−2.02</td>
<td>7.51</td>
<td>2.94</td>
<td>−3.83</td>
</tr>
<tr>
<td>49644.8270</td>
<td>0.9430</td>
<td>3.16</td>
<td>3.08</td>
<td>+4.65</td>
<td>52.75</td>
<td>1.98</td>
<td>+0.43</td>
</tr>
<tr>
<td>49699.8045</td>
<td>0.4223</td>
<td>32.37</td>
<td>3.08</td>
<td>−0.02</td>
<td>17.20</td>
<td>1.98</td>
<td>+1.08</td>
</tr>
<tr>
<td>53693.8806</td>
<td>0.1867</td>
<td>39.09</td>
<td>4.82</td>
<td>+1.44</td>
<td>8.27</td>
<td>3.10</td>
<td>−2.23</td>
</tr>
<tr>
<td>56578.9704</td>
<td>0.7691</td>
<td>17.46</td>
<td>3.95</td>
<td>+0.11</td>
<td>34.61</td>
<td>3.41</td>
<td>+2.43</td>
</tr>
<tr>
<td>56606.9106</td>
<td>0.5373</td>
<td>22.67</td>
<td>3.95</td>
<td>−5.97</td>
<td>26.25</td>
<td>3.41</td>
<td>+6.12</td>
</tr>
<tr>
<td>56650.8257</td>
<td>0.3164</td>
<td>31.34</td>
<td>3.95</td>
<td>−3.96</td>
<td>11.49</td>
<td>3.41</td>
<td>−1.52</td>
</tr>
</tbody>
</table>

Computed from the ephemeris given in Table 3.3.
The nominal temperature corresponds formally to a spectral type of K7, although $T_{\text{eff}}$ is degenerate with the surface gravity in our procedure because of the short wavelength coverage of the Reticon spectra from which we made these determinations. However, the RVs are unaffected by this degeneracy so long as the temperature for the templates is chosen to provide the optimal match to the spectra for a given log $g$ value. Rotational broadening has a much larger effect on the velocities in this case because of the heavy line blending, and we believe our fine-tuning of this parameter for both stars is the reason we are able to derive meaningful velocities from all 23 of the spectra used by Mathieu et al. (1997) (in addition to the other 10 from CfA used here). Their procedures only allowed them to derive separate velocities for 14 of their least blended spectra, the rest providing only an upper limit of 18 km s$^{-1}$ on the velocity separation between the primary and secondary. We list our new radial-velocity measurements from all CfA spectra in Table 3.1.

The velocity zero-point of the Reticon observations was monitored each night by means of dusk and dawn exposures of the twilight sky, and small run-to-run corrections were applied in the manner described by Latham (1992). For TRES the zero point was monitored by observing IAU velocity standards each night. All velocities were placed on the same system, and the measurements listed in Table 3.1 include all corrections.
3.3 Analysis and Results

3.3.1 CO Disk Modeling

We analyze the CO $J=2-1$ line emission using the framework detailed in Czekala et al. (2015) and Rosenfeld et al. (2012). Briefly, we forward-model the ALMA visibilities using a parametric description of the disk structure (densities, temperatures, and velocities). For any set of model parameters, we calculate the excitation conditions in the disk assuming local thermodynamic equilibrium. We then use the RADMC-3D radiative transfer code (v0.38; Dullemond 2012) to ray-trace spectral images, which are Fourier transformed and sampled at the same spatial frequencies as the data. A $\chi^2$ likelihood function is used to assess the fit quality. The posterior parameter-space is explored with a Markov Chain Monte Carlo (MCMC) algorithm.¹

The model parameters can be catalogued into four groups. The first group includes parameters that describe the CO gas densities. We assume the standard Lynden-Bell & Pringle (1974) similarity solution describes the radial surface density profile of the gas, which is described by an index $p$,² a characteristic radius $r_c$, and a normalization that we cast in terms of the CO gas mass $M_{\text{CO}}$. For computational expediency, we fix $p = 1$. The second group describes the gas temperatures. We simplify the scenario by assuming a vertically isothermal structure, with a radial power-law temperature

¹The code used to perform the analysis described here is open source and freely available under an MIT license at https://github.com/iancze/JudithExcalibur

²This is more commonly $\gamma$, but we aim to avoid confusion with the standard terminology in the RV analysis (see Sect. 3.3.2).
profile that has normalization $T_{10}$ (the temperature at 10 AU) and index $q$. This thermal structure is employed in calculating the vertical density distribution, assuming hydrostatic equilibrium. The third group of parameters sets the projected velocity field of the gas, presumed to be in Keplerian rotation. It includes the central binary mass $M_*$, the disk inclination $i_d$ and position angle $\varphi$, and a systemic velocity $v_{\text{sys}}$. Non-thermal line broadening is permitted with a line-width $\xi$ added in quadrature to the normal thermal contribution. In our convention, $i_d = 0^\circ$ is a face-on disk with the angular momentum vector pointed towards the observer, $i_d = 90^\circ$ is edge-on, and $i_d = 180^\circ$ is face-on but with the disk angular momentum vector pointed away from the observer. The position angle $\varphi$ is defined by the projection of the angular momentum axis onto the sky. The fourth group of parameters is utilitarian, including the distance $d$ and nuisance offsets from the observed phase center ($\Delta_\alpha$, $\Delta_\delta$).

We explore the 12-dimensional posterior-space with an ensemble MCMC sampler (Foreman-Mackey et al. 2013), employing uniform (uninformative) priors on all parameters except for $i_d$ and $d$. We adopt a standard geometrical prior on the disk inclination, $p(i_d) = \sin(i_d)/2$, reflecting that there are more disk orientations that result in edge-on than face-on viewing angles.\footnote{As we show below, the disk plane is near the plane of the sky. Given that fact and the substantial cloud contamination near the systemic velocity, we cannot uniquely determine the direction of the angular momentum axis (i.e., the sense of $i_d$, whether it is $\sim 160^\circ$ or $\sim 20^\circ$) from the ALMA data alone. Therefore, we employ the astrometric constraint made by Boden et al. (2009) from infrared interferometry measurements to enforce $i > 90^\circ$ in our analysis.} We choose a conservative Gaussian prior on the distance with a mean of 145 pc and a width ($\sigma$) of 20 pc, meant to represent the range of possible distances to sources in the Taurus clouds (e.g., Torres et al. 2010).
this analysis, we also conservatively exclude from the likelihood calculations 23 spectral channels that show evidence of molecular cloud contamination; these are marked in Fig. 3.1. The resulting inferences on the model parameters are listed in Table 3.2. A comparison of the data with the model is shown in the form of channel maps in Figure 3.1. As was demonstrated clearly in previous work (e.g., Simon et al. 2000; Rosenfeld et al. 2012), the density- and temperature-related parameters have negligible impact on an inference of the host mass. The key parameters are $M_*$ and $i_d$: the \{$M_*$, $i_d$\} joint posterior distribution is shown in Figure 3.2.

We infer a mass of $1.27^{\pm 0.46}_{-0.27} M_\odot$ for the DQ Tau binary, marginalized over the uncertainty contained in our distance prior. This can be expressed in a distance-independent manner as $M_*/d = 0.0086^{+0.0021}_{-0.0018} M_\odot \text{ pc}^{-1}$; the formal uncertainty on $M_*$ is $\sim 25\%$ if the distance is known exactly. This precision is significantly poorer than for most disk-based dynamical mass measurements, due to the unfortunate combination of a relatively face-on viewing geometry ($i_d \approx 160^\circ$) and the cloud contamination around the systemic velocity.

3.3.2 An Updated Spectroscopic Orbital Solution

The orbital solution by Mathieu et al. (1997) used their 14 pairs of CfA RVs along with the velocities measured from the Lick, Keck, and McDonald observatories (hereafter the ‘LKM’ set). Because of difficulties in maintaining a consistent velocity zero point from night to night and instrument to instrument, the latter data were originally derived only as velocity differences between the primary and secondary, rather than individual velocities for each star. To incorporate these LKM data into a conventional
Figure 3.2.—: (top): The joint posterior distribution for \( \{M_*, i_d\} \), marginalized over all other parameters. To compare with the constraints from the updated binary orbit from RV monitoring measurements, we overlay (±1 \( \sigma \)) contours for the measurement of \( M_* \sin^3 i \) (see Sect. 3.3.2). (bottom): The joint posterior distribution combining the RV and disk measurements and assuming \( i = i_d \). Contours denote 1, 2, and 3 \( \sigma \) levels.
Table 3.2:: Inferred Parameters for DQ Tau

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{10}$ (K)</td>
<td>121 ± 10</td>
<td>$i_d$ (°)</td>
<td>160 ± 3</td>
</tr>
<tr>
<td>$q$</td>
<td>0.71 ± 0.02</td>
<td>$\varphi$ (°)</td>
<td>94.2 ± 0.5</td>
</tr>
<tr>
<td>log $M_{CO}$ ($M_\odot$)</td>
<td>−8.0 ± 0.3</td>
<td>$v_{sys}$ (km s$^{-1}$)</td>
<td>+9.24 ± 0.01</td>
</tr>
<tr>
<td>$r_c$ (AU)</td>
<td>28 ± 4</td>
<td>$\Delta \alpha$ (&quot;)</td>
<td>−0.088 ± 0.003</td>
</tr>
<tr>
<td>$M_*$ ($M_\odot$)</td>
<td>1.27$^{+0.46}_{-0.27}$</td>
<td>$\Delta \delta$ (&quot;)</td>
<td>−0.246 ± 0.003</td>
</tr>
<tr>
<td>$\xi$ (km s$^{-1}$)</td>
<td>0.18 ± 0.02</td>
<td>$d$ (pc)</td>
<td>155 ± 15</td>
</tr>
</tbody>
</table>

The quoted uncertainties represent the maximum likelihood estimate and the 68.3% highest density interval computed around this value. The systemic velocity is given in the LSRK frame for the standard radio definition, and corresponds to $+21.95 \pm 0.01$ km s$^{-1}$ in the barycentric frame. Samples from the posterior are published at https://figshare.com/articles/MCMC_Samples/2063424.
double-lined orbital solution, and at the same time to tie those observations to the CfA zero point, Mathieu et al. (1997) constructed primary and secondary “measurements” from each velocity difference. They did this assuming a fixed mass ratio (of unity) and center-of-mass velocity, based on the values inferred from an initial fit based only on their 14 CfA velocity pairs (where they found $M_2/M_1 = 0.97 \pm 0.15$ and $\gamma = 22$ km s$^{-1}$). They then combined all of the measurements into a final fit, but were careful to note that both $\gamma$ and the velocity semiamplitudes $K_1$ and $K_2$ are biased and should be ignored in favor of the values from the CfA-only solution, and similarly with the minimum masses and projected semimajor axes, which depend on the semiamplitudes.

For this work we have preferred to incorporate the LKM velocity differences directly into the fit in their original form. We therefore reconstructed the original velocity differences trivially from the primary and secondary “measurements” reported by Mathieu et al. (1997) in their Table 1. The 22 RV differences were combined in a weighted least-squares fit with our 33 pairs of primary/secondary velocities, yielding the elements listed in Table 3.3. For the individual velocities, weights were calculated from the internal errors. The LKM velocity differences were assigned reasonable nominal errors to begin with, and all uncertainties were then adjusted iteratively so as to obtain reduced $\chi^2$ values near unity for each type of measurement (primary, secondary, RV difference). Final root-mean-square residuals, which are representative of the typical measurement errors, are given in the table. The global fit derives most of the constraint on the mass ratio from the individual primary and secondary velocities. The RV differences strongly constrain the $K_1 + K_2$ sum, but they also help to strengthen the individual $K$ values indirectly to some extent through constraints on the remaining orbital elements. We initially allowed for a difference in the center-of-mass velocities for
the primary and secondary, to account for possible biases in the RVs that may occur as a result of template mismatch, but found the difference to be insignificantly different from zero \((-0.61 \pm 0.71 \text{ km s}^{-1}\)). Consequently, the final fit assumed a common value of \(\gamma\).

A graphical representation of the 33 pairs of primary/secondary velocities from CfA is presented in Figure 3.3, along with our best-fit model from the global fit that includes the LKM velocity differences. In Figure 3.4 we illustrate the good agreement between the same best-fit model (solid curve) and the LKM velocity differences. The deviations between this best-fit model and a separate one that uses only the 33 individual CfA velocities (dotted line in the figure) are minimal.

Our results in Table 3.3 are generally consistent with those of Mathieu et al. (1997), but with uncertainties typically reduced by factors of 2–5. The minimum masses now have relative uncertainties of 5–6% instead of \(\sim21\%\).

3.4 Discussion

We have reported a new dynamical constraint on the mass of the young DQ Tau binary made by reconstructing the velocity field of its circumbinary disk using ALMA observations of its CO line emission, as well as an update on the binary orbital parameters based on a long-term optical spectroscopic monitoring campaign. In the following sections, we compare these constraints in more detail and discuss them in the context of pre-MS evolutionary model predictions and similar measurements for other equal-mass binary systems.
Figure 3.3.— CfA radial-velocity measurements of DQ Tau as a function of orbital phase, including our best fit model that uses also the LKM velocity differences. Primary velocities are represented with filled symbols, and the dotted line marks the center-of-mass velocity. The bottom panels show the residuals.
Figure 3.4.—: Predicted velocity differences as a function of orbital phase according to our best-fit model (solid line), shown with the measured LKM values (triangles). Residuals are displayed at the bottom. Also shown for reference in the top panel is an orbit model that uses only the 33 individual RV measurements from CfA (dotted line), which is nearly indistinguishable from the global fit. The dots represent the velocity differences we compute from the CfA measurements, to show that both types of measurements are fully consistent with each other.
Table 3.3: Updated spectroscopic orbital solution for DQ Tau.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P ) (days)</td>
<td>15.80158 ± 0.00066</td>
</tr>
<tr>
<td>( \gamma ) (km s(^{-1}))</td>
<td>+24.52 ± 0.33</td>
</tr>
<tr>
<td>( K_1 ) (km s(^{-1}))</td>
<td>20.28 ± 0.71</td>
</tr>
<tr>
<td>( K_2 ) (km s(^{-1}))</td>
<td>21.66 ± 0.60</td>
</tr>
<tr>
<td>( e )</td>
<td>0.568 ± 0.013</td>
</tr>
<tr>
<td>( \omega_1 ) (deg)</td>
<td>231.9 ± 1.8</td>
</tr>
<tr>
<td>( T_{\text{peri}} ) (HJD−2,400,000)</td>
<td>47433.507 ± 0.094</td>
</tr>
<tr>
<td>( M_1 \sin^3 i ) (( M_\odot ))</td>
<td>0.0348 ± 0.0017</td>
</tr>
<tr>
<td>( M_2 \sin^3 i ) (( M_\odot ))</td>
<td>0.0326 ± 0.0020</td>
</tr>
<tr>
<td>( (M_1 + M_2) \sin^3 i ) (( M_\odot ))</td>
<td>0.0674 ± 0.0033</td>
</tr>
<tr>
<td>( a_1 \sin i ) (10(^6) km)</td>
<td>3.63 ± 0.13</td>
</tr>
<tr>
<td>( a_2 \sin i ) (10(^6) km)</td>
<td>3.87 ± 0.11</td>
</tr>
<tr>
<td>( a \sin i ) (( R_\odot ))</td>
<td>10.78 ± 0.18</td>
</tr>
<tr>
<td>( q \equiv M_2/M_1 )</td>
<td>0.936 ± 0.051</td>
</tr>
<tr>
<td>( \sigma_1 ) (km s(^{-1}))</td>
<td>3.44</td>
</tr>
<tr>
<td>( \sigma_2 ) (km s(^{-1}))</td>
<td>2.26</td>
</tr>
<tr>
<td>( \sigma_{\text{LKM}} ) (km s(^{-1}))</td>
<td>2.50</td>
</tr>
<tr>
<td>Time span (days)</td>
<td>10668.8</td>
</tr>
<tr>
<td>Time span (orbital cycles)</td>
<td>675.2</td>
</tr>
<tr>
<td>( N_{\text{RV}} )</td>
<td>33 \times 2</td>
</tr>
<tr>
<td>( N_{\text{LKM}} )</td>
<td>22</td>
</tr>
</tbody>
</table>

These results are based on a joint fit of the individual primary/secondary velocities from CfA and the LKM velocity differences. The physical constants used here are those adopted by Torres et al. (2010), consistent with the 2015 IAU Resolution B3.
3.4.1 Comparison of Disk- and Binary-based Constraints

The disk-based dynamical mass approach formally constrains the quantity \((M_*/d) \sin^2 i_d\) by reconstructing the sky-projected Keplerian velocity field of the gas disk. Given some prior information on the distance and sufficient resolution to determine the aspect ratio of the emission \((i_d)\), a precise estimate of \(M_*\) can be made. This is not quite the case for DQ Tau. The disk orbital plane is oriented such that it is viewed nearly in the plane of the sky, which concentrates most of the more spatially extended molecular line emission near the systemic velocity. That would be fine given our ALMA observations, except for the ambient molecular cloud material that also produces extended emission at those same velocities. Taken together, the small projection factor and severe cloud contamination significantly expand the \(\{M_*, \sin^2 i_d\}\) degeneracy (see Fig. 3.2), limiting the precision of our dynamical mass constraint. For a conservative prior on \(d\), we measure a joint constraint of \(M_* \sin^2 i_d = 0.164 \pm 0.016 M_\odot\), or individual measurements of \(1.27^{+0.46}_{-0.27} M_\odot\) and \(i_d = 160 \pm 3^\circ\).

