

The Pennsylvania State University
The Graduate School

**ON DETECTING NEW WORLDS:
THE ART OF PRECISE DOPPLER SPECTROSCOPY
USING IODINE CELLS**

A Dissertation in
The Department of Astronomy and Astrophysics
by
Sharon Xuesong Wang

© 2016 Sharon Xuesong Wang

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2016

The dissertation of Sharon Xuesong Wang was reviewed and approved* by the following:

Jason T. Wright

Associate Professor of Astronomy and Astrophysics
Dissertation Advisor, Chair of Committee

Suvrath Mahadevan

Associate Professor of Astronomy and Astrophysics

Lawrence Ramsey

Distinguished Senior Scholar and Professor of Astronomy

Eric Ford

Professor of Astronomy and Astrophysics

James Kasting

Evan Pugh Professor of Geosciences

Michael Eracleous

Professor of Astronomy and Astrophysics, Head of Graduate Program

*Signatures are on file in the Graduate School.

Abstract

The first discovery of an extra-solar planet (exoplanet) around a main-sequence star, 51 Peg *b*, discovered using Doppler spectroscopy, opened up the field of exoplanets. For more than a decade, the dominant way for finding exoplanets was using precise Doppler spectroscopy to measure the radial velocity (RV) changes of stars. Today, precise Doppler spectroscopy is still crucial for the discovery and characterization of exoplanets, and it has a great chance for finding the first rocky exoplanet in the Habitable Zone of its host star. However, such endeavor requires an exquisite precision of 10-50 cm/s while the current state of the art is 1 m/s.

This thesis set out to improve the RV precision of two precise Doppler spectrometers on two 10-meter class telescopes: HET/HRS and Keck/HIRES. Both of these spectrometers use iodine cells as their wavelength calibration sources, and their spectral data are being analyzed via forward modeling to estimate stellar RVs. Neither HET/HRS or Keck/HIRES deliver an RV precision at the photon-limited level, meaning that there are additional RV systematic errors caused by instrumental changes or errors in the data analysis. HET/HRS has an RV precision of 3-5 m/s, while Keck/HIRES has about 1-2 m/s. I have found that the leading cause behind HET/HRS's "under-performance" in comparison to Keck/HIRES is temperature changes of the iodine gas cell (and thus an inaccurate iodine reference spectrum). Another reason is the insufficient modeling of the HET/HRS instrumental profile. While Keck/HIRES does not suffer from these problems, it also has several RV systematic error sources of considerable sizes. The work in this thesis has revealed that the errors in Keck/HIRES's stellar reference spectrum add about 1 m/s to the error budget and are the major drivers behind the spurious RV signal at the period of a sidereal year and its harmonics. Telluric contamination and errors caused by the spectral fitting algorithm also contribute on the level of 20-50 cm/s. The strategies proposed and tested in this thesis will improve the RV precision of HET/HRS and Keck/HIRES, including their decade worth of archival data.

This thesis also documents my work on characterizing exoplanet orbits using RV data and the discovery of HD 37605*c*. It concludes with a summary of major findings and an outline of future plans to use future precise Doppler spectrometers to move towards the goal of 10 cm/s and detecting Earth 2.0.

Table of Contents

List of Figures	vii
List of Tables	xviii
Acknowledgments	xix
Chapter 1	
Introduction	1
1.1 On Detecting New Worlds	1
1.2 The Art of Precise Doppler Spectroscopy	4
1.3 Precise Doppler Spectroscopy with Iodine Cells as Calibrators, and an Outline for This Thesis	9
Chapter 2	
The California Planet Survey Doppler Code	13
2.1 Introduction and Background	13
2.2 Basic Formulae, Algorithm, and Components	13
2.2.1 The Reference Spectra	15
2.2.2 The Functional Forms of the Instrumental Profile	18
2.3 Code Structure and Work Flow	18
2.4 Adjusting Offsets and Computing Weights for Chunk RVs	21
Chapter 3	
Improving the Radial Velocity Precision of HET/HRS	25
3.1 Introduction and Background	25
3.2 Adoption of REDUCE and the CPS Doppler Code for HET/HRS	27
3.3 The Search for a Better Instrumental Profile	30
3.4 Investigation on the Iodine Cell and the Iodine Atlases	36
3.4.1 Why did we suspect the iodine atlas?	36
3.4.2 Ultra-High Resolution Echelle Spectra of Iodine Cells	40
3.4.3 Measuring iodine cell temperatures using synthetic Iodine spectra	44
3.4.4 Cause for a Higher Iodine Gas Temperature	49
3.4.4.1 Temperature Gradient in the Glass	51

3.4.4.2	Heated Gas within a Cooler Glass	52
3.5	Conclusions and Future Work	54
Chapter 4		
Improving the Radial Velocity Precision of Keck/HIRES		58
4.1	Introduction and Background	58
4.2	Effects of Telluric Contamination and Remedies	59
4.2.1	Introduction	59
4.2.2	Characterization of the Impacts of Micro-tellurics on RV Precision via Simulation	60
4.2.2.1	Methodology of the Simulation	61
4.2.2.2	Results of the Simulation	63
4.2.3	Remedies and Effectiveness	65
4.2.3.1	Masking is an ineffective solution	65
4.2.3.2	How precisely does one need to model the tellurics? . .	68
4.2.3.3	Treating telluric contamination in real observations . .	70
4.2.4	Summary of Recommendations on Treating Telluric Contamination	73
4.3	Errors Induced by Imperfect Stellar Reference Spectra	73
4.4	Numerical and Algorithmic Errors	77
4.5	Conclusions and Future Work	79
Chapter 5		
Characterization of Exoplanet Systems Using Radial Velocities		82
5.1	Introduction and Background	82
5.2	BOOTTRAN: Uncertainties for Orbital Parameters Estimated Using Radial Velocities	83
5.3	Other Applications to Characterize Planetary Systems	86
Chapter 6		
The Discovery of HD 37605c and a Dispositive Null Detection of Transits of HD 37605b		91
6.1	Introduction	91
6.1.1	Context	91
6.1.2	Initial Discovery and Followup	92
6.1.3	TERMS Data	92
6.1.4	Synthesis and Outline	93
6.2	Spectroscopic Observations and Analysis	93
6.2.1	HET and Keck Observations	93
6.2.2	Data Reduction and Doppler Analysis	94
6.2.3	Stellar Analysis	95
6.3	Orbital Solution	96
6.3.1	Transit Ephemeris	96
6.3.2	The 37605 System	96

6.3.3	Comparison with MCMC Results	97
6.3.4	Dynamical Analysis	99
6.3.5	Activity Cycles and Jupiter Analogs	100
6.4	The Dispositive Null Detection of Transits of HD 37605 b	101
6.4.1	APT Observations and Analysis	101
6.4.2	MOST Observations and Analysis	102
6.5	Summary and Conclusion	103
6.6	Note on Previously Published Orbital Fits	104
Chapter 7		
	Summary and Future Directions	115
Bibliography		118

List of Figures

1.1	Habitable Zone for stars with various effective temperature, highlighted in blue (and red for the extended zone). Some known planets in the Habitable Zone of their host stars are plotted, scaled by their sizes. Solar system planets are also plotted and scaled by size. The yellow and orange dashed vertical lines mark the discovery space in terms of planetary periods to which TESS is most sensitive. The white curves are equal-RV-amplitude lines, showing the semi-amplitude of the RV signals induced by any hypothetical Earth-mass planets on each curve. This plot is made by Chester Harman and Ravi Kopparapu and kindly provided by Paul Robertson.	4
1.2	(a) A portion of a HARPS-S spectral image as an example of ThAr calibrated spectrum. (b) The Lick/Hamilton iodine cell in front of its slit plate.	6
1.3	The RV semi-amplitude K of all RV-detected exoplanets vs. date of discovery. The planets are color- and size-coded according to their host star brightness. The RV precision has essentially stalled at ~ 2 m/s level since 2005, except for a very few bright stars observed with high cadence. Current major RV instruments are plotted at their start dates, with the position and height of the rectangles indicating their RV precision. Selected future RV instruments are labeled in thick lines showing their first-light dates and target precision. The black crosses are the <i>Kepler</i> small planets with mass upper limits in Marcy et al. (2014) measure by Keck/HIRES.	10
2.1	An illustration for the forward modeling process for iodine-calibrated stellar spectrum. The top black line represents the stellar reference spectrum (or the DSST), and the middle green line represents the iodine reference spectrum. Both reference spectra are convolved with an IP only for illustration purposes in order to have a clear match with the observed spectrum, plotted in black dots on the bottom. In practice, the reference spectra are multiplied <i>first</i> and then convolved with IP.	16
2.2	Illustration of the making of DSST from deconvolution, using IPs derived from neighboring B star + iodine observations.	17

2.3	Calling sequence and main functionality for core IDL routines in the CPS Doppler code.	19
3.1	Illustration of the effects of temperature stabilization on HET/HRS RV precision, using data on the HD 37605 system (see Chapter 6). The RMS of the RV residuals (bottom panel) against the best-fit two-planet Keplerian solution (solid line in top panel) has dropped from 9 m/s to 6 m/s (gray areas in the bottom plot) after implementation of fine temperature control in the spectrograph room in March 2008. The pre-2009 HET/HRS data are provided by the UT Austin group.	26
3.2	RV time series on HD 185144 as observed by HET/HRS and reduced by our data pipeline. The left panel shows RVs extracted using the Keck HD 185144 DSST, and the right panel shows RVs with DSST derived using HET/HRS data and the CPS code we adopted.	29
3.3	Illustration of convolving the iodine atlas (sharp solid lines) with a kernel (middle right insert) to fit the observed iodine lines (black dots near the bottom, with best-fit model plotted in solid line).	30
3.4	Comparison between fits for a typical iodine-only chunk using HET/HRS data (left panel) and Keck/HIRES data (right panel). Bottom panels are showing the residuals against best-fit models, plotted on the same y -axis scale. HET/HRS fit is significantly worse than Keck/HIRES, which we believe is one of the major drivers behind HET/HRS’s poorer RV precision.	31
3.5	Comparison of fits for iodine-free region/observation (fitting a straight line for each chunk) and for iodine region/chunks. The left panel is for HET/HRS data in one observation, but using spectral orders with or without iodine lines. The right panel is for Keck/HIRES data in two different observations with and without iodine cell in place. The fits for iodine-free spectral chunks turn out to be consistent what is expected with photon noise for both HET/HRS and Keck/HIRES. This eliminates additional noise as a suspect in contributing to the bad fits to HET/HRS iodine spectra.	32
3.6	Illustration of fitting the iodine-only data (bottom panel) using different IPs (top panels), in this case, using GH and sum of Gaussians (GAU). These two IPs are practically “equally bad”, having similarly large reduced χ^2 but neither produces a satisfactory fit. It is also interesting to see how “stable” the best-fit IP can be across the years (i.e., in 2005 vs. 2008) and its smoothness, hinting that the best IP may take a simple, slowly-varying form.	32

3.7 Fourier transform or power spectrum of a HET/HRS iodine-only spectrum (black dots) and its smoothed version (blue line). There is a clear signature of the HET/HRS slit at 4.3 pixel (corresponding to slit width for resolution R = 60k). For comparison, the red curve is for Keck/HIRES data, which shows no clear signature of a slit, because Keck/HIRES is not fiber-fed and the PSF of the star falls mostly within its slit.	33
3.8 Introducing a sharp feature into the HET/HRS IP model, a triangle on top of the GH IP (red curve in the left panel), produces a better fit, somewhat to our surprise. The black in the left panel is the best-fit GH IP. GH+triangle is the IP model that produces the least χ^2_ν among all of our IP models. However, as shown by the right panel, the two fits barely have any visible difference (red curve for GH+triangle IP and black for GH; bottom panel plots the residuals). Such a sharp feature in the IP is nonphysical, and we interpret this results as a hint for an unreliable iodine atlas (the sharp peak at the center is perhaps the IP model trying to “stretch” the iodine lines deeper; see Section 3.4 for more details).	34
3.9 Histogram of goodness of fit, χ^2_ν , values for spectral chunks of an iodine spectrum. The modified Moffat function (red) performs almost equally well while having only 3 parameters, compared with the complicated 11-parameter GH function (black solid). Red dashed histogram is for fits using a ThAr line profile as IP. The insert is showing the modified Moffat function can fit a ThAr line quite well.	35
3.10 Comparison of the KPNO FTS scan (black) and the NIST FTS scan (red) for the HET/HRS iodine cell for a selected 1.5Å chunk. Top: Two scans at their native resolution and original wavelength solution. Middle: Comparison of the two scans after adjusting the normalization, shifting, and convolution for the NIST scan to match the KPNO scan for a more direct comparison of line depths/ratios. Bottom: Residuals of the middle panel, NIST spectrum minus the KPNO spectrum. The median absolute deviation between the two spectra is 0.02 (2%), though at many places, especially at line centers, the two can differ by up to 5–10%.	37

3.11 Both plots are histograms of χ^2_ν values of a single iodine observation. Each χ^2_ν value in the histogram represents the χ^2_ν goodness of fit for a $\sim 2\text{\AA}$ spectral chunk in this iodine observation (each iodine observation is chopped into several hundred of chunks and is fitted independently). Left: χ^2_ν histograms for the fit of the iodine observation using the KPNO (black) and NIST (red) scan as iodine templates, respectively. The KPNO scan obviously performs better. Right: χ^2_ν histograms for the two scans, but both with the normalization as a free parameter for each chunk (as we suspect the NIST scan has problems in normalization). The two scans now perform at essentially the same level. Dashed red line is the same red histogram as plotted in the left panel. Notably, the KPNO scan also performs better when we float the normalization parameter.	38
3.12 Comparison of the median χ^2_ν values for fits of iodine observations using the HET/HRS cell KPNO scan (black solid line), the NIST scan (red dashed), and the Keck/HIRES cell KPNO scan (green dotted-dashed). Each data point represents the median χ^2_ν value for all the chunks in a single iodine observation (these are all lamp-illuminated – no B star observation). Results of 550 HET/HRS observations are plotted here to illustrate the statistically significance. The Keck/HIRES cell KPNO scan provides a better fit than the both HET/HRS scans when fitting HET/HRS iodine observations.	39
3.13 One raw image frame taken using the TS12 setting of Tull Spectrograph. It contains about 1.9\AA of iodine absorption spectrum.	42
3.14 Comparison of the Sandiford iodine cell KPNO FTS spectrum and the spectrum taken with TS12. Left: Comparison of the two spectra in their native resolutions (both about 400,000–500,000). Right: Comparison of the two spectra convolved down to about 60,000 resolution, which is the resolution of typical iodine observations or radial velocity observations (star+iodine). Bottom panel shows the residuals in percentage of the TS12 spectrum minus the KPNO spectrum, with a median absolute deviation of 0.3%.	43
3.15 The same as the right panel of Figure 3.14, the KPNO spectrum and the TS12 spectrum for the Sandiford cell both at 60,000, but for the entire $\sim 30\text{\AA}$ TS12 spectrum available.	44
3.16 Left: HET/HRS cell NIST FTS scan at three different temperatures. Right: HET/HRS cell TS12 spectra at four different temperatures for the same wavelength region, which, unlike the NIST scans, shows significant difference when the temperature of the cell changes by 10°C	45

- 3.21 Comparison between the amounts of variation in a spectral chunk in ~ 500 HET/HRS and Keck/HIRES lamp iodine observations, respectively. The top two panels show the overplotted spectra in this chunk, shifted and normalized so that the spectral lines match (using cross correlation and interpolation; normalized using total counts in this chunk). The black dots in each of the two top panels are the RMS values of the normalized counts in each pixel, multiplied by 4 and added by 0.5 for illustration purposes. The HET/HRS spectra exhibit significantly more variation than the Keck/HIRES spectra, and we discuss potential reasons in Section 3.5. The bottom panel is plotting the pixel counts RMS as a function of the derivative of the spectrum, which shows that the HET/HRS spectra have more variation in both line centers/continuum and line wings than Keck/HIRES. 55
- 4.1 Keck/HIRES RV vs. BC for HD 185144, an RV standard star. The strong correlation between velocities and the barycentric velocity of the star indicates RV systematic errors that are not intrinsic to the star. This is because the barycentric motion of the Earth is the dominant source of redshifts of stellar lines as Earth moves at a large speed of up to 30 km/s relative to the star. As the stellar lines move back and forth annually through the iodine lines, any systematic errors will manifest as correlations between RV and BC. 59
- 4.2 Telluric lines in the iodine region are mostly shallow water lines, with some moderately deep water lines near 5900Å and very deep oxygen lines near 6300Å. The insert plot is showing the pervasiveness of micro-telluric lines, i.e. $\leq 1\text{--}3\%$ in depths. 61
- 4.3 Effects of telluric lines manifested as correlation between RV and BC. Each point represents the difference in RV estimates for a pair of simulated spectra: one without telluric absorption, and one with telluric absorption on top of the stellar and iodine spectra. **Top 2 panels:** To isolate the effects of telluric lines, the simulated spectra used for this plot do not have Poisson noise added, and they have simple one-component Gaussian IPs which have fixed width and thus the IP parameters are all fixed to the true values in the RV extraction. **Middle 2 panels:** same as the top panels, but for simulated spectra with Poisson noise (same noise for the telluric and non-telluric spectrum pairs; and still the same simple IPs). **Bottom 2 panels:** same as above, but for simulated spectra with Poisson noise and complex IPs that are similar to the ones in actual observations. IP parameters are not fixed in this case, so the code is fitting 12 additional parameters for the IP on top of the 3 for wavelength solution and Doppler shift (see Chapter 2 for more details on the code). 64

4.4	Illustration for how we mask telluric contaminated pixels. The top panel shows how we mask the telluric lines (red solid lines) in the epoch observation taken at $t = t_e$. The bottom panel shows why we also need to mask pixels associated with telluric lines in the deconvolved stellar reference spectrum taken at epoch $t = t_0$ and being shifted in order to model the observation.	67
4.5	Improvements in RV RMS for different “level” of telluric modeling/removal. For example, the mid point labeled with “50%, 40 cm/s” means that if you model your telluric absorption lines to 50% of their original depths, the effects of the residual telluric absorption will add 40 cm/s in quadrature to your final RV RMS. The blue point marks the RV RMS for simulations with Poisson noise and complex IP on HD 185144, which represents the photon-limited RV precision (subject to additional numeric or algorithmic errors; see Chapter 7 for more on the limitation of the Doppler code).	69
4.6	Effects of using clean DSST and preliminary telluric modeling on RV precision and accuracy, for HD 185144 (top two rows) and HD 10700 (bottom row). There are two sets of data for HD 185144 because we ran the Doppler code with two DSSTs derived from spectra taken at two different epochs (i.e., at different BCs). The left and middle panels are showing RVs vs. BCs for RV extractions with and without telluric treatment, respectively. The right panels show the RV difference between the two panels.	71
4.7	Effect of imperfect DSST as revealed by simulated data. The top panels are for HD 185144 and the bottom panels are for HD 10700. The left panels are RVs from the simulation run using the perfect DSST or the true spectrum. The middle panels are RVs from the simulation run using the simulated DSST, which has errors from the manufacture process of DSST such as the deconvolution. The right panels are RV differences between the left and the middle ones, illustrating the effects of DSST errors on RV precision and accuracy, because the only difference between these two runs is the DSST.	75