The orbital solution for a double-lined spectroscopic binary determines \(M_* \sin^3 i\) (independent of \(d\)) from a fit to a time series of RV measurements. The updated solution presented here has \(M_* \sin^3 i = 0.0674 \pm 0.0033 M_\odot\). Figure 3.2 confirms that the ALMA disk-based and RV binary-based constraints are in good agreement (well within 1 \(\sigma\)) in the binary mass–inclination plane. This suggests that these constraints can be combined together to yield some informative combined measurements for the system. If we assume that the binary and disk orbital planes are exactly aligned \((i = i_d)\), the joint constraints from the RV and ALMA data indicate \(M_* = 1.21 \pm 0.26 M_\odot\) and \(i = 158 \pm 2^\circ\) (this composite posterior is shown in the bottom panel of Fig. 3.2). If we
consider the ALMA constraint on the quantity $M_*/d$ rather than marginalizing over the prior on $d$, we can use the assumption of coplanarity and the RV data to estimate a dynamical distance to the system. In that case, we estimate $d_{\text{dyn}} = 184 \pm 26$ pc, which, although imprecise, has a most probable value slightly higher (at the $\sim 1\sigma$ level) than standard measurements for the Taurus star-forming region (e.g., Torres et al. 2010) and our adopted prior, which may suggest a larger depth of the Taurus complex. In the context of our nominal prior on $d$, we can also use both datasets to instead infer a limit on the mutual inclination angle between the disk and binary orbital planes: we find that $\psi \equiv i - i_d = -1.3 \pm 1.1^\circ$. Interestingly, we note a small discrepancy between the systemic velocity in the barycentric frame derived from the disk ($+21.95 \pm 0.01$ km s$^{-1}$) and that derived from the binary orbit ($+24.52 \pm 0.33$ km s$^{-1}$). We speculate that this offset may be caused by veiling of the stellar photospheres, which results in a sub-optimal fit of the spectroscopic templates used for the radial velocity determinations.

3.4.2 Comparison to Pre-MS Evolution Models

Having demonstrated that independent dynamical constraints on $M_*$ for the DQ Tau binary yield consistent results, it is of interest to make a comparison with the more common approach of estimating masses (and ages) from theoretical pre-MS evolutionary models.

A range of (combined-light) spectral types have been reported for DQ Tau, with a general consensus around M0–M1. Individual spectral diagnostics often skew towards earlier or later spectral types: for example, Basri et al. (1997) found that ratios of temperature sensitive Sc I lines suggest a K4–K5 classification, while Bary & Petersen
CHAPTER 3. DQ TAU

(2014) showed that many infrared molecular features (e.g., TiO, FeH, and H$_2$O) are better matched with a later type, M2.5–M4.5. Some of this ambiguity may be due to the implicit assumption that both stars have identical photospheric properties. The improved orbital solution in Sect. 3.3.2 suggests otherwise: the inferred mass ratio ($M_2/M_1 = 0.93 \pm 0.05$) indicates that the DQ Tau stars have different temperatures and luminosities.

With that in mind, we explored a two-component fit to the $BVRIJ$ photometry compiled by Rydgren (1984), Kolotilov (1989), and Skrutskie et al. (2006) (previously presented by Andrews et al. 2013). Observations in the $U$-band and at longer infrared wavelengths were excluded due to contamination by accretion activity and dust emission, respectively. The adopted model magnitudes were interpolated for a given $\{T_{\text{eff}}, \log g\}$ from the BT-SETTL synthetic photometry catalog (Allard et al. 2003) for solar metallicity. These were adjusted for extinction using the Fitzpatrick (1999) reddening law (with $R_V = 3.1$) and scaled to account for a given luminosity (assuming the same prior on $d$ as in Sect. 3.3.1). After some experimentation, we found that the effects of surface gravity are relatively small (given the other uncertainties), so we fixed $\log g = 4.0$ for both stars. Each model therefore has five physical parameters, $\{T_1, L_1, T_2, L_2, A_V\}$. We used an additional five nuisance parameters (one per band) to describe the “jitter” (dispersion) in each photometric band due to variability (presumed to be described by a Gaussian with mean zero and this parametric description of the variance). The model quality for a given set of parameters was determined with a $\chi^2$ likelihood function and a reasonable set of priors. At each posterior draw, we calculated the implied mass ratio and imposed a Gaussian prior with mean 0.93 and dispersion 0.05, based on the RV orbital solution. Since photometry alone is a poor diagnostic of $T_{\text{eff}}$ (especially for a
CHAPTER 3. DQ TAU

composite dataset), we adopted a Gaussian prior with mean of 3900 K and dispersion of 250 K on $T_1$ and $T_2$ and enforced the conditions $T_2 \leq T_1$ and $L_2 \leq L_1$.

Figure 3.5 provides a summary of the modeling results. We find that the photometry prefers relatively low temperatures, $T_1 = 3700 \pm 200$ K and $T_2 = 3500 \pm 175$ K, and extinction, $A_V = 0.5 \pm 0.2$, yielding logarithmic luminosities $\log_{10} L/L_{\odot}$ of $-0.73 \pm 0.16$ and $-0.87 \pm 0.16$ for the primary and secondary, respectively. These are not particularly stringent constraints on the binary location in the HR diagram, of course, owing to the relatively ambiguous spectral classifications available for the individual components.

Based on the joint posterior distribution of $\{T_1, L_1, T_2, L_2\}$, we followed the formalism of Jørgensen & Lindegren (2005) to derive component masses and ages $\{\tau, M_1, M_2\}$ from the predictions of pre-MS evolutionary models in the HR diagram, assuming that the binary stars are coeval. Various incarnations of such models (Siess et al. 2000; Dotter et al. 2008; Tognelli et al. 2011; Baraffe et al. 2015) make consistent predictions within the (considerable) uncertainties, indicating a total binary mass of $1.20 \pm 0.16 M_{\odot}$ that is in good agreement with the dynamical constraints from the ALMA and RV data. The corresponding age predictions are considerably more uncertain; favored values are in the 6–10 Myr range, although the permissible ages span from $\sim 6$ to 20 Myr ($1\sigma$).

We note that this analysis is under the assumption of coevality of the two stars, which may not necessarily be true. Additionally, the unusual nature of the DQ Tau system (e.g., colliding magnetospheres during periastron) may also invalidate our assumptions of normal pre-main sequence evolution.
Figure 3.5.— (top): The best-fit models of the broadband photometry overlaid on the data. (middle) The resulting HR diagram, with the marginal posteriors inferred from the photometry modeling shown as 1σ contours. The Dotter et al. (2008) pre-MS model mass tracks and isochrones are overlaid. (bottom): The joint mass and age constraints from the Dotter et al. (2008) pre-MS models assuming the stars are coeval, shown as 1σ contours. The marginalized distributions are shown at the boundaries of the plot. The gray band marks the disk-based constraint on the individual component mass $M_*$ (1σ). Other pre-MS model predictions give generally comparable results.
CHAPTER 3. DQ TAU

3.5 Summary and Context

We have presented a set of new constraints on the fundamental properties of the DQ Tau young binary system, based on ALMA observations of molecular line emission from its circumbinary disk and an updated analysis of optical spectroscopic measurements of its (stellar) radial velocity variations. For a conservative distance prior \( d = 145 \pm 20 \text{ pc} \), we find that the disk-based and binary-based dynamical constraints on the total stellar mass in the DQ Tau system are in excellent agreement: their combined inputs suggest a total mass \( M_\ast = 1.21 \pm 0.26 \, M_\odot \), and therefore individual component masses \( M_1 = 0.63 \pm 0.13 \, M_\odot \) and \( M_2 = 0.59 \pm 0.13 \, M_\odot \) (incorporating the uncertainty on \( q \)). Moreover, we also find that the disk and binary orbital planes are aligned within \( 3^\circ \), showing that the system is coplanar across radial distances from \( \sim 0.1 \text{ AU} \) to \( 100 \text{ AU} \).

In this system, the dynamical mass precision is limited by an unfortunate combination of two factors: an orbital plane that is oriented nearly in the sky plane, and some large-scale contamination of the disk CO spectral emission from the ambient molecular cloud. In the future, an accurate parallax from GAIA will help improve the precision of the disk-based estimate of \( M_\ast \).

We also estimated the stellar mass in the system using the more common technique that compares the component locations in the HR diagram with predictions of theoretical pre-MS evolution models, and generally found good agreement. However, that approach has restricted utility given the lack of component-resolved photometry and substantial ambiguity on the effective temperatures. There is still much to be learned from this fascinating system; our mass constraints lend some quantitative benchmarks that can be adopted in future studies.
DQ Tau is the third nearly-equal mass young binary system that has been analyzed with these two independent dynamical techniques to constrain stellar masses, the others being the older and earlier type systems V4046 Sgr (total mass 1.75 $M_\odot$; Rosenfeld et al. 2012) and AK Sco (total mass 2.50 $M_\odot$; Czekala et al. 2015). Using millimeter-wave interferometric measurements of their CO spectral line emission, model fitting that reconstructs the Keplerian velocity fields of their circumbinary disks finds dynamical masses that are in excellent agreement with constraints from optical RV monitoring of the host binaries (thereby also implying that the binary and disk orbits are co-planar). Granted, this is a small sample, but it does span an important range of system properties: e.g., spectral types from early M to mid F, ages from a few to tens of Myr, and orbital eccentricities from circular to $e \approx 0.6$. Taken together, this work validates the quantitative accuracy of the disk-based dynamical inference of young star masses, provided that it is done carefully in a proper analysis framework. Moving forward, this confirms that ALMA should play a substantial role in young star astrophysics, as the technique used here is the only means of precisely measuring the masses for large samples of single stars.
Chapter 4

Constructing a Flexible Likelihood Function for Spectroscopic Inference


4.1 Introduction

All astronomers recognize that spectroscopy offers a wealth of information that can help characterize the properties of the observing target. In the context of stellar astrophysics, spectroscopy plays many fundamental roles. The relative strengths and widths of stellar absorption lines provide access to physical properties like effective temperature ($T_{\text{eff}}$) and surface gravity ($\log g$), enabling model comparisons in the Hertzsprung-Russell diagram to estimate the masses and ages so crucial to understanding stellar evolution, as well as individual elemental abundances or the collective “metallicity” (typically...
CHAPTER 4. SPECTROSCOPIC INFERENCE

parameterized as $[\text{Fe/H}]$, facilitating studies of the chemical hallmarks of different stellar populations. With sufficient resolution, a spectrum also conveys information about rotation ($v \sin i$) and kinematics (e.g., association with a cluster or companion through the radial velocity, $v_r$). While many fields benefit from such spectroscopic measurements, they are of acute interest to the exoplanet community. There, all estimates of the planet properties are made relative to the host properties (e.g., the mass function and planet-to-host radius ratio are constrained with the radial velocity or transit techniques, respectively). Moreover, essential clues to the planet formation process are encapsulated in the dependences of planet frequency on host mass (e.g., Johnson et al. 2007; Howard et al. 2010) and metallicity (e.g., Fischer & Valenti 2005; Buchhave et al. 2014).

The robust and quantitative extraction of physical (or empirical) parameters from an observed spectrum can be an extraordinary challenge. Stellar models serve as comparative benchmarks to associate observed spectral features with the parameters of interest. Generating a synthetic model spectrum involves a complex numerical treatment of the stellar structure and radiative transfer through the atmosphere (e.g., Kurucz 1993; Castelli & Kurucz 2004; Hauschildt et al. 1999; Husser et al. 2013; Paxton et al. 2011). Detailed models calibrated to individual stars are important, but rare (e.g., the Sun, Vega); therefore, these stellar models are relatively untested in large swaths of parameter-space. Moreover, they necessarily include simplifications to treat complicated physical processes (e.g., convection) or computational limitations (e.g., boundary conditions), and often must rely on incomplete or inaccurate atomic and molecular information (e.g., opacities). In principle, the models could be improved with appropriate reference to spectroscopic data. Nevertheless, they are remarkably successful in reproducing many diagnostic spectral features.
CHAPTER 4. SPECTROSCOPIC INFERENCE

There are various well-tested approaches being used in stellar astrophysics to compare these models with observed spectra and thereby infer basic parameters. Perhaps the most common is a straightforward empirical technique that relies on distilling an information-rich subset of the data, usually in the form of spectral line equivalent widths and/or local continuum shapes. A combined sequence of the ratios of these quantities can be especially sensitive to a given model parameter (e.g., MOOG; Sneden 1973; Gray 1994; Reid et al. 1995; Rojas-Ayala et al. 2010, 2012). This “indexing” approach has the advantage of being trivially fast. But, each condensed relationship is only informative over a limited swath of parameter-space, and it potentially masks degeneracies that are encoded in the spectral line shapes. Another standard approach exploits the cross-correlation of an observed spectrum with a suite of model templates to optimize a set of parameters, usually with some weighting applied to specific spectral regions (e.g., SPC; Buchhave et al. 2012). In this case, the speed advantage is maintained (perhaps enhanced) and more data content is used (particularly in the spectral dimension), thereby achieving higher precision even for data with comparatively low signal-to-noise. The disadvantage is that the model quality and parameter inferences are assessed in a heuristic (rather than probabilistic) sense, making it difficult to quantify uncertainty in the stellar parameters. A more direct method employs a pixel-by-pixel comparison between model and data. This has the benefits of increased parametric flexibility (e.g., one can fit for arbitrary abundances or structures) and a proper inference framework (usually a least-squares approach, although increasingly in a Bayesian format; Shkedy et al. 2007; Schönrich & Bergemann 2014). Ultimately, rather than pre-computing a library of synthetic spectra, one would like to incorporate the spectral synthesis back-end (e.g., SME; Valenti & Piskunov 1996) directly into the likelihood function, bypassing
any interpolation when assessing the fit of stellar parameters in-between grid points in the library. Unfortunately, this is not yet computationally feasible beyond a limited wavelength range.