4.8	Effect of imperfect DSST on simulated data for a single spectral chunk. The top panels are for a chunk near 5160Å for HD 185144, and the bottom panels are for a chunk around 5166Å for HD 10700. The left panels illustrate the differences between the simulated DSST (solid red) and the true spectrum (dashed black), i.e., the errors in DSST. The right panels show the derived RVs for this chunk as a function of BC: the RVs on top, with y -axis labeled as “Observed”, are from real Keck/HIRES data, and RVs below, labeled with “Simulated”, are from simulations using the simulated DSST (black dots) and the “perfect” simulation using the true spectrum as the DSST (gray dots, with no apparent RV-BC trends). The 5100–5200Å spectral region tends to receive high weights due to its high density of stellar and iodine lines and a lack of telluric lines.	76
4.9	Lomb-Scargle periodograms for HD 185144 RVs using real (black line) or simulated (red dashed line) Keck/HIRES data. The spurious peaks near 180 and 360 days are caused by systematic errors which correlates with BCs. In the case of the simulated data, we know that this systematic arises from the errors in the DSST because all other conditions are kept “perfect”. For the real data, it is a combination of DSST errors and other systematics (such as, potentially, telluric contamination).	77
4.10	RV vs. BC for HD 185144, for simulations with fixed simple IPs and no photon noise added. The data plotted here are from simulated spectral data with different IP widths (or spectral resolution, $\sigma = 1.7$ pixels corresponds to original Keck/HIRES resolution). The origins of the RV scatters and trends in these plots are purely algorithmic.	78
4.11	Real (top) or simulated (bottom) Keck RVs from a 2Å chunk from HD 185144 spectra showing RV systematic errors caused by algorithmic errors. The RVs in the bottom panel are from the simulations with complex IPs and photon noise.	79
5.1	Best-fit $100 \times 100 \chi^2$ map for fixed values of P_c and $M_c \sin i_c$ for HD 217107c. This confirms that the period and mass are well-constrained. We have illustrated the contours of the 1σ , 2σ , and 3σ (defined by $\chi^2 = \chi^2_{\min} + \{2.30, 6.17, 11.8\}$) confidence levels, based on for the number of degrees of freedom in the problem (Press et al. 2002). The center and 1σ limits in both parameters are consistent with the bootstrapping uncertainties for these parameters. This figure is published as Figure 9 in Feng et al. (2015) and was co-produced by the author of this thesis and the leading author of the paper, Y. Katherina Feng.	88

6.2	χ^2_ν map for the best Keplerian fits with fixed values of period P and minimum planet mass $M \sin i$ for HD 37605c. This is showing that both P and $M \sin i$ are well-constrained for this planet. The levels of the contours mark the 1σ (68.27%), 2σ (95.45%) and 3σ (99.73%) confidence intervals for the 2-D χ^2 distribution.	107
6.3	Comparison between the Bayesian (MCMC) analysis and RVLIN+BOOTTRAN results. Top four and bottom left: Contours of the posterior distributions of selected orbital parameters (P , e , K , $M \sin i$, and ω) based on the MCMC analysis (dashed dotted line). The x -axes are orbital parameters of the inner planet, b , and the y -axes are those of the outer planet, c . The inner contours mark the 68.27% (' 1σ ') 2-D confidence regions and the outer ones are 95.45% (' 2σ ') ones. Also plotted are the best Keplerian fit from RVLIN (blue squares) and $\pm 1\sigma$ error bars estimated via bootstrapping (blue bars). Bottom right: Marginalized posterior distribution of time of conjunction (mid-transit) T_c of HD 37605b in dashed dotted line. The solid grey vertical line is the median of the distribution, and the dashed grey vertical lines mark 1σ confidence interval. The solid blue vertical line is the best estimate of T_c from RVLIN+BOOTTRAN, with $\pm 1\sigma$ error bars plotted in blue dashed vertical lines. See § 6.3.3 for details.	108
6.4	Dynamic evolution of the best-fit MCMC system. On the left we plot the short-term evolution over 10 years, on the right we plot the evolution over 10^7 years (< 1/10 of our dynamic simulation time scale). The top plots describe the evolution of the semi-major axes and eccentricities of the inner planet (a_b & e_b , blue lines) and the outer planet (a_c & e_c , red lines), while the bottom plot describes the parameter space covered by the $e \cos \omega$, $e \sin \omega$ quantities over 10^8 years (blue for inner planet and red for outer planet). We find that over the short-term (e.g., our RV observation window of ~ 10 years), the parameter variations are negligible, but in the long term significant eccentricity oscillations can take place (particularly noticeable in the eccentricity of the outer planet). See § 6.3.4 for details.	109

6.5 Photometric observations of HD 37605 acquired over five years with the T12 0.8m APT. The top panel shows the entire five-year data set; the dotted line represents the mean brightness of the first observing season. A long-term brightening trend is evident with a total range in the seasonal means of 0.002 mag. The middle panel shows the photometric data normalized so that each season has the same mean as the first and then phased to the orbital period of HD 37605 b (55.01307 day). The solid line is the predicted transit light curve, with Phase 0.0 being the predicted time of mid-transit, T_c . A least-squares sine fit of the phased data produces the very small semi-amplitude of 0.00031 ± 0.00011 mag, providing strong evidence that the observed radial-velocity variations are not produced by rotational modulation of surface activity on the star. The bottom panel plots the observations near T_c at an expanded scale on the abscissa. The horizontal bar below the transit window represents the $\pm 1\sigma$ uncertainty in T_c . Unfortunately, none of the APT observations fall within the predicted transit window, so we are unable to rule out transits with the APT observations. See § 6.4.1 for more.	110
6.6 Brightness variability in HD 37605 possibly induced by stellar rotation at $P = 57.67 \pm 0.30$ days. Top panel is the periodogram of the complete, normalized data set. Bottom panel shows the normalized photometry folded with this possible rotation period. The peak-to-peak amplitude is 0.00120 ± 0.00021 mag. See § 6.4.1 for more.	111
6.7 Photometric observations of HD 37605 by the MOST satellite, which rule out the edge-on transit of HD 37605 b at a $\gg 10\sigma$ level. The solid line is the predicted transit light curve, and the dashed vertical lines are the 1σ transit window boundaries defined by adding σ_{T_c} (0.069 day) on both sides of the predicted transit window (0.352-day wide). See § 6.4.2 for more details.	112
7.1 Initial result from the new Doppler pipeline: 1-D and 2-D marginalized posterior distribution for 5 parameters (out of 17, including two for Gaussian processes) in a fit for a 2Å spectral chunk.	116

List of Tables

1.1	Current Precise Doppler Spectrometers	7
1.2	Future Precise Doppler Spectrometers	9
1.3	List of Frequently Used Acronyms	12
2.1	Parameters for Forward Modeling Keck/HIRES RV Spectra	14
3.1	Hardware Settings for TS12 Iodine Spectrum Test	41
4.1	List of Simulations	63
4.2	Comparison of RV RMS between Simulations with or without Masking	66
4.3	Summary of Keck/HIRES RV Errors Found in This Work	80
6.1	STELLAR PARAMETERS	105
6.2	KEPLERIAN FIT PARAMETERS	105
6.3	COMPARISON WITH MCMC RESULTS	112
6.4	PHOTOMETRIC OBSERVATIONS OF HD 37605 FROM THE T12 0.8m APT	113
6.5	PHOTOMETRIC OBSERVATIONS OF HD 37605 ON MOST	113
6.6	Updated $M \sin i$ and Errors for HD 114762 b and HD 168443 b, c	114
6.7	HET and KECK RADIAL VELOCITIES FOR HD 37605	114

Acknowledgments

This thesis would not have been possible without the help from many, especially my advisor, Prof. Jason Wright. His tireless mentoring and encouragement provided strong support for the later years of my graduate study. His intellectual brilliance, academic integrity, and genuine altruism as an advisor and mentor has made him my role model for the academic path.

I thank my families for their support in many ways, especially my husband, Dr. Enshi Xu, for his generous support and whine-free devotion to the housework. I thank my daughter for truly understanding the importance of my work to me (especially for her willingness to get out of the house quickly in the mornings even when she did not want to but had to go to daycare).

I thank all my committee members, Profs. Larry Ramsey, Jim Kasting, Suvrath Mahadevan, and Eric Ford for guidance and helpful suggestions. Their brain storming ideas on job hunting had put me in contact with the PFS team at the Carnegie Institute for Science and eventually landed me my dream job as a Carnegie Fellow at DTM after graduation.

I thank all the Penn State Astronomy faculty, staff, and students for providing a warm home for me for the past eight years. I truly think we are the best department on Earth. In particular, I would like to thank Profs. Peter Mészáros and Niel Brandt for taking me in on their projects, providing me financial support, and teaching me many valuable skills as an astronomer. I would like to thank Profs. Robin Ciadullo, Steinn Sigurdsson, and Don Schneider for their mentoring and moral support. I thank my fellow graduate students, especially Megan Comins, Prakash Arumugasamy, and Jason Young, for moral support.

Many collaborators have contributed to the work in this thesis. I am in debt of Dr. John A. Johnson for providing his copy of the California Planet Survey (CPS) Doppler code and for his generous advice and help on Doppler reduction in general. I thank Dr. Debra Fischer for her help with Doppler reduction in general, and for inviting me to host the breakout session on telluric contamination at the EPRV II workshop in Yale in July 2015. I thank Dr. Chad Bender for his generous help with TERRASPEC and on modeling the telluric lines. I am grateful for Dr. Bill Cochran and the UT Austin team for providing archival HD 37605 data and for their collaboration on the discovery of HD 37605c. I am very grateful for CPS for the same reason, and also for providing many

other data, such as Keck spectra on HD 185144, which are very crucial for my work. I thank the HET staff for helping us on the engineering run for characterizing modal noise in HRS, and I am extremely thankful for the staff at McDonald Observatory, especially David Doss and Coyne Gibson, for their help with the TS12 runs. Many thanks to Drs. Gillian Nave and Stephen Redman for performing the NIST FTS scan for the HET/HRS iodine cell. Last but not least, I would like to thank Dr. Ming Zhao for his work on TS12, as well as his generous help and advice throughout our overlapping years at Penn State and even beyond.

I am very grateful for the funding provided by NSF (grant AST-1211441), NASA NESSF, CEHW, and NAI at Penn State for the work in this thesis.

Chapter 1

Introduction

“There are infinite worlds both like and unlike this world of ours. For the atoms being infinite in number, as was already proven, ...there nowhere exists an obstacle to the infinite number of worlds.”

Epicurus (~341-270 B.C.)

“How vast those Orbs must be, and how inconsiderable this Earth, the Theatre upon which all our mighty Designs, all our Navigations, and all our Wars are transacted, is when compared to them.”

Christiaan Huygens, *Cosmotheoros* (1698)

“Of all of the topics of study in astronomy, exoplanets hold a special place in the imagination. More than stars, nebulae, or galaxies, they are *places*, ...”

Wright (2006)

1.1 On Detecting New Worlds

Even before human beings realized that other stars are like our Sun, the existence of other worlds have been speculated by ancient Greek philosophers such as Epicurus and Democritus. In the blooming age of astronomy in the 1400s and 1600s, early pioneers such as Giordano Bruno and Christiaan Huygens have also pondered upon the existence of planets around other stars (extra-solar planets, or exoplanets).¹ In modern times, 40 years before the discovery of the first exoplanet, Otto Struve stated that exoplanets, especially “super-Jupiters” on short orbits, should be detectable via spectroscopy and photometry (Struve 1952).

¹In fact, Huygens conducted the first documented search on exoplanets. For more on the history of exoplanet searches, see these three websites:

The NASA PlanetQuest, <http://www.nasa.gov/externalflash/PQTimeline/>;
Search for Exoplanets, <http://www.hao.ucar.edu/research/stare/search.html>;
ESO, <https://www.eso.org/public/outreach/eduoff/cas/cas2004/casreports-2004/rep-228/>.

Unfortunately, the earliest claims of exoplanet detections before 1980s all turned out to be erroneous (Jacob 1855; van de Kamp 1969). These and a later retracted claim of a planet around a pulsar by Bailes et al. (1991) made all astronomers extremely cautious about exoplanet detection claims. In 1988, Campbell, Walker and Yang announced potential planetary signal from the star Gamma Cephei, but they were hesitant in calling it a detection due to limitations of early instruments. Their detection method was to measure the radial velocity (RV) variation of the star using precise Doppler spectroscopy, which is described in the next section and is also the theme of this thesis. It was not until 2003 that the planet around γ Cephei A was confirmed (Hatzes et al. 2003), which means the work of Campbell et al. (1988) is the first real exoplanet detection though only in retrospect. The detection of HD 114762b (“Latham Planet”; Latham et al. 1989) also belongs to the family of first exoplanet detections, though the planet was thought to be a brown dwarf at the time due to its large mass. A similar story to γ Cephei Ab is the discovery of β Gemini b (Hatzes & Cochran 1993), where the existence of the planet was not confirmed until 2006 (Hatzes et al. 2006) because of the strong activity-induced RV signals of the giant host star.²

The more commonly recognized first detection of exoplanets belongs to Wolszczan & Frail (1992), who reported two planets around the pulsar PSR B1257+12, detected via the pulsar timing method using radio data (later on it turned out this system hosts one more planet). It was a surprising detection in many aspects, and these planets remains the only known planetary system around a pulsar to date (as of May 2016). If exoplanets could exist around exotic stars like pulsars, then it is only natural to expect them to exist around more “regular” main-sequence stars like our Sun.

Finally, in 1995, a team in Geneva announced the first definitive detection of a planet around a main-sequence star, 51 Peg b (Mayor & Queloz 1995). Their results were quickly confirmed by other planet hunters such as Geoffrey W. Marcy and R. Paul Butler, who quickly caught up with the game (Butler & Marcy 1996) and went on to detect more than half of the hundreds of known exoplanets up until the launch of NASA’s *Kepler* mission Borucki et al. (2010). The method adopted by Mayor & Queloz (1995) and Butler & Marcy (1996) was again precise Doppler spectroscopy. Today, there are over 585 exoplanets discovered by precise Doppler spectroscopy. The discoveries by precise Doppler spectroscopy that happened beyond this point are briefly accounted for in the next section.

Several years later, Henry et al. (2000) and Charbonneau et al. (2000) detected the first exoplanet transiting event, where the planet moves in between the disk of the star and our line of sight periodically, leaving signals in the stellar light curves. This transiting planet, HD 209458b, was discovered via precise Doppler spectroscopy first. Nonetheless, this discovery opened up the age of transit detections, where projects such as OGLE, TRES, WASP, XO, HAT, and CoRoT etc. added more than 200 new exoplanet discoveries to date (Konacki et al. 2003; Alonso et al. 2004; Christian et al. 2006; McCullough et al. 2006; Bakos et al. 2007).

In 2009, the discovery of exoplanets entered a new era with the launch of NASA’s *Kepler* satellite, which is a dedicated space mission to detect transiting exoplanets. This

²See Chapter 4 of Wright & Gaudi (2013) for a more detailed history on these early detections.

extremely fruitful mission has made new exoplanet discoveries in the counts of thousands (Rowe et al. 2014; Morton et al. 2016), with over 2000 more planet candidates (see, e.g., NASA Exoplanet Archive for statistics on exoplanet discoveries). The science of exoplanets expanded from the philatelic style to including population and statistical studies which inform planet formation and evolution in powerful ways more than ever (e.g., Fressin et al. 2013 on occurrence rate and Wolfgang et al. 2015 on composition distribution). As of May 23 2016, there are 3268 confirmed exoplanets, to which the transit method contributed 2569 (585 discovered by Doppler spectroscopy).

Besides using precise Doppler spectroscopy and transits, other methods have also made unique and important discoveries of exoplanets, as they probe different stellar population and are subject to different observational biases. Bond et al. (2004) made the first micro-lensing detection of exoplanet, where planets act as additional gravitational lenses beside their host star and leave characteristic signatures in the light curves of the background star. There are 37 exoplanets discovered via micro-lensing so far. Astronomers also directly detected light from young exoplanets around young stars via direct imaging, the first of which are Fomalhaut b (Kalas et al. 2008) and the four planets around HR 8799 (Marois et al. 2008). Today, there are 41 directly imaged planets.

More exoplanets around more diverse host stars are expected to be discovered in the near future, with many new missions and surveys being carried out, built, or planned. Post-2013, *Kepler* continued as the K2 mission (*Kepler* on two reaction wheels) and kept churning out planets (e.g., Vanderburg et al. 2016b). The Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2014; expected to launch in late Summer 2017) will survey the whole sky, targeting nearby and bright stars, including the previously relatively unexplored population of M dwarf stars. With TESS, exoplanet discoveries will be made in counts of thousands once again. Ongoing surveys with the Gemini Planet Imager on Gemini South (Macintosh et al. 2014) and the SPHERE instrument on the Very Large Telescope (Beuzit et al. 2008) are populating exoplanets in a new parameter space (young stellar/planetary age and moderate to long orbital distances). The future for micro-lensing discoveries also remains bright as thousands of exoplanets are expected to be found by NASA's WFIRST-AFTA mission (Yee et al. 2014).

Among all these exciting discoveries happening or on the horizon, precise Doppler spectroscopy continues to play an important role. It is the most important method for measuring planetary masses,³ and it will remain a crucial independent method for discovering new exoplanets (after all, only a small fraction of exoplanets happen to pass in between their host star and the Earth). The synergy between the *Kepler* mission and the ground-based Doppler spectroscopy follow-ups has demonstrated the power of this new exoplanet discovery and characterization scheme, where RVs are presented as the convincing evidence for the planetary nature of the transit signal, and they also provide valuable information on the planetary masses and thus their bulk densities (e.g.,

³Planetary masses can also be measured via studies on the transit timing variations (TTVs) due to the dynamic interactions of multiple planets (e.g., Jontof-Hutter et al. 2016), but TTVs are only measurable for a small fraction of all transiting planets (Mazeh et al. 2013). de Wit & Seager (2013) have also developed an innovative method to estimate planetary mass via transmission spectroscopy of the planetary atmosphere, but the method is model-dependent and requires a large amount of large-aperture space telescope time (e.g., hundreds of orbits of JWST).