In this article, we construct a flexible forward-modeling method for the general spectroscopic inference problem in a Bayesian framework, building on the best aspects of the latter two approaches highlighted above. The key developments in this design include a spectral emulator to address the difficult task of interpolation in coarsely sampled synthetic spectral libraries and a non-trivial covariance matrix parameterized by both global (stationary) and local (non-stationary) Gaussian process kernels. When combined with an appropriately sophisticated set of quantitative metrics for the relevant physical parameters, this method will efficiently propagate systematic uncertainties into the parameter inferences. Ultimately, this approach could be employed to leverage spectroscopic data as a reference for improving the models.

A complete overview of the methodology behind this approach is provided in Section 4.2. Some tests and example applications (for a high resolution optical spectrum of an F star, and a medium-resolution near-infrared spectrum of a mid-M star) are described in Section 4.3. Finally, a discussion of its potential utility, especially the possibility of extending it to develop data-driven spectral models, is provided in Section 4.4.
CHAPTER 4. SPECTROSCOPIC INFERENCE

4.2 Methodology

This section describes a generative Bayesian modeling framework that confronts some of the key technical obstacles in the spectroscopic inference problem. The goal is to conservatively extract the maximal amount of information about a prescribed (and usually degenerate) parameter set by forward-modeling an observed spectrum, while also recognizing and explicitly accounting for the covariances (and potentially biases) introduced by pathologically imperfect models. The method is modular, and therefore can easily incorporate additional physical or nuisance parameters as desired without sacrificing an accurate reflection of the limitations in the data. The specific applications discussed here are related to the spectra of individual stars, but the methodology is generic (and could be used for the composite spectra of unresolved stellar clusters, galaxies, etc.).

Figure 4.1 serves as a graphical guide to the mechanics of this modeling framework, and the remainder of this section. First, a model spectrum is generated for a given set of physical parameters (Section 4.2.1; Appendix 4.5), and then post-processed to mimic reality using a set of observational and practical nuisance parameters (Section 4.2.2). Next, a direct, pixel-by-pixel comparison between the data and model spectra is made with a prescribed likelihood function and a parametric treatment of the covariances between pixel residuals (Section 4.2.3). That process is iterated using Markov Chain Monte Carlo (MCMC) simulations in a multi-stage Gibbs sampler to numerically explore the posterior probability density of the model conditioned on the data, and thereby to determine constraints on the parameters of interest (Section 4.2.5). Along the way, these procedures are illustrated with observations of the high resolution optical spectrum from
a nearby F star. That specific application, along with some alternative demonstrations of the method, are discussed in more detail in Section 4.3.

4.2.1 Generating a Model Spectrum

There are many approaches for generating a model spectrum, $f_{\lambda}$, for a specific set of parameters, $\theta_\ast = \{T_{\text{eff}}, \log g, [\text{Fe/H}]\}$. In the most direct case of spectral synthesis, a model atmosphere structure is assembled and simulations of energy transport through it are conducted with a radiative transfer code (e.g., Kurucz 1993; Hauschildt et al. 1999). However, in general this approach is often computationally prohibitive for most iterative methods of probabilistic inference. One partial compromise is to interpolate over a library of atmosphere structures that were pre-computed for a discrete set of parameter values, $\{\theta_\ast\}^{\text{grid}}$, for some arbitrary $\theta_\ast$. Then, perform a radiative transfer calculation with that interpolated atmosphere to synthesize $f_{\lambda}$ (e.g., SME; Valenti & Piskunov 1996). A more common variant is to interpolate over a pre-synthesized library of model spectra, $f_{\lambda}(\{\theta_\ast\}^{\text{grid}})$ (e.g., Husser et al. 2013; Schönrich & Bergemann 2014). Although the former approach is preferable, the computational cost of repeated spectral synthesis is enough to make a detailed exploration of parameter space less appealing (although see Section 4.4). Although the framework we are advocating is applicable for any “back-end” that generates a model spectrum, it is illustrated here using the latter approach with the Husser et al. (2013) PHOENIX library.

In practice, this reliance on spectral interpolation within a model library requires a sophisticated treatment of associated uncertainties. The key problems are that the spectra themselves do not vary in a straightforward way as a function of $\theta_\ast$ (especially
CHAPTER 4. SPECTROSCOPIC INFERENCE

within spectral lines), and that the typical model library is only sparsely sampled in \( \theta_* \). Because of these issues, standard interpolation methods necessarily result in some information loss. The practical consequence is that the inferred posteriors on the model parameters are often sharply peaked near a grid point in the library, \( \{ \theta_* \}_{\text{grid}} \), potentially biasing the results and artificially shrinking the inferred parameter uncertainties (e.g., Cottaar et al. (2014)). To mitigate these effects, we develop a spectral “emulator” that smoothly interpolates in a sparse model library and records a covariance term to be used in the likelihood calculation that accounts for the associated uncertainties. The emulator is described in detail in Appendix 4.5. We first decompose the model library into a representative set of eigenspectra using a principal component analysis. At each gridpoint in the library, the corresponding spectrum can be reconstructed with a linear combination of these eigenspectra. The weights associated with each eigenspectrum contribution vary smoothly as a function of the parameters, and so are used to train a Gaussian process to interpolate the weights associated with any arbitrary \( \theta_* \). In this way, the emulator delivers a probability distribution that represents the range of possible interpolated spectra. By then marginalizing over this distribution, we can modify the likelihood function to propagate the associated interpolation uncertainty. In the remainder of this section, the details of generating the reconstructed (interpolated) spectrum are not especially relevant (see Appendix 4.5).

4.2.2 Post-Processing

Typically, the “raw” interpolated model spectrum \( f_{\lambda}(\theta_*) \) that was generated above is highly over-sampled, and does not account for several additional observational and
instrumental effects that become important in comparisons with real data. Therefore, a certain amount of post-processing is required before assessing the model quality. We treat that post-processing in two stages. The first stage deals with an additional set of “extrinsic” parameters, $\theta_{\text{ext}}$, that incorporate some dynamical considerations as well as observational effects related to geometry and the relative location of the target. The second stage employs a suite of nuisance parameters, $\phi$, designed to forward model some imperfections in the data calibration.

We can further divide $\theta_{\text{ext}}$ into those parameters that impact the model primarily in the spectral or flux dimensions. For the former, we consider three kernels that contribute to the line-of-sight velocity distribution function. The first, $F_{\text{inst}}^v$, treats the instrumental spectral broadening. For illustrative purposes, we assume $F_{\text{inst}}^v$ is a Gaussian with a mean of zero and a constant width $\sigma_v$ at all $\lambda$, although more sophisticated forms could be adopted. The second, $F_{\text{rot}}^v$, characterizes the broadening induced by stellar rotation, parameterized by $v \sin i$ as described by Gray (2008, his Eq. 18.14), the rotation velocity at the stellar equator projected on the line of sight (where $i$ is the inclination of the stellar rotation axis). And the third, $F_{\text{dop}}^v = \delta(v - v_r)$, incorporates the radial velocity through a Doppler shift. The model spectrum is modified by the parameters $[\sigma_v, v \sin i, v_r]$ through these kernels, using a convolution in velocity-space,

$$f_\lambda(\theta_*, \sigma_v, v \sin i, v_r) = f_\lambda(\theta_*) * F_{\text{inst}}^v * F_{\text{rot}}^v * F_{\text{dop}}^v,$$

and then re-sampled onto the discrete wavelengths corresponding to each data pixel,

$$f_\lambda(\theta_*, \sigma_v, v \sin i, v_r) \mapsto M(\theta_*, \sigma_v, v \sin i, v_r),$$

1In practice, these convolutions are performed as multiplications in Fourier-space to better preserve spectral information (cf., Tonry & Davis 1979); the mathematical formalism is presented for clarity.
where the \( \mapsto \) symbol denotes a re-sampling operator that maps the model spectrum onto the \( N_{\text{pix}} \)-element model vector \( M \) (\( N_{\text{pix}} \) is the number of pixels in the spectrum). Figure 4.2 shows a (condensed) graphical representation of these post-processing steps.

At this stage, the model is further modified in the flux dimension. A typical synthetic spectrum is computed as the flux that would be measured \textit{at the stellar surface}, and so needs to be diluted by the subtended solid angle, \( \Omega = (R_*/d)^2 \), where \( R_* \) is the stellar radius and \( d \) is the distance. An additional wavelength-dependent scaling factor is applied to account for interstellar extinction, assuming some previously-derived extinction law \( A_\lambda \) (e.g., Cardelli et al. 1989) that is parameterized by \( A_V \). The parameters \([\Omega, A_V]\) are then applied as

\[
M(\Theta) = M(\theta_*, \theta_{\text{ext}}) = M(\theta_*, \sigma_v, v \sin i, v_r) \times \Omega \times 10^{-0.4 A_\lambda},
\]

with simplified notation such that \( \Theta \equiv [\theta_*, \theta_{\text{ext}}] \), where \( \theta_{\text{ext}} = [\sigma_v, v \sin i, v_r, \Omega, A_V] \). Some spectral libraries provide spectra as with peak fluxes normalized to a constant value, in that case, \( \Omega \) will simply serve as an arbitrary scaling parameter.

The procedure so far is composed of straightforward operations demanded by practical astronomical and computing issues. If the data were \textit{perfectly} calibrated, we could proceed to a likelihood calculation that makes a direct comparison with \( M(\Theta) \). However, the calibration of the continuum shape for data with reasonably large spectral range is often not good enough to do this. A common example of this imperfect calibration can be readily seen when comparing the overlaps between spectral orders from echelle observations. Even if such imperfections (e.g., in the flat field or blaze corrections, or perhaps more likely in the flux calibration process) induce only minor,
low-level deviations in the continuum shape, they can add up to a significant contribution in the likelihood function and thereby potentially bias the results.

The traditional approach to dealing with this issue has been avoidance; a low-order polynomial or spline function is matched (separately) to the model and the data and then divided off to normalize the spectra. While this is straightforward to do for earlier type stars, it only masks the problem. This normalization procedure disposes of useful physical information content available in the continuum shape, and can be considerably uncertain in cases where the spectral line density is high (e.g., for cooler stellar photospheres). Moreover, it cannot propagate the uncertainty inherent in deriving the normalization functions into a proper inference framework.

Instead, we employ a more rigorous approach that forward-models the calibration imperfections with a set of nuisance parameters that modify the shape of the model spectrum. By later marginalizing over these nuisance parameters, we properly account for any uncertainties that these kinds of calibration imperfections induce on the stellar parameters of interest while also maintaining the useful information in the continuum shape. In practice, this is achieved by distorting segments of the model with polynomials, \( P \) (e.g., Eisenstein et al. 2006; Koleva et al. 2009). Figure 4.3 demonstrates how these nuisance parameters are applied to the model. For \( N_{\text{ord}} \) spectral orders, each denoted with index \( o \), the model spectrum can be decomposed as

\[
M(\Theta, \phi_P) = \{ M_o(\Theta) \times P_o \} = \{ M_o(\Theta) \times \sum_n c_o^{(n)} T_o^{(n)} \}, \tag{4.4}
\]

\(^2\)For instance, the imperfect calibration would still in principle be discernible through the slight differences of the normalization functions derived for the data and model.
where \( T^{(n)} \) is an \( n^{th} \) degree Chebyshev function. The \( nN_{\text{ord}} \) coefficients are considered a set of nuisance parameters, \( \phi_P = \{ [c^{(0)}_o, c^{(1)}_o, \ldots, c^{(n-1)}_o] \} \). Judicious priors can ensure that the real spectral features (e.g., molecular bands) are not treated as residual calibration artifacts. The lowest-degree (scaling) coefficient, \( c^{(0)}_o \), is degenerate with the solid angle, \( \Omega \). Therefore, we enforce an additional constraint that the mean of the polynomial is unity. For data with a single spectral order, this means simply setting \( c^{(0)}_o = 1 \). In the multiple order case, we assign \( c^{(0)}_o = 1 \) in an arbitrary order as an anchor, but permit the \( c^{(0)}_o \) in other orders to be different.

### 4.2.3 Model Evaluation

The fit of the model spectrum is assessed by comparing to the data with a pixel-by-pixel likelihood calculation. If we denote the data spectrum as \( D \), then a corresponding residual spectrum (an \( N_{\text{pix}} \)-element vector) can be defined for any input parameter set,

\[
R \equiv R(\Theta, \phi_P) \equiv D - M(\Theta, \phi_P).
\]  

To quantify the probability of the data conditioned on the model, we adopt a standard multi-dimensional Gaussian likelihood function

\[
p(D|M) = \frac{1}{(2\pi)^{N_{\text{pix}}} \det(C)^{1/2}} \exp\left(-\frac{1}{2}R^T C^{-1} R\right)
\]

that penalizes models which yield larger residuals and explicitly allows for covariances in the residual spectrum through the \( N_{\text{pix}} \times N_{\text{pix}} \) matrix \( C \). For practical reasons, the log-likelihood is used as the quality metric, where

\[
\ln p(D|M) = -\frac{1}{2} \left(R^T C^{-1} R + \ln \det C + N_{\text{pix}} \ln 2\pi \right).
\]
CHAPTER 4. SPECTROSCOPIC INFERENCE

The covariance matrix $C$ characterizes both the measurement uncertainty ($\sigma$; “noise”) in each pixel and the covariance between pixels. When using a spectral emulator to interpolate model spectra, $C$ will be the sum of the covariance matrix described here and the emulator matrix derived in Appendix 4.5 (Eq. 4.59). In the special case where each pixel is independent, the covariance matrix is diagonal, $C_{ij} = \delta_{ij} \sigma_i^2$, where $\sigma_i$ is the uncertainty in pixel $i$ and $\delta_{ij}$ is the Kronecker delta function, and Eq. 4.7 reduces to the familiar

$$\ln p(D|M) - \text{constant} = -\frac{1}{2} \sum_i N_{\text{pix}} \frac{R_i^2}{\sigma_i^2} \equiv -\chi^2_2,$$

the sum of the square of the residuals weighted by their inverse variances. However, that simplification rarely applies in practice. A more complex covariance matrix is required, so that additional off-diagonal terms can be used to explicitly characterize (1) pixel-to-pixel covariances imposed by the discrete over-sampling of the line-spread function, and (2) highly correlated residuals as manifestations of systematic imperfections in the model library. The following sections describe how these issues are addressed by constructing a more sophisticated $C$.

Global Covariance Structure

Astronomical spectrographs are designed to have the detector over-sample the instrumental line-spread function with at least a few pixels. Therefore, adjacent pixels never record independent samples of the true spectrum. In that case, a difference between an observed and modeled spectral feature creates a correlated residual that spans multiple pixels. This can be demonstrated clearly in the autocorrelation of $R$: a slight model mismatch will produce correlated residuals over a characteristic scale similar to the instrumental or rotation broadening kernel width (whichever is larger). Figure 4.4
shows an example of these correlated residuals in real data; a significant autocorrelation signal is seen on an $\sim 8$ pixel scale, corresponding to the $6.8 \text{ km s}^{-1}$ FWHM of $F_v^{\text{inst}}$.

It is important to distinguish here between “noise” and the fit residuals. Noise introduced to the spectrograph by astrophysical or instrumental effects is generally uncorrelated with wavelength. The arrival and propagation of each photon through the instrument and into the detector can be considered an independent event. In essence, the noise itself is not correlated, but the fit residuals likely are. However, from a mathematical perspective the correlated residuals can be treated in the same way as correlated noise, by constructing a non-trivial covariance matrix with off-diagonal terms. In practice, this is achieved by parameterizing $C$ with a kernel that describes the covariance between any pair of pixels, indexed $ij$, representing wavelengths $\lambda_i$ and $\lambda_j$.

For a well-designed spectrograph and sufficiently accurate model, this global (i.e., present throughout the spectrum) covariance should have a relatively low amplitude and small correlation length. To describe that structure, we use a stationary covariance kernel (or radial basis function) with an amplitude that depends only on the velocity separation between two pixels,

$$r_{ij} = r(\lambda_i, \lambda_j) = \frac{c}{2} \left| \frac{\lambda_i - \lambda_j}{\lambda_i + \lambda_j} \right|, \quad (4.9)$$

where $c$ is the speed of light. This kernel is used to characterize the covariance between pixel residuals,

$$K_{ij}^G = \langle R_i, R_j \rangle. \quad (4.10)$$

A variety of useful kernels have been developed in the field of Gaussian processes to parameterize such a covariant structure (e.g., Rasmussen & Williams 2005; Santner et al. 2013), and are seeing increased use in many areas of astrophysics (for some specific
examples in stellar and planetary applications, see Foreman-Mackey et al. 2014; Aigrain et al. 2015; Barclay et al. 2015). After some experimentation, we adopted the Matérn kernel with $\nu = 3/2$ because it performed well at reproducing the appearance of realistic residuals for this specific problem. In this case,

$$K_{ij}^G(\phi_{C,G}) = w_{ij}a_G \left( 1 + \frac{\sqrt{3} r_{ij}}{\ell} \right) \exp \left( -\frac{\sqrt{3} r_{ij}}{\ell} \right),$$  \hspace{1cm} (4.11)

with $\phi_{C,G} = [a_G, \ell]$, an amplitude ($a_G$) and a scale ($\ell$). The $\phi_{C,G}$ are termed hyperparameters here; because a Gaussian process describes a population of functions generated by random draws from a probability distribution set by a mean vector and a covariance matrix, the kernel parameters are naturally part of a hierarchical model. In this specific case, the functions described by these hyperparameters represent many realizations of covariant residuals from a spectral fit. Figure 4.5 shows an example of the Gaussian process kernel and the covariant residuals that can be generated from it. To ensure that $C$ remains a relatively sparse matrix (for computational expediency), we employ a Hann window function

$$w_{ij}(r_0) = \begin{cases} \frac{1}{2} + \frac{1}{2} \cos \left( \frac{\pi r_{ij}}{r_0} \right) & r_{ij} \leq r_0 \\ 0 & r_{ij} > r_0 \end{cases}$$  \hspace{1cm} (4.12)

to taper the kernel. The truncation distance $r_0$ can be set to a multiple of the scale (we set $r_0 = 4\ell$).
Figure 4.1.—: A flowchart showing how the parameters of the model are combined to forward model a spectrum. Before starting inference for a particular star, a Bayesian emulator is tuned to efficiently interpolate a grid of synthetic spectra (Appendix 4.5) for any queried set of “intrinsic” stellar parameters ($\Theta^*$). The spectrum is then modified according to “extrinsic” stellar parameters ($\Theta_{\text{ext}}$) like $v \sin i$ and $v_r$. Then, calibration polynomials ($\phi_P$) provide slight adjustment to the continuum shape of the model to account for uncertainties in flux calibration. The second major component of the framework is accounting for covariant residual structure by using kernels to set the structure of the “noise” matrix to downweight erroneous residual structure. Then, the multidimensional likelihood function is evaluated using the sum of these covariance matrices.
Figure 4.2.—: (top) The line-of-sight velocity distribution function (solid black curve) and its decomposition into broadening kernels. The instrumental kernel (dotted) is treated as a Gaussian, the rotation kernel (dashed) is a parabola-like function of the projected rotational velocity, and the Doppler kernel (solid) is a δ-function that introduces the radial velocity. In this specific case, $\sigma_v = 2.9\, \text{km s}^{-1}$, $v \sin i = 5\, \text{km s}^{-1}$, and $v_r = 7\, \text{km s}^{-1}$, appropriate for the example in Section 4.3.1. (bottom) A segment of a raw, full-resolution model spectrum and its post-processed equivalent after convolution and re-sampling at the coarser resolution of the detector pixels.
Figure 4.3.—: A demonstration of our treatment for residual calibration mismatches. The observed spectra at the overlap of two echelle orders (top) have slightly (∼1–3%) discrepant continuum levels. By using Chebyshev polynomials (middle, top) one can correct for that mismatch by adjusting the data (middle, bottom); instead, in practice we equivalently distort the model by these polynomials (bottom) such that the model remains linear in the Chebyshev coefficients (Eq. 4.4). Note that this procedure preserves the natural units of flux and any intrinsic shape of the spectral energy distribution—the spectrum is not continuum normalized.
Figure 4.4.—: (*top*) A comparison of the data and a typical model with parameters drawn from the posterior distribution, along with the corresponding residual spectrum. (*middle*) A zoomed view of the gray band in the top panels, highlighting the mildly covariant residual structure that is produced by slight mismatches between the data and model spectra. (*bottom*) The autocorrelation of the residual spectrum. Notice the substantial autocorrelation signal for offsets of $\lesssim 8$ pixels, demonstrating clearly that the residuals are not well described by white (Poisson) noise alone.
Figure 4.5.—: A decomposition of the modeling procedure, explicitly highlighting the roles of the various contributions to the covariance matrix. The top panels show a typical comparison between the data and model spectrum, along with the associated residual spectrum. The subsequent panels focus on the illustrative region shaded in grey. The left column of panels show the corresponding region of the covariance matrix $C$, decomposed into its primary contributions: from top to bottom, the trivial noise matrix, then combined with the global covariance kernel, and finally including an appropriate local covariance kernel. In the right column, we show the zoomed-in residual spectrum (black) along with example random draws from the subsets of $C$ exhibited to the left. The shaded contours (orange) represent the 1, 2, and 3$\sigma$ dispersions of an ensemble of 200 random draws from $C$. Note that the trivial noise matrix ($\delta_{ij}\sigma_i$) poorly reproduces both the scale and structure of the residual spectrum. The addition of a global kernel ($K^G$) more closely approximates the structure and amplitude of the residuals, but misses the outlier line at 5202.2 Å. Including a local kernel ($K^L$) at that location results in a covariance structure that does an excellent job reproducing all the key residual features.
CHAPTER 4. SPECTROSCOPIC INFERENCE

Figure 4.5.—
Local Covariance Structure

In addition to the global covariance structure, there can be local regions of highly correlated residuals. These patches of large $R$ are usually produced by pathologically incorrect spectral features in the model, due to systematic imperfections like missing opacity sources or poorly constrained atomic/molecular data (e.g., oscillator strengths). Some representative examples are shown in Figure 4.6. To parameterize such regions in $C$, we introduce a sequence of non-stationary kernels that explicitly depend on the actual wavelength values of a pair of pixels ($\lambda_i$ and $\lambda_j$), and not simply their separation ($r_{ij}$).

Assuming that these local residual features are primarily due to discrepancies in the spectral line depth (rather than the line shape or central wavelength), a simple Gaussian is a reasonable residual model. In that case, the pixel residuals of the $k$-th such local feature could be described as

$$R_j \equiv R(\lambda_j) = A_k \exp \left[ -\frac{r^2(\lambda_j, \mu_k)}{2\sigma_k^2} \right]$$  \quad (4.13)

with peak amplitude $A_k$, central wavelength $\mu_k$, and width $\sigma_k$. We assume that the amplitude of this Gaussian feature is drawn from a normal distribution

$$A_k \sim \mathcal{N}(0, \sigma_k^2)$$  \quad (4.14)

with mean 0 and variance $\sigma_k^2$. The pixels in this Gaussian-shaped residual are correlated because each pixel shares a common random scale factor ($A_k$). Then, the covariance of any two pixels in this region is given by Eq. 4.10, where the expectation value is taken.

81
CHAPTER 4. SPECTROSCOPIC INFERENCE

with respect to the probability distribution in Eq. 4.14

\[ K_{ij}^{L,k} = \langle A_k \exp \left[ -\frac{r^2(\lambda_i, \mu_k)}{2\sigma_k^2} \right] A_k \exp \left[ -\frac{r^2(\lambda_j, \mu_k)}{2\sigma_k^2} \right] \rangle \]

\[ = \langle A_k^2 \rangle \exp \left[ -\frac{r^2(\lambda_i, \mu_k) + r^2(\lambda_j, \mu_k)}{2\sigma_k^2} \right] \]

\[ = a_k^2 \exp \left[ -\frac{r^2(\lambda_i, \mu_k) + r^2(\lambda_j, \mu_k)}{2\sigma_k^2} \right]. \tag{4.15} \]

The full local covariance kernel covering all of the possible Gaussian residuals is composed of a linear combination of kernels,

\[ K_{ij}^{L}(\phi_{C,L}) = \sum_k^N w_k^{ij} K_{ij}^{L,k}, \tag{4.16} \]

with a corresponding set of hyperparameters \( \phi_{C,L} = \{a_1, \mu_1, \sigma_1, \ldots, a_N, \mu_N, \sigma_N\} \). Note that we again taper the kernels with Hann windows (Eq. 4.12) to ensure a sparse covariance matrix; in this case, the truncation distance \( r_0 \) can be set to some multiple of the width parameter (e.g., \( r_0 = 4\sigma_k \)). In effect, these kernels systematically down-weight the influence of strong residuals in the likelihood calculation, mitigating any potential bias they might induce on inferences of the interesting parameters (\( \Theta \)). Similar in spirit to robust linear regression and “bad data” mixture models (Hogg et al. 2010), these kernels provide a means for (correlated) outlier rejection that preserves the integrity of the probabilistic framework (as opposed to the common manual or threshold-based techniques of masking or clipping).

In principle, the concept of these local kernels can be extended to account for more complex residual structures. For example, late-type stars with imperfectly modeled molecular bandheads may produce a complicated pattern of positive and negative residuals or a pronounced mismatch over a relatively large spectral scale. This phenomenologically different local covariance behavior can still be treated in this framework if an appropriate kernel morphology is adopted.
CHAPTER 4. SPECTROSCOPIC INFERENCE

Figure 4.6.—: A particularly illustrative spectral region with substantial localized structure in the residuals due to “outlier” spectral lines in the model library. For any specific line, there might exist a set of model parameters, $\Theta$, that will improve its match with the data, but a $\Theta$ that will properly fit all the outlier lines does not exist in a pre-computed library with (necessarily) limited parametric flexibility. Out of concern that such intrinsic mismatches can bias the inference on $\Theta$, the methodology advocated here introduces local kernels to inflate the covariance around these outliers, self-consistently down-weighting their influence on the fit.

**Composite Covariance Matrix**

We can now compute the covariance matrix employed in the likelihood calculation (Eq. 4.7) as the linear combination of the trivial pixel-by-pixel noise matrix and the global and local kernels discussed above,

$$
C_{ij}(\phi_C) = b \delta_{ij} \sigma_i^2 + K_{ij}^G(\phi_{C,G}) + K_{ij}^L(\phi_{C,L}),
$$

(4.17)
with hyperparameters $\phi_C = [\phi_{C,G}, \phi_{C,L}]$. The factor $b$ is a parameter that scales up the Poisson noise in each pixel by a constant factor to account for additional detector or data reduction uncertainties (e.g., read noise, uncertainties in the spectral extraction procedure, etc.); typically $b < 1.1$ for well-calibrated optical spectra. If there are $N_{loc}$ local covariance patches (see Section 4.2.5 on how this is determined), then there are $4N_{loc} + 2$ elements in the set of covariance hyperparameters, $\phi_C$. Figure 4.5 provides a graphical illustration of how the kernels that comprise the covariance matrix are able to reproduce the structure present in a typical residual spectrum.

4.2.4 Priors

The Bayesian framework of this inference approach permits us to specify prior knowledge about the model parameters, $p(M)$. As will be discussed further in Sections 4.3 and 4.4, in most cases it is necessary to utilize some independent information (e.g., from asteroseismology constraints or stellar evolution models) as a prior on the surface gravity. But otherwise we generally recommend a conservative assignment of uniform priors, such that $p(\theta_*)$ is flat over the spectral library grid (and zero elsewhere) and $p(\theta_{\text{ext}})$ is flat for physically meaningful values (e.g. $v \sin i \geq 0$, $\Omega > 0$, and $A_V \geq 0$).

For (early type) stars with a clear continuum, it makes sense to assume flat priors on the polynomial parameters $\phi_P$. However, information about the calibration accuracy (e.g., from comparisons of multiple calibration sources in the same observation sequence) can be encoded into a simple prior on the Chebyshev coefficients; for example, Gaussian priors with widths that represent the fractional variance between different derived calibration functions would be reasonable. For (late type) stars with a poorly defined
continuum, some judicious tapering of the priors (such that small coefficients at high $n$ are preferred) may be required to ensure that broad spectral features are not absorbed into the polynomial (see Section 4.3).

In general, uniform (non-negative) priors are recommended for the global kernel hyperparameters. For the local kernels, we typically adopt uniform priors for the amplitudes and means $\{a_k, \mu_k\}$, but construct a logistic prior for the widths $\{\sigma_k\}$ that is flat below the width of the line-of-sight velocity distribution function (defined as the

![Probabilistic Graphical Model](image)

Figure 4.7.—: A probabilistic graphical model representing how the parameters of the model are combined to forward model a spectrum and evaluate the likelihood function (Eq. 4.6). When interpolating models using a spectral emulator (Appendix 4.5), the stellar parameters ($\theta_*$) determine the weights ($w$) of the eigenspectra ($\Xi$), which are modified according to the observational parameters ($\theta_{\text{ext}}$) and polynomial parameters ($\phi_P$). Together, these parameters specify the model spectrum ($M$). If one uses linear interpolation instead of a spectral emulator, then there would be no intermediate nodes for $w$ and $\Xi$. The structure of the covariance matrix, which is included in the likelihood function, is determined by the covariance hyperparameters ($\phi_C$). Together, the model spectrum and the covariance matrix predict the resulting dataset ($D$).
convolution of three broadening kernels in Eq. 4.1), \( \sigma_{\text{los}} \), and smoothly tapers to zero at larger values:

\[
p(\sigma_k) = \frac{1}{1 + e^{\sigma_k - \sigma_{\text{los}}}}. \tag{4.18}
\]

Such a prior formulation prevents local kernels from diffusing to large \( \sigma_k \) and low \( a_k \), since that kind of behavior is better treated by the global kernel. When modeling real data, there is no \textit{a priori} information about the locations \( \{\mu_k\} \) of the local kernels; they are instantiated as needed (see Section 4.2.5). However, using the knowledge gained from previous inferences of similar targets, one could instead start by instantiating kernels at the outset with priors on \( \{\mu_k\} \) where there are known to be systematic issues with the synthetic spectra.

### 4.2.5 Exploring the Posterior

The inference framework developed here has a natural blocked structure between the collections of “interesting” parameters, \( \Theta = [\theta^*, \theta_{\text{ext}}] \), the nuisance parameters \( \phi_P \), and the covariance hyperparameters \( \phi_C \). The conditional dependencies of these parameters are shown graphically in Figure 4.7 as a directed acyclic graph (Bishop 2006; Mandel et al. 2009). To explore the posterior distribution,

\[
p(\Theta, \phi_P, \phi_C | D) \propto p(D | \Theta, \phi_P, \phi_C) p(\Theta, \phi_P, \phi_C) \tag{4.19}
\]

for this type of structure, it is convenient to employ Markov Chain Monte Carlo (MCMC) simulations with a blocked Gibbs sampler coupled to the Metropolis-Hastings algorithm. This procedure works by sampling in a subset of parameters (with Metropolis-Hastings proposals) conditioned on the current (fixed) values of the other parameters. After each iteration, the Gibbs sampler updates the sampled parameters and then cycles through
all the (previously fixed) different parameter subsets in the same way (for a more
mathematical description, see Chapter 11 of Gelman et al. 2013). A step-by-step pre-
scription follows, where the $i^{th}$ iteration of the Gibbs sampler is indexed with a superscript:

1) Initialize the parameters. One might set $\Theta^0$ based on estimates in the literature or
scaling behaviors, and make simple assumptions about $\phi_p^0$. Here, we set the Chebyshev
coefficients ($\phi_p^0$) so that the polynomials are constant ($c_o^{(0)} = 1$ and $c_o^{(>0)} = 0$, $\forall o$) and
assume only the trivial noise spectrum (and spectral emulator kernel; see Appendix 4.5)
contributes to the $C$ (i.e., $\phi_C^0 = 0$).