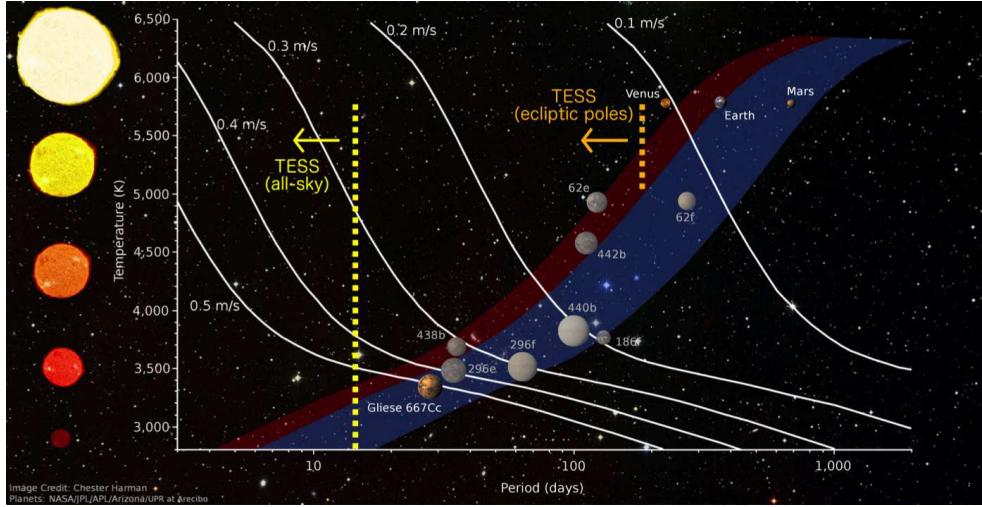


Figure 1.1 Habitable Zone for stars with various effective temperature, highlighted in blue (and red for the extended zone). Some known planets in the Habitable Zone of their host stars are plotted, scaled by their sizes. Solar system planets are also plotted and scaled by size. The yellow and orange dashed vertical lines mark the discovery space in terms of planetary periods to which TESS is most sensitive. The white curves are equal-RV-amplitude lines, showing the semi-amplitude of the RV signals induced by any hypothetical Earth-mass planets on each curve. This plot is made by Chester Harman and Ravi Kopparapu and kindly provided by Paul Robertson.

Marcy et al. 2014). Such measurements are crucial for mapping out the demographics of exoplanets. However, only a small fraction of *Kepler* targets have been followed up by Doppler spectroscopy, and the future discoveries of TESS will put an even higher demand on RV follow up (see, e.g., a summary in Plavchan et al. 2015).

Doppler spectroscopy also remains the most promising avenue for detecting Earth-like planet in the Habitable Zone (Kasting et al. 1993; Kopparapu et al. 2013) in the near future. Figure 1.1 illustrates the Habitable Zones for different types of stars and the discovery space that TESS will access, which does not include the Habitable Zone around Sun-like to early-M stars due to TESS’s short lifespan. The next generation Doppler spectroscopy, with an RV precision of <0.5 m/s, bears great hope for detecting rocky or even Earth-like planets in the Habitable Zone. Can we fulfill such a great expectation? The next section focuses on the art of precise Doppler spectroscopy, on how we achieved the current precision of 1 m/s today, and on how the field will carry on and aim for an RV precision of ≤ 10 cm/s in the coming decade.

1.2 The Art of Precise Doppler Spectroscopy

Although it has been common practice to measure the radial velocities (RVs) of stars for over a century, the idea of measuring them *precisely*, to the order of 10s m/s and below, was first proposed in 1973 by Griffin & Griffin (1973), where they proposed to

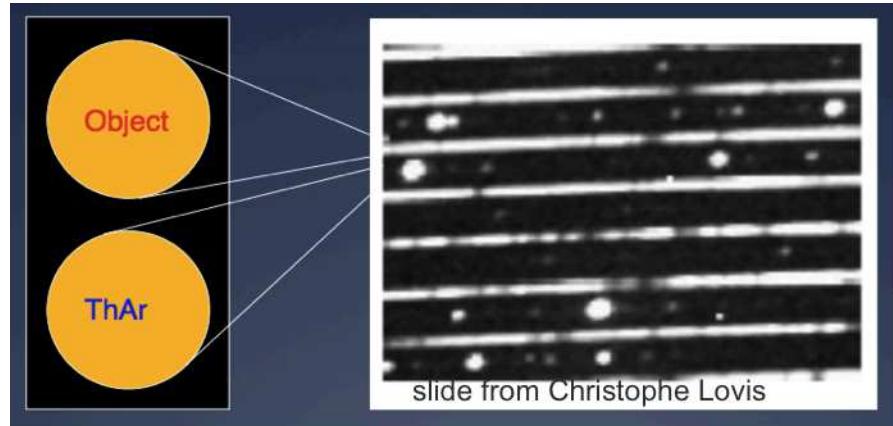
use atmospheric oxygen lines as wavelength calibration and estimated that a precision of 10 m/s should be attainable. The source of wavelength calibration and how the calibration is done are two of the key aspects in measuring precise RVs. From very early on,⁴ there have been two types of ideas, which also heavily shaped the landscapes of today’s precise Doppler spectrometers. One idea uses simultaneous wavelength calibration, where the spectral lines of the calibrator are blended with the stellar lines, such as the oxygen lines in the atmosphere or absorption lines of a gas cell in the light path. The other one is similar to the traditional approach of measuring RVs, where a spectrum with known wavelength solution (e.g., from a ThAr lamp) is taken before and/or after or simultaneously with the observed frame (but not interwoven with the stellar lines), and to obtain more precise RVs, the spectrometer is further stabilized to minimize changes in wavelength solution on the image plane.

The pioneering work by Campbell et al. (1988) calibrated their wavelengths using an absorption gas cell filled with hydrogen fluoride (HF), which provides well-known and clean and evenly (though sparsely) spaced absorption lines in the red optical band. They were able to achieve a stunning precision of 10-15 m/s back in the late 1980s. The discovery of HD 114762b by Latham et al. (1989) at Harvard/CfA used the parallel calibration approach, where they used solar spectrum at dusk and dawn as calibrators. They later stabilized their spectrometer by employing strategies such as mechanical temperature stabilization for the spectrometer and fiber feeds and improved the precision to about 100 m/s.

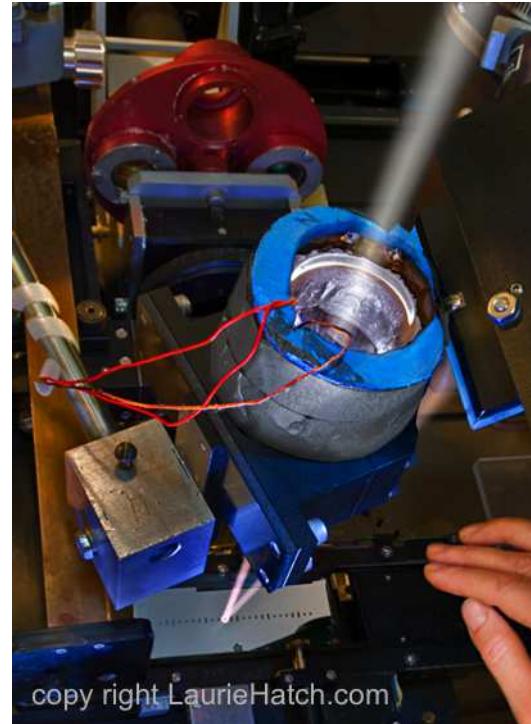
The spectrometer Mayor and Queloz used to discover 51 Peg *b* was ELODIE (Baranne et al. 1996), which achieved 15 m/s thanks to its excellent mechanical stability. The Geneva team lead by Mayor and the European community continued along the path with parallel calibrators (with ThAr lamps in 4000-6900Å) and stabilized spectrometers. More and more exoplanet discoveries are made with two successors of ELODIE: first CORALIE (1999), which achieved 2-7 m/s precision using simultaneous ThAr calibration in the spectral image (Figure 1.2), and later on SOPHIE (2006; Perruchot et al. 2008) and HARPS-S (2004; Mayor et al. 2003), with its stunning stability of 1 m/s enabled by its ultra-stabilized vacuum enclosure. The European team later went on to build HARPS-N (2012; essentially a copy of HARPS-S, with an RV precision of 1 m/s) in collaboration with the Harvard/CfA group including David Latham (Cosentino et al. 2012).

Meanwhile across the ocean, in the early 1990s, Geoffrey Marcy and Paul Butler adopted the absorption gas cell approach and began their survey on nearby solar type stars. Their choice of gas cell was the iodine cell (Marcy & Butler 1992; Figure 1.2), which had dense and sharp absorption lines across the green part of the optical spectrum (5000-6200Å), where it is also rich of stellar lines for solar type stars. They first began with the Hamilton spectrograph at Lick Observatory, which achieved 3-7 m/s (Fischer et al. 2014). Later on, they achieved 1-3 m/s with HIRES on Keck I (1996; then in 2004, a CCD upgrade brought the precision to 1-2 m/s). Chapter 4 is on our efforts in improving the RV precision of Keck/HIRES. The surveys done by Cochran and Hatzes also used iodine calibrators, first with the coudé spectrograph 2.1 m telescope at McDonald Observatory

⁴A more detailed recount on the early history of precise Doppler spectroscopy can be found in Section 4.3 in Wright & Gaudi (2013).