2a) Start the $i^{th}$ iteration of the Gibbs sampler. For each iteration of the Metropolis-
Hastings algorithm, sample in $\Theta$ to evaluate the posterior (Eq. 4.19) following the
framework laid out in Sections 4.2.3 and 4.2.4. This represents a “slice” through the
posterior space conditioned on the other parameters being held fixed ($\phi_p = \phi_p^{i-1}$ and
$\phi_C = \phi_C^{i-1}$). Then update $\Theta^{i-1} \rightarrow \Theta^i$.

2b) For each spectral order, sample in the polynomial parameters $\phi_p$ and covariance
hyperparameters $\phi_C$, conditioned on the other parameters being held fixed $\Theta^i = \Theta^{i-1}$.
Then update $\phi_p^{i-1} \rightarrow \phi_p^i$ and $\phi_C^{i-1} \rightarrow \phi_C^i$.

3) Repeat Step (2) for 20,000 samples.

4) Repeat the procedure in Steps (1)–(3) with different initializations, storing the
samples for each Markov chain. After removing the burn-in samples for each chain, we compute the Gelman-Rubin convergence diagnostic, $\hat{R}$ (Gelman et al. 2013, their Eq. 11.4). If $\hat{R} < 1.1$, we can be reasonably sure that all of the chains have converged to the posterior distribution.

In Step (2b), local covariance kernels are instantiated according to the following procedure. First, an “average” residual spectrum is generated by combining $\sim 500$ residual spectra that were stored during a burn-in period using only the global kernels (prior to this storage, the Markov chain is thinned to account for autocorrelation of the posterior samples). This average spectrum is then iteratively examined for deviations outside a critical threshold. When a large residual is identified, a local kernel is introduced with a mean ($\mu_k$) at its location.\(^3\) After some experimentation with different threshold criteria, we chose to instantiate when the local residual is $>4\times$ the standard deviation in the average residual spectrum.\(^4\) Alternative schemes, such as re-evaluating the kernel locations with each iteration of the Gibbs sampler, yield similar results; however, the adopted approach consistently converges with minimal computational

\(^3\)Although this may seem similar to the procedure of “sigma-clipping”, there is a crucial difference. Rather than rejecting outlier data once it is found (i.e., setting its weight in the inference problem to zero), this procedure will actually self-consistently determine how to weight the outliers inside the probabilistic framework.

\(^4\)Lower thresholds result in more local kernels, thereby reducing the amplitude of the global kernel. In the extreme case of a very low threshold, a local kernel would be instantiated for every spectral line (and no global kernel would be required). We found that ultimately the posterior inferences on the parameters of interest are relatively insensitive to the choice of a threshold level, so long as it is set low enough to capture the egregious outliers.
CHAPTER 4. SPECTROSCOPIC INFERENCE

overhead. Once all the local kernels are instantiated, the Gibbs sampler is run for another period of burn-in.\footnote{There is no practical reason to delete local kernels once instantiated. If the parameters have changed such that a given local kernel is no longer required, that kernel amplitude will be driven towards zero and represent a negligible contribution to $C$; in effect, the model will act as if the kernel were deleted automatically.}

This entire procedure can be a significant computational challenge. A typical spectrum has $N_{\text{pix}} > \mathcal{O}(10^3)$, and therefore the many evaluations of the matrix product $R^T C^{-1} R$ in the likelihood calculation can be numerically expensive. However, because $C$ is a symmetric, positive semi-definite matrix, we can employ Cholesky factorization to optimize the evaluation of the matrix product and avoid the direct calculation of the matrix inversion ($C^{-1}$). For multi-order echelle spectra or multiple spectra of the same target (perhaps taken with different instruments), the nuisance parameters for each segment of the spectrum are independent. This means that the computationally intensive steps of generating a model spectrum for a specific wavelength range and evaluating the likelihood can be parallelized. The only segment of the code that needs to be synchronized is the MCMC proposal of stellar parameters, which are shared between all chunks of the spectrum. The massive parallelization of this algorithm on a computer cluster therefore enables the simultaneous inference of interesting parameters over wide spectral ranges at high resolution, or from multiple datasets. To sample the posteriors in this mode, we extend the Metropolis-Hastings sampler included in the \texttt{emcee} package (Foreman-Mackey et al. 2013) to function within a parallelized blocked Gibbs sampler.

The time required to thoroughly explore the posterior depends on both the data volume and the desired precision on the inference of the covariance hyperparameters. If
only the stellar parameters $\Theta$ are of interest, one can first optimize the kernel parameters and then proceed with them fixed, since the stellar parameter posteriors are relatively insensitive to the precise value of the kernel parameters (once near their optimal value). A fit of an $R \approx 40,000$ spectrum with $>30$ echelle orders takes $\sim 2$ hours (parallelized on a cluster). If the full posteriors for the nuisance parameters are desired, the computation might take an order of magnitude longer.

### 4.3 Demonstrations

In this section, we illustrate how the modeling framework operates for two real datasets. The first is an elaboration of the example shown throughout Section 4.2, using a high resolution optical spectrum of the F5 star (and transiting exoplanet host) WASP-14 (Joshi et al. 2009; Torres et al. 2012). The second uses a medium resolution near-infrared spectrum of the M5 dwarf Gliese 51 (hereafter Gl 51), observed as part of the NASA/IRTF library of spectral standards (Cushing et al. 2005; Rayner et al. 2009). In both cases, we sequentially build up the complexity of the modeling framework to demonstrate how each of the components described in Section 4.2 affects the posteriors on the parameters of interest ($\theta_*$). We adopt the recent incarnation of the PHOENIX library (Husser et al. 2013) for the models, although comment on systematic differences between libraries in Section 4.3.3.
CHAPTER 4. SPECTROSCOPIC INFERENCE

4.3.1 WASP-14

A high resolution ($R \approx 44,000$) optical spectrum of WASP-14 was obtained on 2009 June 14 using the Tillinghast Reflector Echelle Spectrograph (TRES; Fűrész 2008) on the Fred Lawrence Whipple Observatory 1.5 m telescope. TRES delivers an echelle spectrum with 51 orders that cover the full optical range, from 3860–9100 Å. The data were reduced and calibrated using standard techniques in the TRES pipeline (cf., Buchhave et al. 2010; see Torres et al. 2012 for more specific details). At 5100 Å, the S/N is $\sim$150 per resolution element. Following Torres et al. (2012), we focus here on the central three TRES orders, covering $\sim$5100-5400 Å.

We start with a “standard” inference, using the most commonly employed likelihood function (i.e., $\propto \chi^2$, with a trivial covariance matrix using only the Poisson uncertainties). Interpolation in the model library is performed with a basic tri-linear algorithm (in this specific case, $\theta_*$ has only three dimensions). To avoid a prominent systematic (see Section 4.3.3), we fix the surface gravity to $\log g = 4.29$ (with a $\delta$-function prior). This independent prior information comes from the combination of a constraint on the mean stellar density based on exoplanet transit depth measurements and a comparison of optical photometry with stellar models in the color-magnitude diagram (Joshi et al. 2009). The resulting marginal posteriors on $T_{\text{eff}}$ and [Fe/H], listed in Table 4.1 and shown in Figure 4.8, are remarkably narrow – unbelievably so, given how subtly the spectrum changes over such small parameter deviations.

For the second test, we increase the complexity of the covariance matrix by introducing the global kernel treatment discussed in Section 4.2.3. We find non-negligible amplitudes and correlation lengths for these kernels, as would be expected for a typical
correlated residual spectrum. With respect to the standard inference, the uncertainty associated with \( T_{\text{eff}} \) has increased by roughly a factor of three, but the [Fe/H] posterior is only marginally broadened (by \( \sim 50\% \)). Upon closer inspection of the latter, it becomes clear that the posterior has an artificially sharp peak located at a grid point in the model library (\([\text{Fe/H}] = -0.5\)). This ‘noding’ is an artifact of naive interpolation over a sparsely-sampled dimension in the library grid; when the uncertainty in the interpolation itself constitutes a significant fraction of the total error budget, the fit will be driven toward grid points (where the interpolation error is naturally minimized; see also Cottaar et al. 2014). To mitigate this behavior, we need to employ an interpolation scheme that properly incorporates this kind of uncertainty.

Therefore, in a third test we implement the Bayesian emulator described in Appendix 4.5 to propagate uncertainty in the interpolation. This procedure successfully avoids the ‘noding’ behavior in [Fe/H], and inflates the associated uncertainty by a factor of 2.5 compared to the “standard” inference approach. The uncertainty on \( T_{\text{eff}} \) is now 5\( \times \) larger than in the original test.

Finally, in a fourth test we fold in the methodology for the local covariance kernels described in Section 4.2.3. This has little effect on the widths of the parameter posteriors (\( \lesssim 10\% \) increase), but does shift their peaks to slightly higher values in both \( T_{\text{eff}} \) and [Fe/H]. We suspect this is likely driven by a bias in the inference of [Fe/H], produced because the PHOENIX models tend to have more ‘outlier’ spectral lines with over-predicted line depths. Without the local covariance kernels to downweight these outliers, the models tend toward lower metallicity to account for them. But when the local kernels are included, this bias is reduced and a more appropriate higher [Fe/H] value is inferred. Figure 4.9 demonstrates how well the modeling framework can match
Figure 4.8.—: The marginal posterior probability distributions for the WASP-14 $T_{\text{eff}}$ and $[\text{Fe/H}]$ based on the PHOENIX model library, for various levels of model complexity, including: (1) a simple linear interpolation scheme and trivial covariance matrix (blue-green); (2) including global covariance terms from Gaussian process kernels (orange); (3) employing a Bayesian emulator for more appropriate interpolation (purple); and (4) also including local covariance kernels to downweight systematic outlier spectral lines (magenta).
Figure 4.9.—: (top) A representative segment of the TRES spectrum of WASP-14 (blue), overlaid with a PHOENIX model (red) generated by drawing parameters from the posterior distribution (under the assumption of a fixed log $g = 4.29$). (bottom) The corresponding residual spectrum overlaid on contours representing the distributions of a large number of random draws from the covariance matrix (the shading is representative of the 1, 2, and $3\sigma$ spreads of that distribution of draws), as in Fig. 4.5. Note the utility of local patches of increased residual variance in accounting for outlier features, which are introduced by the local covariance kernels described in Sect. 4.2.3.
### Table 4.1: Demonstration Tests for WASP-14

<table>
<thead>
<tr>
<th>Test</th>
<th>Interp</th>
<th>C</th>
<th>$T_{\text{eff}}$</th>
<th>[Fe/H]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>linear</td>
<td>trivial</td>
<td>6280 ± 5</td>
<td>−0.471 ± 0.004</td>
</tr>
<tr>
<td>(2)</td>
<td>linear</td>
<td>$+ \mathcal{K}^G$</td>
<td>6297 ± 16</td>
<td>−0.500 ± 0.006</td>
</tr>
<tr>
<td>(3)</td>
<td>emulator</td>
<td>$+ \mathcal{K}^G$</td>
<td>6281 ± 26</td>
<td>−0.482 ± 0.012</td>
</tr>
<tr>
<td>(4)</td>
<td>emulator</td>
<td>$+ \mathcal{K}^G + \mathcal{K}^L$</td>
<td>6301 ± 29</td>
<td>−0.431 ± 0.012</td>
</tr>
</tbody>
</table>

The best-fit parameter values (peak of the marginal posteriors) and associated (1 $\sigma$) uncertainties (68.3% confidence intervals) for the four tests of increasing complexity described in the text. Note that log $g$ is fixed to 4.29 (cf., Joshi et al. 2009).
the character of the residual spectrum when employing the sophisticated covariance matrix (test 4) advocated here.

### 4.3.2 Gl 51

A moderate resolution \( R \approx 2,000 \) near-infrared spectrum of Gl 51 was obtained on 2000 Nov 6 using the SPEX instrument (Rayner et al. 2003) on the 2.3 m NASA Infrared Telescope Facility (IRTF). SPEX is a cross-dispersed echelle spectrograph that covers the red-optical to thermal-infrared spectrum (0.7–5.5 \( \mu \)m) in two settings. These data were obtained as part of the IRTF spectral standard library project (Cushing et al. 2005; Rayner et al. 2009), and were processed through the well-vetted Spextool reduction pipeline (Cushing et al. 2004; Vacca et al. 2003) to deliver a fully calibrated spectrum. At 2.1 \( \mu \)m, the S/N is \( \sim 400 \) per resolution element.

Modeling late-type stellar atmosphere structures and their spectra is considerably more complex than for Sun-like stars, due to lingering uncertainties in the atmosphere physics and molecular opacities. Especially confounding is the presence of complex condensates (clouds) at the coolest temperatures (Allard et al. 2013), making it considerably more challenging to determine (sub-) stellar properties (Rajpurohit et al. 2014). Various approaches have been taken to infer the key parameters in the face of these difficulties, including iteratively masking regions with poor spectral agreement (e.g., Mann et al. 2013). Astutely, Mann et al. note that such a scheme may exclude regions of the spectrum that contain intrinsically useful information for discriminating between physical properties, and that a more sophisticated approach would weight each spectral region based on its consistency with the data. The modeling framework that we
Figure 4.10.—: The marginal posterior probability distributions for the WASP-14 $T_{\text{eff}}$ and [Fe/H] based on the PHOENIX model library, for various levels of model complexity, including: (1) a simple linear interpolation scheme and trivial covariance matrix (blue-green); (2) including global covariance terms from Gaussian process kernels (orange); (3) employing a Bayesian emulator for more appropriate interpolation (purple); and (4) also including local covariance kernels to downweight systematic outlier spectral lines (magenta).
Figure 4.11.—: The $K$-band SPEX spectrum of Gl 51 (blue) compared with a PHOENIX model (red) generated by drawing parameters from the inferred posterior distribution. (bottom) The residual spectrum along with contours representing the distributions of a large number of random draws from the covariance matrix (the shading is representative of the 1, 2, and 3σ spreads of that distribution of draws), as in Fig. 4.9. Note how the ‘outlier’ features (Na I at 2.21 $\mu$m and Ca I at 2.26 $\mu$m) are identified and treated by the local covariance kernels.
Table 4.2: Demonstration Tests for Gl 51

<table>
<thead>
<tr>
<th>Test</th>
<th>Interp</th>
<th>C</th>
<th>$T_{\text{eff}}$</th>
<th>[Fe/H]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>linear</td>
<td>trivial</td>
<td>3256 ± 3</td>
<td>0.89 ± 0.01</td>
</tr>
<tr>
<td>(2)</td>
<td>linear</td>
<td>+ $K^G$</td>
<td>3022 ± 35</td>
<td>0.00 ± 0.03</td>
</tr>
<tr>
<td>(3)</td>
<td>emulator</td>
<td>+ $K^G$</td>
<td>3230 ± 30</td>
<td>0.27 ± 0.03</td>
</tr>
<tr>
<td>(4)</td>
<td>emulator</td>
<td>+ $K^G + K^L$</td>
<td>3180 ± 35</td>
<td>0.28 ± 0.04</td>
</tr>
</tbody>
</table>

The best-fit parameter values and associated (1σ) uncertainties (as in Table 4.1) for the four tests of increasing complexity described in the text. Note that log $g$ is fixed to 5.0 (cf., Rojas-Ayala et al. 2012).
As another demonstration of this framework, we carried out the sequence of tests outlined in the previous section for the \( K \)-band portion of the SPEX spectrum of Gl 51. Following the analysis of similar data for this star by Rojas-Ayala et al. (2012), we fix the surface gravity to \( \log g = 5.0 \) based on a comparison with standard stellar evolution models. The test results are listed in Table 4.2; the posteriors for \( T_{\text{eff}} \) and \([\text{Fe/H}]\) are shown together in Figure 4.10. Like WASP-14, we find that a more appropriate treatment of the covariance matrix results in a substantial broadening of the parameter posteriors; the uncertainties on \( T_{\text{eff}} \) and \([\text{Fe/H}]\) are inflated by a factor of \( \sim 10 \) and 4, respectively.

However, in this case the parameter values (posterior peaks) also exhibit substantial movement along the sequence of tests. The underlying cause of this behavior lies with the Na I and Ca I resonance line depths, which are systematically under-predicted in the PHOENIX library (even for high metallicities; see also Rojas-Ayala et al. 2012). Rajpurohit et al. (2010) suggest that these discrepancies may be the consequence of inaccurate atomic data (oscillator strengths and/or opacities). In the first test with a trivial covariance matrix, these ‘outlier’ lines drive the model to favor a very high \([\text{Fe/H}]\). But when we consider the more sophisticated versions of \( \mathbf{C} \) that employ Gaussian processes to treat correlated residuals, the contribution of these features to the likelihood calculation is reduced, and therefore \([\text{Fe/H}]\) returns to a more appropriate range. Because this portion of the spectrum has only these two outlier features, their influence can be mitigated either with a larger global covariance kernel amplitude \( (a_G) \), or with a smaller \( a_G \) but significant contributions from local covariance kernels (which explains why there
is little difference between the posteriors in the third and fourth tests in the sequence). For reference, Figure 4.11 compares the observations with draws from the posterior distribution for the advocated modeling approach (corresponding to the fourth test).

The methodology behind the likelihood calculations we have developed could prove especially useful for spectroscopic inferences of the parameters of cool stars like Gl 51, where substantial uncertainties in their more complex atmospheres will naturally produce systematic deviations between models and data. However, many of those discrepancies will be manifested in molecular features, which likely result in considerably more complex residual structures than noted here (e.g., the TiO bands in the red-optical; see Mann et al. 2013, their Fig. 9). The overall framework we have employed should still function, although more appropriate local covariance kernels may need to be developed to capture the different nature of these outliers. For example, one might employ hybrid kernels (like the product of a truncated exponential and a Matérn kernel) or empirically-motivated parametric shapes (e.g., a saw-tooth pattern) to provide a better representation than a simple Gaussian feature.

4.3.3 Synopsis and Systematics

The results of the sequence of tests in the previous two sections illustrate some key issues in the spectroscopic inference of stellar parameters. First, the residual spectra derived from (typically) imperfect models exhibit correlated structure (e.g., see Fig. 4.4) that cannot be explained well with a trivial (diagonal) covariance matrix. If that naive assumption is made (as is usually the case), the resulting posteriors are unrealistically narrow and may end up being biased (particularly for [Fe/H] or for cases influenced by
prominent ‘outlier’ lines).

This issue of implausibly small formal uncertainties has long been recognized in the stellar spectroscopy community. The standard solution has been to add (in quadrature) a ‘floor’ contribution, imposed independently on each parameter and meant to be representative of the systematics (e.g., see Torres et al. 2012 or Schönrich & Bergemann 2014 for clear and open discussions of this approach). The key problems with this tactic are that these systematics are in reality degenerate (and so should not be applied independently) and that they dominate the uncertainty budget, but are in a large sense arbitrary – they are not self-consistently derived in the likelihood framework. Our goal here has been to treat one aspect of this systematic uncertainty budget internal to the forward-modeling process, by employing a non-trivial covariance matrix that accounts for generic issues in the pixel-by-pixel inference problem. Given the results above, we have demonstrated that this procedure successfully accounts for a substantial fraction of the (empirically motivated) ad hoc systematic ‘floor’ contribution typically adopted in inference studies.

However, although a likelihood function that can properly account for the character of the residuals is important, it does not by itself treat all of the important kinds of systematics in the general spectroscopic inference problem. In future work that can build on the flexible likelihood formalism we have advocated here, there are three other important sources of systematic uncertainty that should be considered: (1) data calibration; (2) optimized parameter sensitivity; and (3) model assumptions, or flexibility. We discuss each of these issues briefly, with attention paid to potential remedies that fit within the likelihood framework developed here.
Perhaps the most familiar source of systematics lies with issues in the data calibration. In the idealized case of perfect calibration, the physical parameters inferred from different observations of the same (static) source should be indistinguishable. But given the complexity of a detailed spectroscopic calibration, that is not typically the case in practice. The common approach to quantify the systematic uncertainties contributed by calibration issues is to compare the inferences made using different spectra (e.g., from different instruments and/or observations). The final parameter values are usually presented as an average of these separate inferences, with the uncertainties inflated by adding in quadrature some parameter-independent terms that account for their dispersion. The more appropriate way of combining these inferences is to model the individual spectra simultaneously in a hierarchical framework like the one discussed in Section 4.2: in that way, the dispersion is appropriately propagated into the parameter uncertainties while any intrinsic degeneracies are preserved (which is not possible in the standard ‘weighted average’ approach). Ultimately, one could also introduce some empirically-motivated nuisance parameters that are capable of forward-modeling imperfections in the data calibration, similar to the approach adopted in Section 4.2.2 (e.g., see Fig. 4.3).

Another important source of systematic bias comes from the fact that certain physical parameters have only a relatively subtle effect on the spectrum. Stellar spectroscopists are familiar with this being an issue when inferring the surface gravity, \( \log g \), since it is primarily manifested as low-level modifications to the wings of certain spectral lines like Mg b and in the equivalent widths of lines from singly-ionized elements like Ti II and Fe II. When modeling data with a large spectral range, the effects of varying \( \log g \) are small compared to the residuals introduced by the many other model
imperfections. Consequently, the surface gravity will not be constrained well, and
inferences on log $g$ (and therefore other degenerate parameters) can be substantially
biased. As an example, when fitting the WASP-14 data in Section 4.3.1 without prior
information on the surface gravity, we find a shift of $\sim 0.9$ dex to lower log $g$ (and
accompanying shifts in $T_{\text{eff}}$ and [Fe/H]). If we instead use a customized version of the
Castelli & Kurucz (2004) models designed to more accurately reproduce this part of the
optical spectrum for Sun-like stars (as employed by SPC; Buchhave et al. 2012), there is
still a 0.2 dex shift compared to the independent, accurate constraints from the transiting
planet (Joshi et al. 2009). Similar work with larger samples indicate a typical scatter in
the log $g$ values inferred solely from spectra relative to independent, accurate constraints
from other data ($\sim 0.5$ dex; Cottaar et al. 2014; Schönrich & Bergemann 2014).

There are two commonly utilized, and not mutually exclusive, approaches to
mitigating this kind of bias. First is the judicious use of a prior, based on either
independent and accurate measurements (e.g., from asteroseismology, dynamical masses
and distances, etc.) or stellar evolution models (as is demonstrated here). Of course,
such information is unfortunately not always readily available for the target of interest.
A second approach is to severely limit the spectral range of the data being modeled,
focusing primarily on those spectral features especially sensitive to the parameter of
interest. But that carries its own risk, since the models derived from the inferred
posteriors might well be wildly inconsistent with the rest of the spectrum. Recently,
Brewer et al. (2015) proposed a sophisticated, iterative approach that apparently resolves
this issue, employing a sequence of conditional inferences based on sets of specific spectral
features that are especially sensitive to individual physical parameters. This seems like a
promising component for future inclusion in the likelihood framework we have developed.
in Section 4.2.

Finally, and perhaps most significant, there are also sources of systematic bias and uncertainty introduced by limitations in the synthetic stellar models themselves. Different models make varied assumptions in their treatments of the atmosphere structures, boundary conditions (e.g., convection), fundamental atomic/molecular data (e.g., opacities), and radiative transfer. Taken together, these variations produce notably different model spectra for the same values of the physical parameters. As a benchmark for estimating the scope of this source of bias, we re-performed the inference described in the fourth test of Section 4.3.1, but using the customized Castelli & Kurucz (2004) model library instead of the Husser et al. (2013) library. The resulting inferences for $T_{\text{eff}}$ and [Fe/H] are in excellent agreement with those derived by Torres et al. (2012) using the SPC method, but are shifted by 150 K (higher) and 0.15 dex (higher), respectively, compared to the PHOENIX results. While the relevant physics included in these models is very similar for these temperatures and the inferred stellar parameters are similar in an absolute sense, it is still striking that the systematic shift between models is several times larger than the statistical uncertainties derived from our likelihood function. At this point, there is little to be done to rectify these model-dependent differences; in the future, one hopes that the model inputs can be refined based on feedback from the data (see Sect. 4.4). Any inferences of physical parameters should only be considered in the context of the assumed models.

Aside from these different assumptions and inputs, the limited flexibility of these models certainly also contributes to the systematic uncertainty budget, and is possibly also a source of systematic bias. Model spectral libraries typically have neglected dimensions in parameter-space that, if made available, would be expected to broaden and
possibly shift the posteriors for the primary physical parameters. One typical example lies with element-specific abundance patterns, often distilled to the enhancement of $\alpha$-elements (i.e., [$\alpha$/Fe]). If the target star has a non-zero [$\alpha$/Fe] (an enhancement or deficit relative to the solar ratios), but is fit with a single, global metallicity pattern, it is not clear that the sophisticated covariance formalism developed here would be capable of appropriately capturing such residual behavior. Another prominent example of an important hidden parameter dimension is the microturbulence, which for some spectral types and spectral resolution may impact the spectrum in a similar way as the surface gravity (and may be partly responsible for the log $g$ bias discussed above; Gray et al. 2001). To mitigate the resulting deficiencies in precision (and potentially accuracy) on the inference of other parameters, we would ideally employ libraries or modeling front-ends that can incorporate some flexibility in these hidden (i.e., ignored) dimensions of parameter-space (e.g., individual elemental or group-based abundance patterns, microturbulence, etc.).

4.4 Discussion

Astronomers exploit spectroscopy to retrieve physical information about their targets. Ideally, such inferences are made with the maximal precision afforded by the measurement noise, and accurately reflect the uncertainties with minimal systematic bias. But in practice, the spectral models used as references are never perfect representations. Even modest mismatches between data and model can propagate substantial systematic uncertainty into the inference problem. In high-sensitivity applications (e.g., stellar and exoplanetary astrophysics), ignoring these systematics can give a false sense of both
precision and accuracy in the inferences of key parameters. Typically, the more egregious
of these imperfections are “mitigated” by dismissal (explicitly not considering a subset of
the data; e.g., masking, clipping). Rarely, they are confronted directly with painstaking,
computationally expensive fine-tuning of more general (nuisance) parameters in the
model (e.g., oscillator strengths, opacities), albeit only over a very limited spectral range
and region of physical parameter-space.

We have presented an alternative approach to dealing with this fundamental issue,
grounded in a generative Bayesian framework. The method advocated here constructs
a sophisticated likelihood function, employing a non-trivial covariance matrix to treat
the correlated pixel-to-pixel residuals generated from intrinsically imperfect models.
That matrix is composed of a linear combination of global (stationary) and local
(non-stationary) Gaussian process kernels, which parameterize an overall mild covariance
structure as well as small patches of highly discrepant outlier features. In the context of
a given model parameterization (i.e., synthetic spectral library, or a more complex and
flexible model generator), the framework we have developed provides a better inference
than the standard $\chi^2$ (or cross-correlation) comparison. We have built up a series of
tests that demonstrates how the emulator, global kernels, and local kernels affect the
nature of the inference on the stellar parameters. To demonstrate how the framework
is used, we determined the surface parameters of main-sequence stars with mid-F and
mid-M spectral types from high-S/N optical and near-infrared data, with reference to
pre-computed model libraries (Sect. 4.3). The source code developed here is open and
freely available for use: see http://iancze.github.io/Starfish.

The novelty of employing this kind of likelihood function in the spectroscopic
inference problem is that the treatment of data–model mismatches (in essence, the fit
quality) is explicitly built into the forward-modeling framework. This offers the unique advantage that discrepant spectral features (outliers), which may contain substantial (even crucial) information about the parameters of interest, can still effectively propagate their useful information content into the posteriors with a weighting that is determined self-consistently. From a practical standpoint, this means that a larger spectral range can be used and model imperfections can be downweighted by the usage of covariance kernels. The global covariance framework provides more appropriate estimates of the posterior probability distribution functions (i.e., the precision or uncertainty estimates) for the model parameters. The automated identification and disciplined downweighting of problematic “outlier” spectral lines (those that cannot be reproduced with any combination of the model parameters) with local covariance kernels can prevent them from overly influencing (and possibly biasing, especially in cases with few spectral features available) the inferences. In many cases, the underlying physical problem lies with incorrect (or inaccurate) atomic and/or opacity data used in the models. In this sense, the posteriors of the hyperparameters of the local covariance kernels can actually indicate in what sense and scale these inputs need to be modified to better reproduce observational reality.

The approach we describe is generally applicable to any spectroscopic inference problem (e.g., population synthesis in unresolved star clusters or galaxies, physical/chemical models of emission line spectra in star-forming regions, etc.). Moreover, it has the flexibility to incorporate additional information (as priors) or parametric complexity (if desired), and could be deployed as a substitute for a simplistic $\chi^2$ metric in already-established tools (e.g., SME). Another potential application might be in the estimation of radial velocities using traditional Doppler-tracking pipelines for
CHAPTER 4. SPECTROSCOPIC INFERENCE

exoplanet or binary star research. Poorly modeled micro-tellurics can lead to incorrect measurements of radial velocities for certain contaminated chunks of the spectrum, causing them to give unrealistically precise but biased velocity measurements. A flexible noise model would broaden the posteriors on these points and allow them to be combined into a more accurate systemic velocity.

Ultimately, the benefits of employing covariance kernels to accommodate imperfect models could be extended well beyond modeling the spectra of individual targets. In principle, the approach we have described here can be used to systematically discover and quantify imperfections in spectral models and eventually to build data-driven improvements of those models that are more appropriate for spectroscopic inference. By fitting many stellar spectra with the same family of models, we can catalog the covariant structure of the fit residuals – especially the parameters of the local covariance kernels – to collate quantitative information about where and how the models tend to deviate from observational reality. That information can be passed to the spectral synthesis community, in some cases enabling modifications that will improve the quality of the spectral models. On a large enough scale, this feedback between observers and modelers could be used to refine inputs like atomic and molecular data (oscillator strengths, opacities), elemental abundance patterns, and perhaps the stellar atmosphere structures. If one has access to the radiative synthesis process that generates the model spectra, there are many possible means to improve their quality. In particular, a process of history matching can be used to rule out regions of parameter space where the models do not fit well (e.g., for a use in galaxy formation simulations, see Vernon et al. (2010)). For example, if one had full control over the radiative synthesis code, stellar structure code, and atomic line database, one could improve the performance of the spectral emulator by
CHAPTER 4. SPECTROSCOPIC INFERENCE

ruling out regions of parameter space for these separate components that are inconsistent with a collection of observed spectra, such as a set of standard stars spanning the full range of spectral classifications.