(a) ThAr-Calibrated HARPS Spectrum



(b) Iodine Cell on Lick/Hamilton

Figure 1.2 (a) A portion of a HARPS-S spectral image as an example of ThAr calibrated spectrum. (b) The Lick/Hamilton iodine cell in front of its slit plate.

(20 m/s using oxygen lines and later 10-20 m/s with an iodine cell), then later on with the Tull spectrograph (1998; 5 m/s) on the McDonald 2.7 m and HRS on the Hobby-Eberly Telescope (HET) at the same observatory (3 m/s), whose “under-performance” in comparison to Keck/HIRES is the focus of Chapter 3 of this thesis.

The efforts and instruments described above all operate in the optical band. The RV precision in the near infrared (NIR) is about an order of magnitude lower. Blake et al. (2010), Figueira et al. (2010), and Bailey et al. (2012) demonstrated 10-50 m/s precision in the NIR using telluric lines as wavelength calibrators. The “state of the art” for NIR precise RV in 2016 is the 5 m/s precision achieved by Bean et al. (2010) using CRIRES and a methane gas cell as the calibrator. The two bottlenecks are hardware stability and telluric contamination. Fortunately, with the cryogenic NIR spectrometer CARMENES (whose commissioning has began in November 2015; it also has an arm in the optical band; Quirrenbach et al. 2014) and HPF (scheduled to begin in 2016; Mahadevan et al. 2012), the situation in the NIR will change quickly in the very near future.

As of 2016, the precision of Keck/HIRES and HARPS represents the state of the art of the field for each calibration method. These two methods of wavelength calibration have contributed almost equally to the ensemble of exoplanets discovered using precise Doppler spectroscopy. There are, of course, many other precise Doppler spectrometers that are running today, and Fischer et al. (2016) summarize their performance in terms of RV precision in details. Table 1.1 summarizes these spectrometers and their reported precision mostly as in Fischer et al. (2016), which provides an overview of the landscape of precise Doppler spectroscopy in 2016.

It took about a decade to breach the 10 m/s precision barrier (Butler et al. 1996a), and optimistically, one would expect the 1 m/s barrier to be breached a few years ago around 2010, extrapolating from the previous progress. Unfortunately, this was not the case. Figure 1.3 shows the evolution of RV precision and how we have basically stalled around the 1-2 m/s level for the past two decades. Except for a couple of extremely bright stars that were observed heavily with HARPS-S, there has not been detections of planets regularly made with a precision below 1 m/s. The reasons are complicated and many.

Stars are known to be “uncooperative” in terms of staying RV quiet. (1) The choice of stars is important, because magnetically active stars such as giants and fast rotators are intrinsically tough for precise RV surveys. There is also the concern of stellar brightness: because the signal-to-noise ratio of the spectrum directly affects the RV precision, this makes large telescopes such as Keck and HET better for large RV surveys. (2) Stellar activity induced RV signals on the level of 1-2 m/s are very hard to model. The physical mechanisms behind such stellar jitter, such as macro turbulence, are still relatively poorly understood, which makes stellar activity arguably the most difficult piece in the puzzle of getting down to 10 cm/s. We have some handle on spots or plages induced RV signals (Dumusque et al. 2014), and it is encouraging to see successful cases such as CoRoT 7b (Haywood et al. 2014), where the stellar RV signal is neatly modeled with the help of photometric data and advanced statistical tools (see more in Vanderburg et al. 2016a).

Hardware stability builds the foundation for high RV precision. (3) Changes in temperature and pressure will induce mechanical changes of the spectrometer and also change the index of refraction of the air, and both will translate into wavelength solution

Table 1.1. Current Precise Doppler Spectrometers

Facility	RV Precision ^a
Iodine Cell Calibrated	
CHIRON ^b	1.0
PFS ^b	1.2
HIRES ^b	1.5
APF/Levy ^b	1.5-2.0
SONG ^b	2.0
UCLES ^b	3.0
Tull	5.0
HRS ^{b,c}	3.0
Hamilton ^{b,c}	3.0
ThAr Lamp Calibrated	
HARPS-S	0.8
HARPS-N	0.8
TOU	1.0
PARAS	1.0
SOPHIE+	1.1

^d

^aEstimated single measurement precision, which represents the photon-limited precision of the instrument at an SNR of 200.

^bPlanet search programs on these instruments all use some version of the CPS Doppler code for RV extraction (or once have used).

^cThe old HRS is taken down in 2013. The Lick/Hamilton program ended in 2011 due to an incidence that damaged the iodine cell.

^dAs reported in Ma et al. (2016).

drifts on the CCD and thus RV noise. The successful story of HARPS has shown that it is best to have an ultra stabilized spectrometer sealed in a temperature and pressure controlled vacuum vessel. (4) Moreover, the light going into the spectrometer needs to be stabilized, in the sense that the fiber optics which feeds the spectrometer need to be evenly illuminated and stay stable (e.g., Halverson et al. 2015). (5) Another extremely important piece is the fidelity and stability of the wavelength calibrator, and the next generation precise Doppler spectrometers are all making the switch from using ThAr lamps to laser frequency combs (Li et al. 2008; Yi et al. 2016). Chapter 3.4 shows the importance of the stability and a full knowledge of the iodine cell to RV precision.

Elimination of telluric contamination and robust data analysis tools are crucial “software” aspects that ensure the delivery of high RV precision out of the raw spectral data. (6) As mentioned above, telluric contamination is a bottleneck for NIR precise RV, and it has not been considered a problem for the optical RV community until recently because it contributes to the RV error budget by about 10-50 cm/s, well below the current precision (see more in Chapter 4.2). (7) Robust data analysis tools are also crucial because they are the last avenue for battling instrumental and stellar RV systematic errors. For the ThAr calibrated spectra, precise RVs are measured via cross correlation between a “mask” with varying weights at the positions of stellar lines and the observed stellar spectrum that are flattened and wavelength-calibrated (Baranne et al. 1996; Pepe et al. 2002b). For iodine calibrated spectra, precise RVs are extracted through the forward modeling method, which is described in detail in Chapter 2 and is also the theme for Chapter 4.3 and 4.4. Interpreting the RVs and translating them into planetary orbits is also challenging and crucial (Chapter 5), especially with the existence of stellar activity induced signals (e.g., the coming and going of α Cen Bb; Dumusque et al. 2012; Rajpaul et al. 2016).

Here I have only briefly discussed each item, with the goal to present a broad picture of how precise Doppler spectroscopy works and provide a general background for this thesis. For a more detailed description and summary of the status and the community’s plans on breaching the 10 cm/s barrier, see Plavchan et al. (2015) and Fischer et al. (2016). A list of future precise Doppler spectrometers is presented in Table 1.2, and most of them have a target RV precision of below 50 m/s or even 10 cm/s.

1.3 Precise Doppler Spectroscopy with Iodine Cells as Calibrators, and an Outline for This Thesis

Precise Doppler spectroscopy with iodine cells as calibrators has had a glorious past full of exciting discoveries, such as the first stars with multiple planets (Butler et al. 1999), the first Earth-density planet Howard et al. (2013); Pepe et al. (2013), and also the characterization of first sample of sub-Neptune and super-Earth planets (Marcy et al. 2014) which enabled the first studies on the demographics of small exoplanets (Wu & Lithwick 2013; Weiss & Marcy 2014; Rogers 2015; Wolfgang & Lopez 2015; Wolfgang et al. 2015). However, all planned future precise Doppler spectrometers are going to be ultra-stabilized and calibrated by laser combs with <1 m/s precision. How would iodine-calibrated Doppler spectrometers such as HET/HRS and Keck/HIRES fit into

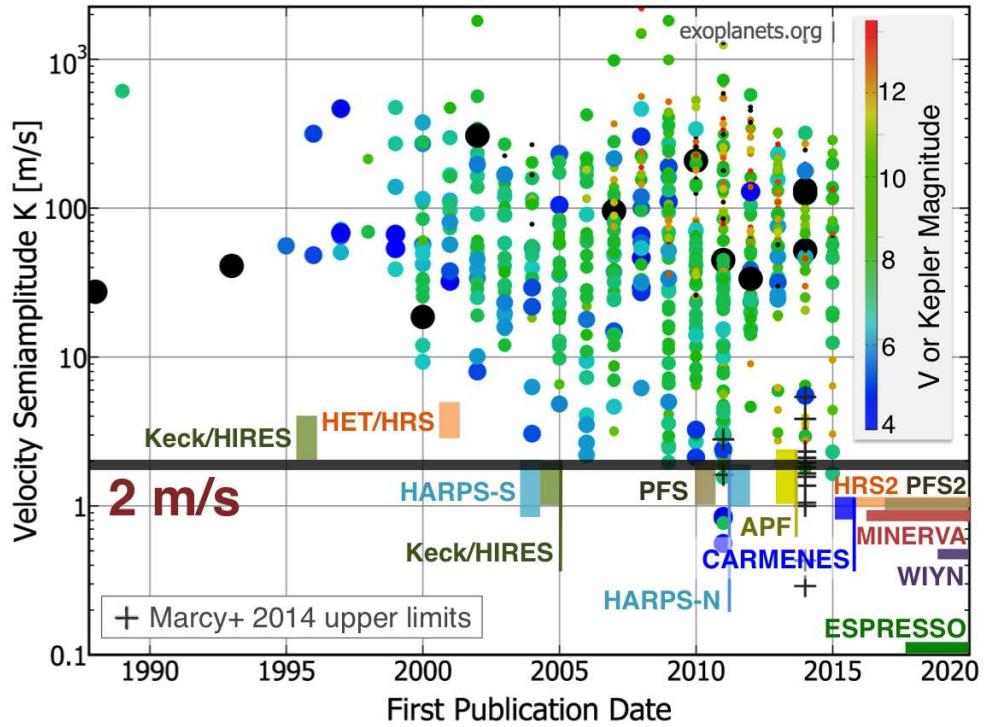


Figure 1.3 The RV semi-amplitude K of all RV-detected exoplanets vs. date of discovery. The planets are color- and size-coded according to their host star brightness. The RV precision has essentially stalled at ~ 2 m/s level since 2005, except for a very few bright stars observed with high cadence. Current major RV instruments are plotted at their start dates, with the position and height of the rectangles indicating their RV precision. Selected future RV instruments are labeled in thick lines showing their first-light dates and target precision. The black crosses are the *Kepler* small planets with mass upper limits in Marcy et al. (2014) measure by Keck/HIRES.