In a similar vein, we could also simultaneously use several synthetic spectral libraries to infer the stellar parameters while also identifying discrepant regions of the spectrum. A treatment using multiple synthetic libraries would likely reveal interesting correlations between model discrepancies, such as a specific signature among many lines (e.g. deviations in spectral line shape that cannot be explained by variations in $\theta$). Conversely, if a discrepant feature is seen for all models, it could be due to either an anomaly with the given star (e.g., a chromospheric line due to activity or perhaps an intervening interstellar absorption line) or a correlated difficulty among all models (e.g., an incorrect atomic constant).

Alternatively, this kind of feedback could be used to make data-driven modifications to the already existing models, creating a new semi-empirical model library. This could be accomplished by linking the parameters of the covariance kernels while fitting many stars of similar spectral type in a hierarchical Bayesian model, which would add confidence to the assessment that certain spectral features are systematic outliers and offer general quantitative guidance on how to weight them in the likelihood calculation. Rather than simply assembling an empirical spectral library using only observations, this combined machine-learning approach would naturally provide a physical anchoring for the key physical parameters, since they are reflected in the spectra based on the physical assumptions in the original models. This kind of large-scale analysis holds great promise in the (ongoing) era of large, homogeneous high resolution spectroscopic datasets (e.g., like those being collected in programs like the APOGEE and HERMES surveys; Nidever}
et al. 2012; Zucker et al. 2012), since they provide enormous leverage for identifying and improving the underlying model systematics.

4.5 Spectral Emulator for Interpolation

The spectral emulator is designed to serve as an improved interpolator for the synthetic spectral library. Rather than a (tri-)linear interpolator, which would deliver a single spectrum for a given $\theta^*$, the spectral emulator delivers a probability distribution of possible interpolated spectra. In this manner, it is possible to incorporate realistic uncertainties about the interpolation process into the actual likelihood calculation. In the limit of moderate to high signal-to-noise spectra, these interpolation uncertainties can have a significant effect on the posterior distribution of $\theta^*$. A schematic of the emulator is shown in Figure 4.12, which is a continuation of Figure 4.1. Briefly, the emulator consists of a set of eigenspectra, representing the synthetic spectral library, that can be summed together with different weights to reproduce any spectrum originally in the library. To produce spectra that have $\theta^*$ in between $\{\theta^*\}^{\text{grid}}$, the weights are modeled with a smooth Gaussian process (GP). This GP delivers a probability distribution over interpolated spectra, which can then be incorporated into the covariance matrix introduced in Section 4.2.3. Here we describe the design and construction of our spectral emulator.

Model library spectra are stored as (1-dimensional) arrays of fluxes, sampled on high resolution wavelength grids. In the case of interest here, the sets of model parameters $\{\theta^*\}^{\text{grid}} = [T_{\text{eff}}, \log g, \text{[Fe/H]}]$ define the dimensions of the library grid. The full spectral
Figure 4.12.—: A continued flowchart explaining the contribution of the spectral emulator to the likelihood function. The synthetic library is first decomposed into an eigenspectrum basis. Then, the extrinsic parameters $\theta_{\text{ext}}$ modify the eigenspectra. The intrinsic stellar parameters $\theta_*$ are fed into a Gaussian process (GP), which delivers a probability distribution of weights used to sum the eigenspectra. The mean weights can be used to reconstruct a mean model spectrum, while the variances of the weights are used to propagate interpolation uncertainties into the likelihood function.

library, $f_\lambda(\{\theta_*\}^{\text{grid}})$, is therefore encapsulated in a 4-dimensional array. The libraries used here have grid spacings of 0.5 dex in log $g$ and 0.5 dex in [Fe/H]; the Castelli &
CHAPTER 4. SPECTROSCOPIC INFERENCE

Figure 4.13.—: (top) The mean spectrum, standard deviation spectrum, and five eigenspectra that form the basis of the PHOENIX synthetic library used to model Gl 51, generated using a subset of the parameter-space most relevant for M dwarfs. (bottom) The original synthetic spectrum from the PHOENIX library ($\theta_* = [T_{\text{eff}} = 3000\,\text{K}, \log g = 5.0\,\text{dex}, [\text{Fe/}\text{H}] = 0.0\,\text{dex}]$) compared with a spectrum reconstructed from a linear combination of the derived eigenspectra using Eqn 4.23 (with the weights $w_k$ listed in the top panel figure).

Kurucz (2004) library steps by 250 K in $T_{\text{eff}}$, but the PHOENIX library has finer coverage in 100 K increments.

The first step in designing a spectral emulator is to break down the library into
an appropriate basis (Habib et al. 2007; Heitmann et al. 2009). We chose the principal component basis to decompose the library into a set of “eigenspectra”, following the techniques of Ivezić et al. (2013). Prior to this decomposition, we isolate a subset of the library (containing $M$ spectra) with parameter values that will be most relevant to the target being considered (e.g., for Gl 51, this means considering only effective temperatures below $\sim 3800$ K). We then standardize these spectra by subtracting off their mean spectrum and then “whiten” them by dividing off the standard deviation spectrum measured in each pixel across the grid. The mean spectrum is

$$
\xi_\mu = \frac{1}{M} \sum_{i=1}^{M} f_\lambda(\{\theta_\ast\}_{i}^{\text{grid}})
$$

and the standard deviation spectrum is

$$
\xi_\sigma = \sqrt{\frac{1}{M} \sum_{i=1}^{M} [f_\lambda(\{\theta_\ast\}_{i}^{\text{grid}}) - \xi_\mu]^2},
$$

where $\{\theta_\ast\}_{\text{grid}}$ denotes the full collection of the $M$ sets of stellar parameters under consideration in the library and $\{\theta_\ast\}_{i}^{\text{grid}}$ denotes a single set of those parameters drawn from this collection. Both $\xi_\mu$ and $\xi_\sigma$ are vectors with length $N_{\text{pix}}$, the same size as a raw synthetic spectrum ($f_\lambda$). In effect, all library spectra are standardized by subtracting the mean spectrum and dividing by the standard deviation spectrum

$$
\hat{f}_\lambda(\{\theta_\ast\}_{\text{grid}}) = \frac{f_\lambda(\{\theta_\ast\}_{\text{grid}}) - \xi_\mu}{\xi_\sigma}.
$$

The eigenspectra are computed from this standardized grid using principal component analysis (PCA; Ivezić et al. 2013). Each eigenspectrum is a vector with length $N_{\text{pix}}$, denoted as $\xi_k$, where $k$ is the principal component index $k = \{1, 2, \ldots, m\}$. We decided to truncate our basis to the first $m$ eigenspectra, where $m$ is decided by the minimum number of eigenspectra required to reproduce any spectrum in the grid to
better than 2\% accuracy for all pixels (the typical error for any given pixel is generally much smaller than this, \( \lesssim 0.5\% \)). As an example, the eigenspectra basis computed for Gl 51 using the Phoenix library is shown in the top panel of Figure 4.13.

Using the principal component basis, we can lossily reconstruct any spectrum from the library with a linear combination of the eigenspectra

\[
    f_{\lambda}(\{\theta_{*}\}_{i}^{\text{grid}}) \approx \xi_{\mu} + \xi_{\sigma} \sum_{k=1}^{m} w_{k}(\{\theta_{*}\}_{i}^{\text{grid}}) \xi_{k}
\]

where \( w_{k} \) is the weight of the \( k^{th} \) eigenspectrum. These weights are 3-dimensional scalar functions that depend on the stellar parameters \( \theta_{*} \). Any given weight, which is generally a smooth function of the stellar parameters (see the left panel of Figure 4.14), can be determined at any grid point in the library by taking the dot product of the standardized synthetic spectrum with the eigenspectrum

\[
    w_{k}(\{\theta_{*}\}_{i}^{\text{grid}}) = \sum_{\lambda} \hat{f}_{\lambda}(\{\theta_{*}\}_{i}^{\text{grid}}) \xi_{k}.
\]

To simplify notation, we can write the collection of eigenspectra weights in a length-\( m \) column vector

\[
    w(\theta_{*}) = \begin{bmatrix}
        w_{1}(\theta_{*}) \\
        w_{2}(\theta_{*}) \\
        \vdots \\
        w_{m}(\theta_{*})
    \end{bmatrix}
\]

and horizontally concatenate the eigenspectra into a matrix with \( N_{\text{pix}} \) rows and \( m \) columns

\[
    \Xi = \begin{bmatrix}
        \xi_{1} & \xi_{2} & \cdots & \xi_{m}
    \end{bmatrix}.
\]

Then, we can rewrite Eq. (4.23) as

\[
    f_{\lambda}(\{\theta_{*}\}_{i}^{\text{grid}}) \approx \xi_{\mu} + \xi_{\sigma} \circ \left( \Xi w(\theta_{*})^{\text{grid}} \right)
\]
Figure 4.14.—: The Gaussian process modeling of the principal component weights for the Gl51 PHOENIX spectral library. (left) the blue dots mark the weights $w^\text{grid}_2$ of the 2nd eigenspectrum $\xi_2$ computed at a one-dimensional slice of the spectral library for grid points with $\log g = 5.0$, $[\text{Fe/H}] = 0.0$ and various values of $T_{\text{eff}}$. In reality, the weights are a three-dimensional function of $\theta_\ast$. The thin blue lines show 50 random draws of possible functional forms described by the Gaussian process. (inset) a zoomed portion showing the scatter in the possible functional forms. The black vertical line represents a slice through the scatter of the predicted weight value at $\theta_\ast = [T_{\text{eff}} = 3150 \text{ K}, \log g = 5.0 \text{ dex}, [\text{Fe/H}] = 0.0 \text{ dex}]$. (right) The posterior predictive probability of the collection of all weights $w$ at this value of $\theta_\ast$ is completely described by Eqn 4.49, allowing us to analytically marginalize over all probable values of the weights, and thus marginalize over all probable spectral interpolations.
CHAPTER 4. SPECTROSCOPIC INFERENCE

where $\circ$ represents the element-wise multiplication of two vectors.

To recapitulate, the framework described above can be used to decompose the synthetic spectra in a model library into a principal component basis, allowing us to reconstruct any spectrum in the library as a (weighted) linear combination of $m$ eigenspectra. The weights corresponding to each eigenspectrum are moderately-smooth scalar functions of the three stellar parameters, $\theta_*$. Therefore, to create a spectrum corresponding to an arbitrary set of these parameters that is not represented in the spectral library, we must interpolate the weights to this new set. In practice, it may be possible to use a traditional scheme like spline interpolation to do this directly. However, we found that with sensitive spectra (e.g., for Gl 51 the S/N is $>400$), the uncertainty in the interpolated representation of the spectrum can constitute a significant portion of the total uncertainty budget. This, combined with the under-sampling of the synthetic grid can cause artificial “noding” of the posterior near grid points in the synthetic library, because the interpolated spectrum is not as good as the raw spectrum at the grid point. Even explicitly accounting for interpolation error by doing “drop-out” interpolation tests and empirically propagating it forward does not relieve this noding issue. So instead, we address this problem by employing a Gaussian process to model the interpolation of the eigenspectra weights over $\theta_*$, thereby encapsulating the range of possible interpolated spectra.

Each weight is modeled by a Gaussian process for each eigenspectrum. For a single eigenspectrum with index $k$, we denote the collection of $w_k(\{\theta_*\}^\text{grid}_i)$ evaluated for all the spectra in the library as a length $M$ vector $w_k^\text{grid}$. The Gaussian process treats $w_k^\text{grid}$ as a collection of random variables drawn from a joint multi-variate Gaussian distribution
CHAPTER 4. SPECTROSCOPIC INFERENCE

(Rasmussen & Williams 2005),

$$w_k^{\text{grid}} \sim \mathcal{N} \left( 0, \Sigma_k^{\text{grid}} \right),$$

(4.28)

with $\Sigma_k^{\text{grid}}$ denoting the covariances. The kernel that describes the covariance matrix for this distribution is assumed to be a 3-dimensional squared exponential,

$$K(\theta_i, \theta_j | \phi_{\text{int},k}) = a_{\text{int}}^2 \exp \left[ -\frac{(T_{\text{eff}i} - T_{\text{eff}j})^2}{2 \ell_{T_{\text{eff}}}^2} \right] \times \exp \left[ -\frac{(\log g_i - \log g_j)^2}{2 \ell_{\log g}^2} \right] \times \exp \left[ -\frac{([\text{Fe/He}]_i - [\text{Fe/He}]_j)^2}{2 \ell_{[\text{Fe/He}]}^2} \right],$$

(4.29)

with hyperparameters $\phi_{\text{int},k} = \{a_{\text{int}}, \ell_{T_{\text{eff}}}, \ell_{\log g}, \ell_{[\text{Fe/He}]}\}$ representing an amplitude and length scale for each dimension of $\theta_i$. Unlike the Matérn kernel used in Section 4.2 (which produces a more structured behavior reminiscent of the spectral residuals), this squared exponential kernel has a smooth functional form that is more appropriate to represent the behavior of the eigenspectra weights across the library grid, as demonstrated in Figure 4.14. The $M \times M$-dimensional covariance matrix is

$$\Sigma_k^{\text{grid}} = K(\{\theta_i\}^{\text{grid}}, \{\theta_j\}^{\text{grid}} | \phi_{\text{int},k}),$$

(4.30)

the evaluation of the covariance kernel for all pairings of stellar parameters at library gridpoints.

Once the Gaussian processes for each $k$ are specified, we can construct the joint distribution.

$$\begin{bmatrix} w_1^{\text{grid}} \\ \vdots \\ w_m^{\text{grid}} \end{bmatrix} \sim \mathcal{N} \left( 0, \begin{bmatrix} \Sigma_1^{\text{grid}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \Sigma_m^{\text{grid}} \end{bmatrix} \right),$$

(4.31)
We use $\mathbf{w}^\text{grid}$ to denote the concatenation of $\mathbf{w}_k^\text{grid}$ vectors into a single length $Mm$ vector, and $\Sigma^\text{grid}$ as the $Mm \times Mm$ covariance matrix,

$$
\mathbf{w}^\text{grid} \sim \mathcal{N}(0, \Sigma^\text{grid})
$$

(4.32)

Although we could optimize the hyperparameters of each Gaussian process independently based upon how well it reproduces the collection of weights for that eigenspectrum, ideally we would like to optimize the hyperparameters according to a metric that describes how well the emulator actually reproduces the original library of synthetic spectra.