Table 1.2. Future Precise Doppler Spectrometers

Optical	NIR
Commissioning in 2016	
CARMENES on CAO 3.5 m (duo bands)	
MINERVA at Mt Hopkins	MINERVA Red
HET/HRS2, Magellan/PFS2	HET/HPF
Being Built	
Magellan/MAROON-X	CFHT/SPIRou
VLT/ESPRESSO	VLT/CRIRES+
LCOGT/NRES	Subaru/IRD
AAT/Veloce	
LBT/iLocator	
WIYN/NEID	
DCT/EXPRES	
Concept and Planning	
EELT/HIRES (duo bands)	EELT/SIMPLE
TNG/GIARPS (duo bands)	Keck/NIRSPEC ^a
Keck/SHREK	TMT/NIRES ^a TMT/HROS ^a
GMT/G-CLEF	
GTC/HORUS	

^aProposed but unfunded.

the current and future pictures of precise RV?

First, the iodine-calibrated method is a relatively cheap way of measuring RVs precisely at the 1-2 m/s level, because it does not require pairing with a ultra-stabilized spectrometer like HARPS. Even with the existence of Doppler spectrometers with ~ 10 cm/s precision, instruments with 1 m/s precision will still have plenty of science to do such as RV follow up of transiting planets and asteroseismic studies (e.g., the SONG telescopes, Grundahl et al. 2011).

Second, iodine-calibrated instruments will stay competitive in the next 3-5 years, especially during TESS era, because most 10-meter class telescopes will only have iodine-calibrated precise Doppler spectrometer (the only exceptions are HPF on HET in the NIR and ESPRESSO on VLT). There are several next-generation spectrometers being built or planned for large telescopes such as iLocator, MAROON-X, and SHREK (Table 1.2), but it is unlikely that they would be scientifically productive during most of the TESS mission era. Large apertures are highly desirable for *Kepler* and TESS follow up. For example, HARPS-N can only target stars brighter than the 10th *Kepler* magnitude, while Keck/HIRES can follow up an order of magnitude more (median brightness for *Kepler* stars is about 13-14 mag). As a result, majority of the *Kepler* RV follow up is done by Keck/HIRES. Even for TESS, which primarily target “bright” stars, the median *V* magnitude for Sun-like stars is actually below 11 (Sullivan et al. 2015). It is of no doubt that HET/HRS and Keck/HIRES will play an important role in TESS follow up.

Improving the precision of HET/HRS and Keck/HIRES will enable more efficient

Table 1.3. List of Frequently Used Acronyms

Acronym	Meaning
RV	radial velocity
BC	barycentric velocity correction
RMS	root mean square
SNR	signal to noise ratio
CCD	charge-coupled device
IP	instrumental profile (or spectrograph response function)
DSST	deconvolved stellar spectral template
IDL	Interactive Data Language, a programming language
GH	Gauss-Hermite polynomials
FTS	Fourier Transform Spectrometer
KPNO	Kitt Peak National Observatory
NIST	National Institute of Standards and Technology
HET	Hobby-Eberly Telescope
HRS	High Resolution Spectrograph (on HET)
HIRE	High Resolution Echelle Spectrometer (on Keck)
CPS	California Planet Survey

follow up on TESS targets and also benefit the independent long-term RV surveys being carried out at HET and Keck. This is the goal of this thesis. Improving the precision of the fiber-fed HET/HRS also helps us to gain a better understanding on modeling fiber-fed instruments and to prepare for future ones such as MINERVA (Swift et al. 2015). The work in this thesis will also improve the RV precision on archival HET/HRS and Keck/HIRES data, which span over a decade – the longest baseline in the history of precise Doppler spectroscopy.

The scope of this thesis is mostly on RV data analysis, including improving the data analysis tools and diagnosing hardware problems through data. Chapter 2 contains the documentation on the software package I used for getting RVs out of HET/HRS and Keck/HIRES data, and it also serves as an introduction on extracting RVs from iodine-calibrated stellar spectra. Chapter 3 and 4 document my efforts in improving the RV precision of HET/HRS and Keck/HIRES, respectively. The introduction sections in these two chapters contain more information on these two instruments and a general background on my work. Chapter 5 and 6 are from peer-reviewed, published papers. Chapter 5 documents my work in characterizing exoplanet orbits using RV data, and Chapter 6 reports the discovery of HD 37605c, which is the first exoplanet discovered by the Wright group. Chapter 7 contains a brief summary of the findings in this thesis and a description of future directions beyond the ones mentioned in its preceding chapters, where the readers will see that the tool set built for this thesis and the lessons learned on how to model the stellar spectra precisely are valuable not only for iodine-calibrated data, but also for general future work on detecting new worlds using precise Doppler spectroscopy.

Chapter 2

The California Planet Survey Doppler Code

2.1 Introduction and Background

For all of the work in this thesis, we used the California Planet Survey (CPS) Doppler code to extract precise RVs from the iodine-calibrated HET/HRS and Keck/HIRES spectra. The main authors for this code are Paul Butler, Geoffrey Marcy, and John Johnson, with contributions from Jeff Valenti and so on. This chapter contains a brief documentation describing the algorithm and structure of the CPS Doppler code, which is used extensively for all the work in this thesis. As of March 2016, no documentation in published or unpublished form existed for this code, although Butler et al. (1996a) describes the basics for the technique of iodine-calibrated precise RV, and some CPS publications contain description for certain elements of the code (e.g., Johnson et al. 2006; Howard et al. 2009, 2011; Johnson et al. 2011a).

The earliest date mark in the code is 1991, which is roughly when Paul Butler and Geoffrey Marcy started drafting the code. The code was heavily modified by John Johnson around 2002-2008. Later on, two versions of the code were maintained by two separate groups: the California Planet Survey (CPS) team (John Johnson) and the Lick-Carnegie Planet Survey (LCPS) team (Paul Butler). Post 2014, the CPS version of the code was maintained mostly by Howard Isaacson at UC Berkeley. This code is widely used by many iodine-calibrated precise RV instruments, such as Lick/Hamilton, HET/HRS (this thesis; Chapter 3), AAT/UCLES, Magellan/PFS, and the Automated Planet Finder (APF) at Lick (both LCPS and CPS have their own version of this code for APF). Our copy of the code was kindly provided by John Johnson in 2009, and the copy was checked to still be consistent with the CPS version in 2013.

2.2 Basic Formulae, Algorithm, and Components

First, we describe the basic mathematics and algorithm behind RV extraction from iodine-calibrated stellar spectra using the CPS code. The overall algorithm is to forward model the stellar spectra using synthetic or empirically derived reference spectra, finding

Table 2.1. Parameters for Forward Modeling Keck/HIRES RV Spectra

Parameter	Unit and Meaning
z	no unit, the stellar redshift
w_0	Å, wavelength of the first pixel of a spectral chunk
w_d	Å/pixel, wavelength dispersion scale for a spectral chunk
$A_n, n = 1, \dots, 12$	no unit, amplitudes of side gaussians for IP ^a

^aSee Section 2.2.2 for more information.

best-fit model parameters including the Doppler shift, z .

The reference spectra include¹: a model spectrum for the iodine absorption lines, $F_{\text{I}_2}(\lambda)$ and a model spectrum for the star, $F_{\star}(\lambda)$. The goal is to use the model the observed, extracted, and normalized 1-D spectrum, $F_{\text{obs}}(x)$, at any given pixel position (and spectral order), x , using these reference spectra and model parameters. The broadening effect of the spectrograph is described by the spectral response function, or the spectral point spread function, or the instrumental profile (IP), which we will refer to as the IP throughout this thesis and is denoted as $\mathcal{P}(x)$. Hence,

$$F_{\text{obs}}(x) = [F_{\text{I}_2}(\lambda(x)) \times F'_{\star}(\lambda(x))] * \mathcal{P}(x), \quad (2.1)$$

where $\lambda(x)$ is the wavelength solution for the 1-D spectrum, and F'_{\star} is the redshifted stellar spectrum defined by $F'_{\star}(\lambda) = F_{\star}(\lambda \cdot (1 + z))$. The Doppler shift z contains two components: the stellar RV v_{\star} and the barycentric (BC) velocity of the telescope with respect to the star v_{BC} . The BC component is corrected by $v_{\star} = v_{\text{measured}} + v_{\text{BC}} + z \cdot v_{\text{BC}}$ (the last term being relativistic correction; Wright & Eastman 2014).