Following Habib et al. (2007), we write down a likelihood function describing how well the reconstructed spectra match the entirety of the original synthetic grid

$$
\mathcal{L}(\mathcal{F}|\mathbf{w}^\text{grid}, \lambda_\xi) \propto \lambda_\xi^{MN_{\text{pix}}/2} \exp \left[ -\frac{\lambda_\xi}{2} (\mathcal{F} - \Phi \mathbf{w}^\text{grid})^T (\mathcal{F} - \Phi \mathbf{w}^\text{grid}) \right]
$$

(4.33)

Here, $\mathcal{F}$ represents a length $MN_{\text{pix}}$ vector that is the collection of all of the synthetic flux vectors concatenated end to end. The precision of the eigenspectra basis representation, or the statistical error in the ability of the emulator to reproduce the known eigenspectra is represented by $\lambda_\xi$. Because we have truncated the eigenspectra basis to only $m$ components, where $m < M$ is much smaller than the number of raw spectra in the library, the emulator will not be able to reproduce the synthetic spectra perfectly. By including this “nugget” term in the emulator, we are also forward propagating the interpolation uncertainty for $\theta_*$ near or at values of $\{\theta_*\}^\text{grid}$. We specify a broad $\Gamma$ function prior on $\lambda_\xi$ because we expect it to be well constrained by the data.

$$
p(\lambda_\xi) = \Gamma(a_{\lambda_\xi}, b_{\lambda_\xi})
$$

(4.34)
where shape $a_{\lambda\xi} = 1$ and rate $b_{\lambda\xi} = 0.0001$. To facilitate the manipulation of Eqn 4.33, we create a large $MN_{\text{pix}} \times Mm$ matrix that contains all of the eigenspectra

$$\Phi = [I_M \otimes \xi_1, \ldots, I_M \otimes \xi_m] \quad (4.35)$$

where $\otimes$ is the Kronecker product. This operation creates a matrix, which, when multiplied by the vector $\mathbf{w}^{\text{grid}}$, enables (lossy) reconstruction of the entire synthetic library

$$\mathcal{F} \approx \Phi \mathbf{w}^{\text{grid}} \quad (4.36)$$

up to truncation error in the eigenspectrum basis ($\lambda\xi$). For a given $\lambda\xi$, the maximum likelihood estimate for Eqn 4.33 is $\hat{\mathbf{w}}^{\text{grid}} = (\Phi^T\Phi)^{-1}\Phi^T\mathcal{F}$. Using $\hat{\mathbf{w}}^{\text{grid}}$, we can factorize Eqn 4.33 into

$$L(\mathcal{F}|\mathbf{w}^{\text{grid}}, \lambda\xi) \propto \lambda_{\xi}^{Mm/2} \exp \left[ -\frac{\lambda\xi}{2} (\mathbf{w}^{\text{grid}} - \hat{\mathbf{w}}^{\text{grid}})^T (\Phi^T\Phi) (\mathbf{w}^{\text{grid}} - \hat{\mathbf{w}}^{\text{grid}}) \right] \times \lambda_{\xi}^{M(N_{\text{pix}} - m)/2} \exp \left[ -\frac{\lambda\xi}{2} \mathcal{F}^T (I - \Phi(\Phi^T\Phi)^{-1}\Phi^T) \mathcal{F} \right]$$

(4.37)

Now, only the middle line of this distribution depends on $\hat{\mathbf{w}}^{\text{grid}}$, so we can reformulate this equation into a dimensionality reduced likelihood function and absorb the other terms into a modified prior on $\lambda\xi$.

$$L(\hat{\mathbf{w}}^{\text{grid}}|\mathbf{w}^{\text{grid}}, \lambda\xi) \propto \lambda_{\xi}^{Mm/2} \exp \left[ -\frac{\lambda\xi}{2} (\mathbf{w}^{\text{grid}} - \hat{\mathbf{w}}^{\text{grid}})^T (\Phi^T\Phi) (\mathbf{w}^{\text{grid}} - \hat{\mathbf{w}}^{\text{grid}}) \right]$$

(4.38)

To summarize, we have reduced the dimensionality of the distribution from

$$L(\mathcal{F}|\mathbf{w}^{\text{grid}}, \lambda\xi) = \mathcal{N}(\mathcal{F}|\Phi\mathbf{w}^{\text{grid}}, \lambda_{\xi}^{-1}I_{MN_{\text{pix}}})$$

(4.39)
CHAPTER 4. SPECTROSCOPIC INFERENCE

\[
\mathcal{L} (\mathbf{w}^{\text{grid}} | \mathbf{w}^{\text{grid}}, \lambda) = \mathcal{N} (\mathbf{w}^{\text{grid}} | \mathbf{w}^{\text{grid}}, (\lambda \Phi^T \Phi)^{-1})
\]  

(4.40)

Although we introduced the likelihood function in Eqn 4.33, we have yet to include the Gaussian processes or the dependence on the emulator parameters \( \phi_{\text{int}} \). We do this by multiplying Eqn 4.40 with our prior distribution on the weights (Eqn 4.32),

\[
p (\mathbf{w}^{\text{grid}} | \lambda, \phi_{\text{int}}) = \\
\mathcal{N} (\mathbf{w}^{\text{grid}} | \mathbf{w}^{\text{grid}}, (\lambda \Phi^T \Phi)^{-1}) \mathcal{N} (\mathbf{w}^{\text{grid}} | 0, \Sigma^{\text{grid}})
\]

(4.41)

and integrate out the dependence on \( \mathbf{w}^{\text{grid}} \). We perform this integral using Eqn A.7 of Rasmussen & Williams (2005) for the product of two Gaussians, which yields

\[
p (\hat{\mathbf{w}}^{\text{grid}} | \lambda, \phi_{\text{int}}) = (2\pi)^{-Mm/2} | \left( \lambda \Phi^T \Phi \right)^{-1} + \Sigma_w |^{-1/2} \\
\times \exp \left[ -\frac{1}{2} \hat{\mathbf{w}}_d^T \left( \left( \lambda \Phi^T \Phi \right)^{-1} + \Sigma_w \right)^{-1} \hat{\mathbf{w}}_d \right]
\]

(4.42)

The dimensionality reduction operation changes the priors on \( \lambda \) (Eqn 4.34) to

\[
a'_{\lambda} = a_{\lambda} + \frac{M(N_{\text{pix}} - m)}{2}
\]

(4.43)

\[
b'_{\lambda} = b_{\lambda} + \frac{1}{2} \Phi^T \left( I - \Phi \left( \Phi \Phi^T \right)^{-1} \Phi^T \right) \Phi
\]

(4.44)

To complete the posterior distribution for the emulator, we specify \( \Gamma \) function priors on the Gaussian process length scale kernel parameters \( \phi_{\text{int}} \). Typically, these priors are broad and peak at lengths corresponding to a few times the spacing between grid points, which helps the Gaussian process converge to the desired emulation behavior. The full posterior distribution is given by

\[
p(\lambda, \phi_{\text{int}} | \hat{\mathbf{w}}^{\text{grid}}) \propto p(\hat{\mathbf{w}}^{\text{grid}} | \lambda, \phi_{\text{int}}) p(\lambda, \phi_{\text{int}})
\]

(4.45)
where the prior is given by

\[ p(\lambda_\xi, \phi_{\text{int}}) = \Gamma(a'_{\lambda_\xi}, b'_{\lambda_\xi}) \Gamma^{m}(a_{T_{\text{eff}}}, b_{T_{\text{eff}}}) \times \Gamma^{m}(a_{\log g}, b_{\log g}) \Gamma^{m}(a_{[\text{Fe/H}]}, b_{[\text{Fe/H}]}). \] (4.46)

Now that we have fully specified a posterior probability distribution, we can sample it and find the joint posteriors for the parameters \( \lambda_\xi \) and \( \phi_{\text{int}} \) for all \( k \) simultaneously. Once we have identified the best-fit parameters for the emulator, we fix these parameters for the remainder of the spectral fitting.

Now, the emulator is fully specified and can be used to predict the values of the weights at any arbitrary set of stellar parameters \( \theta_* \) by considering them drawn from the joint distribution

\[
\begin{bmatrix}
\hat{w}_{\text{grid}} \\
w
\end{bmatrix} \sim \mathcal{N}
\begin{bmatrix}
0 \\
0
\end{bmatrix},
\begin{bmatrix}
\lambda_\xi \Phi^T \Phi & 0 \\
0 & 0
\end{bmatrix} + \Sigma_{w_{\text{grid}},w}
\]

(4.47)

where \( \Sigma_{w_{\text{grid}},w} \) is an augmented covariance matrix that includes the point \( \theta_* \). To simplify notation, we let

\[
\begin{bmatrix}
V_{11} & V_{12} \\
V_{21} & V_{22}
\end{bmatrix} = \begin{bmatrix}
\lambda_\xi \Phi^T \Phi & 0 \\
0 & 0
\end{bmatrix} + \Sigma_{w_{\text{grid}},w}
\]

(4.48)

With this notation, the \( Mm \times Mm \) matrix \( V_{11} \) is the region of the covariance matrix that describes the relations between the set of parameters in the grid, \( \{ \theta_* \}^{\text{grid}} \). The \( Mm \times m \) matrix \( V_{12} \) (and its transpose \( V_{21} \)) describe the relations between the set of parameters in the grid and the newly chosen parameters to interpolate at \( \theta_* \). The structure of \( V_{12} \) is set by evaluating \( K_i \) (Eqn 4.30) across a series of rows of \( \{ \theta_* \}^{\text{grid}} \) like in \( \Sigma_{\text{grid}} \), for \( i = 1, 2, \ldots, m \), and across \( m \) columns of \( \theta_* \). \( V_{22} \) is a \( m \times m \) diagonal matrix that represents \( K_i \) evaluated at the zero-spacing parameter pair \( (\theta_*, \theta_*) \), \( i = 1, 2, \ldots, m \). Then,
to predict a vector of weights at the new location, we use the conditional probability

$$p(w | \hat{w}^{\text{grid}}, \theta_s) = \mathcal{N}(w | \mu_w(\theta_s), \Sigma_w(\theta_s))$$

where

$$\mu_w(\theta_s) = V_{21} V^{-1}_{11} \hat{w}^{\text{grid}}$$

$$\Sigma_w(\theta_s) = V_{22} - V_{21} V^{-1}_{11} V_{12}$$

These equations are also commonly referred to as kriging equations (Cressie & Cassie 1993). Though the notation is complex, the interpretation is straightforward: the probability distribution of a set of eigenspectra weights $w$ is a $m$-dimensional Gaussian distribution whose mean and covariance are a function of $\theta_s$, conditional upon the (fixed) values of $\hat{w}^{\text{grid}}$ and the squared exponential hyperparameters (an example for a single $w_k$ is shown in Figure 4.14, right panel).

If we desired actual values of the interpolated weights, for example to reconstruct a model spectrum, we could simply draw a Gaussian random variable $w$ from the probability distribution in Eq. (4.49). However, because we now know the probability distribution of the weight as a function of $\theta_s$, we can rewrite our data likelihood function (Eq. 4.7) in such a way that it is possible to analytically marginalize over all possible values of $w$, and thus all probable spectral interpolations.

Up until this point, we have described the reconstruction of a spectrum as a linear combination of the eigenspectra that characterize the synthetic library (Figure 4.13). But in practice, that reconstructed spectrum must be further post-processed as detailed in Section 4.2.2. Fortunately, because convolution is a linear operation, we can first post-process the raw eigenspectra according to $\theta_{\text{ext}}$, and then represent the reconstructed spectrum as a linear combination of these modified eigenspectra without loss of
information. Unfortunately, the Doppler shift and resampling operations are not linear operations, and there will be some loss of information when trying to approximate them in this manner. However, we find that in practice when the synthetic spectra are oversampled relative to the instrument resolution by a reasonable factor, the flux error due to resampling is smaller than 0.2% across all pixels, and thus any effect of that information loss is negligible. For notational compactness, we let \( \tilde{\xi}_\mu, \tilde{\xi}_\sigma, \) and \( \tilde{\Xi} \) represent the post-processed eigenspectra, with an implied dependence on the current values of the extrinsic observational parameters (\( \theta_{\text{ext}} \)) and the polynomial nuisance parameters (\( \phi_P \)).

Now, the model spectrum is a function of the vector of eigenspectra weights

\[
M(w) = \tilde{\xi}_\mu + Xw
\]

(4.52)

where

\[
X = \tilde{\xi}_\sigma I_{N_{\text{pix}}} \tilde{\Xi}.
\]

(4.53)

Because the Gaussian process describes a probability distribution of the weights, we now have a distribution of possible (interpolated) models and the likelihood function (Eq. 4.6) is specified conditionally on the weights,

\[
p(D|M(w)) = p(D|w) = \mathcal{N} \left( D|\tilde{\xi}_\mu + Xw, C \right).
\]

(4.54)

The final task of designing the spectral emulator is to combine this data likelihood function with the posterior predictive distribution of the eigenspectra weights (Eq. 4.49) and then marginalize over the weights

\[
p(D|\theta_*) = \int p(D|w)p(w|\theta_*)dw
\]

(4.55)

such that we are left with a modified posterior distribution of the data that incorporates the range of probable interpolation values for the model. To perform this
multidimensional integral, we use a convenient lemma found in Gelman et al. (2013, their Appendix A): if the probability distributions of \(w\) and \(D|w\) are specified conditionally as in Eq. 4.49 and 4.54, respectively, then the marginal distribution (Eq. 4.55) is

\[
p(D|\theta_\star, \theta_{\text{obs}}, \Phi) = \mathcal{N} \left( D \bigg| \tilde{\xi}_\mu + X\mu_w, X\Sigma_wX^T + C \right),
\]

where the dependence on the model parameters is now made explicit. We can couch this modified likelihood function in the form of Eqn 4.7 by rewriting

\[
M' = \tilde{\xi}_\mu + X\mu_w
\]

\[
R' = D - M'
\]

\[
C' = X\Sigma_wX^T + C
\]

where \(M'\) can be thought of as the “mean model spectrum” given the model parameters, and the covariance matrix has been modified to account for the various probable manifestations of the model spectrum about that mean spectrum.
References


Alencar, S. H. P., Melo, C. H. F., Dullemond, C. P., et al. 2003, A&A, 409, 1037 [2.1, 2.1, 2.4.1, 2.4.2, 2.4.2]


Ambartsumian, V. 1947, Armenian Acad. of Sci [1.1]


REFERENCES

Bessell, M. S. 1979, PASP, 91, 589 [2.4.2]
Bishop, C. M. 2006, Pattern recognition and machine learning (springer) [4.2.5]
REFERENCES


Fűrész, G. 2008 [3.2.2, 4.3.1]


Fitzpatrick, E. L. 1999, PASP, 111, 63 [2.4.2, 3.4.2]

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306 [2.4.2, 3.3.1, 4.2.5]


Frigo, M., & Johnson, S. G. 2005, Proceedings of the IEEE, 93, 216, special issue on “Program Generation, Optimization, and Platform Adaptation” [2.3]


REFERENCES

Gray, D. F. 2005, The observation and analysis of stellar photospheres (Cambridge University Press) [2.4.2]

Gray, D. F. 2008, The Observation and Analysis of Stellar Photospheres [4.2.2]


Hayashi, C., Hoshi, R., & Sugimoto, D. 1962, Progress of Theoretical Physics Supplement, 22, 1 [1.1]


Herbig, G. H. 1962, Advances in Astronomy and Astrophysics, 1, 47 [1.1]


REFERENCES

Kolotilov, E. A. 1989, AZh, 66, 335 [3.4.2]
Kurucz, R. L. 1993, SYNTHE spectrum synthesis programs and line data [4.1, 4.2.1]
Mathieu, R. D., Baraffe, I., Simon, M., Stassun, K. G., & White, R. 2007, Protostars and Planets V, 411 [1.3, 2.1]
Mathieu, R. D., Stassun, K., Basri, G., et al. 1997, AJ, 113, 1841 [3.1, 3.2.2, 3.2.2, 3.3.2, 3.3.2, 3.3.2]
REFERENCES

REFERENCES


Santner, T. J., Williams, B. J., & Notz, W. I. 2013, The design and analysis of computer experiments (Springer Science & Business Media) [4.2.3]


Schönrich, R., & Bergemann, M. 2014, \textit{MNRAS}, 443, 698 [4.1, 4.2.1, 4.3.3]


Sneden, C. A. 1973, PhD thesis, The University of Texas at Austin. [4.1]


Tognelli, E., Prada Moroni, P. G., & Degl’Innocenti, S. 2011, \textit{A&A}, 533, A109 [2.4.2, 2.2, 3.4.2]


Torres, G., Andersen, J., & Giménez, A. 2010, \textit{A&A Rev.}, 18, 67 [3.3.1, 3.3, 3.4.1]


Tremaine, S., & Dong, S. 2012, \textit{AJ}, 143, 94 [2.4.1]
REFERENCES