The stellar reference spectrum is empirically derived from iodine-free stellar observations taken on an epoch, say, T_0 . As a result, all measured RVs for the star using a stellar template from T_0 represent relative stellar velocities between epoch T_0 and epoch T_{obs} (i.e., $v_{\star,T_{\text{obs}}} - v_{\star,T_0}$), instead of the absolutely RVs of the star. The following subsection describes the origins of the stellar (and also the iodine) reference spectra.

In practice, for Keck/HIRES, for example, each 1-D spectrum taken at an epoch is divided into ~ 700 spectral chunks, each with 80 pixels and about 2Å in wavelength. One model is created and fitted for each spectral chunk, with the model parameters listed in Table 2.1. Model parameter optimization is done through least- χ^2 fitting using the Levenberg-Marquardt (LM) algorithm. Errors on the extracted 1-D spectrum are assumed to be Poisson noise plus a 2% addition representing potential errors in the raw reduction. Initial guesses of the parameters come from the solution for nearest B star + iodine observation, with the exception of the initial guess for z , which is set to be v_{BC} because that is usually on the order of km/s and dominates the Doppler shift signal.

¹It can also include a model spectrum for a faint secondary star, telluric absorption lines (see Chapter 4 Section 4.2), and so on.

To sample the reference spectra into the observed pixel grid, the code first re-samples (using cubic spline interpolation) each of them onto a grid finer than the observation pixel grid by a factor of four (i.e., using a wavelength dispersion of $w_d/4$). The wavelengths of this fine grid are provided by the proposed wavelength solution parameters w_0 and w_d . Next, it shifts the stellar reference spectrum according to the proposed Doppler shift parameter z . Then it multiplies the shifted stellar reference spectrum with the iodine reference spectrum, and then convolves the product spectrum by the IP. Then it re-bins the finely sampled model spectrum onto the observed pixel grid, which yields the final normalized model spectrum $F_{\text{model,norm}}(x)$. It is worth noting that the code only resamples the reference spectra and never modifies the reduced observed spectrum in any way at any time.

In reality, the observed spectrum used in the fitting is not normalized (i.e. no blaze or continuum removal). To account for blaze and stellar continuum, the code divides the observed spectrum by the normalized model spectrum, i.e., $F_{\text{obs}}(x)/F_{\text{model,norm}}(x)$, and then it fits a straight line $S(x)$ through the divided spectrum (for a small 2Å chunk, this linear approximation seems sufficient). It then computes a new model spectrum by adding this model “continuum” on top, i.e., $F_{\text{model,final}}(x) = F_{\text{model,norm}}(x) \times S(x)$. This way, the continuum component in the observed spectral chunk is modeled by a linear function but imposes no explicit parameters for the model.

2.2.1 The Reference Spectra

Ideally, the reference spectra are the “ground truth” spectra, i.e. the intrinsic spectra of the sources (e.g., the iodine cell, or the star) without Doppler shift or being broadened by the spectrometer. In reality, there is no way of knowing such “ground truth”, so the reference spectra are empirically derived from observations.

The iodine reference spectrum, often referred to as the iodine atlas, originates from a Fourier Transform Spectrometer scan of the iodine cell illuminated by a continuum source. It is often of very high signal-to-noise ratio (SNR) with high resolution (normally $\sim 500,000$ or larger). Therefore, it is generally regarded as basically the “ground truth” for the cell, especially for the purpose of forward modeling lower-resolution ($\sim 60,000$) spectra. However, there can be problems with the iodine atlas, for various reasons. See Chapter 3 Section 3.4 for more on this topic. The current FTS iodine atlas being used for Keck/HIRES RV work is from a scan in 1993, using the Babar FTS at NSO/KPNO, and so is the atlas for HET/HRS. See Section 3.4 for more on iodine reference spectra.

The stellar reference spectrum for any star, or internally to CPS referred to as the Deconvolved Stellar Spectral Template (DSST), is empirically derived from observed spectra of the target star. For most of the CPS targets (bright stars), a few (4-5) observations of the star with a narrower slit ($R \geq 80,000$) are taken without the iodine cell in the light path consecutively. Then the extracted 1-D stellar spectra are shifted and stacked together to boost the SNR (> 500), and then deconvolved with proper IPs derived from two groups of B star + iodine observations taken immediately before and after the iodine-free stellar observations. The wavelength solution for the DSST also comes from the bracketing B star + iodine observations. For faint stars where obtaining stellar template is expensive or unfeasible, Johnson et al. (2006) developed a technique

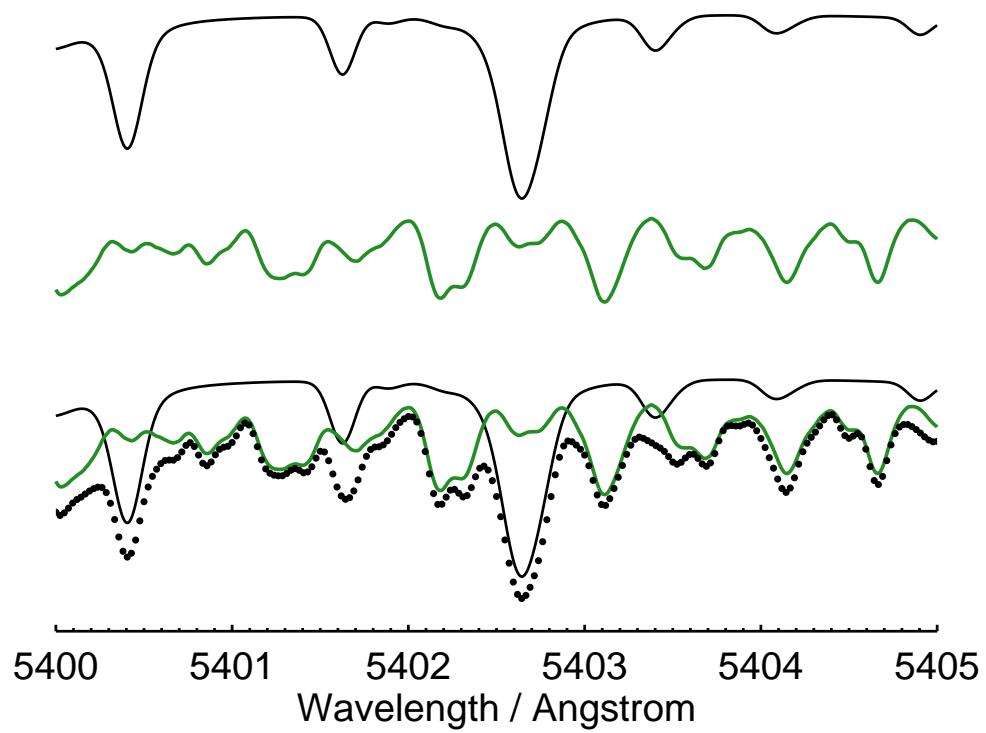


Figure 2.1 An illustration for the forward modeling process for iodine-calibrated stellar spectrum. The top black line represents the stellar reference spectrum (or the DSST), and the middle green line represents the iodine reference spectrum. Both reference spectra are convolved with an IP only for illustration purposes in order to have a clear match with the observed spectrum, plotted in black dots on the bottom. In practice, the reference spectra are multiplied *first* and then convolved with IP.

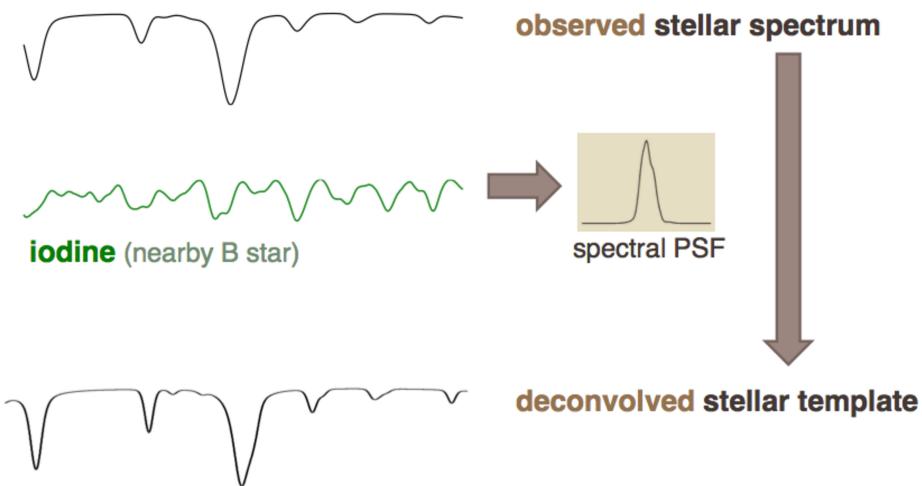


Figure 2.2 Illustration of the making of DSST from deconvolution, using IPs derived from neighboring B star + iodine observations.

where they “morph” a synthetic stellar spectrum or an existing DSST of another star with similar stellar properties to fit the stellar iodine observation, and then they use this new morphed DSST for RV extraction.

Here is an outline of how DSSTs for high SNR stars are made (also illustrated in Figure 2.2) with highlights on potential sources of errors:

1. A few high SNR, high resolution (using a narrower slit than the one for standard star+iodine frames) observations on the target stars are taken without the iodine cells. These are the bases for the DSST. The 1-D stellar spectra extracted from these frames are stacked together to raise the SNR, and the stacking process could have errors from interpolation and rebinning. Each chunk of the stacked 1-D stellar spectrum is also normalized before the final deconvolution, and normalization errors can sneak in.
2. Bracketing these stellar observations, both before and after, a couple of frames on nearby B star are taken through the iodine cell. These are the frames which anchor the wavelength solution and IPs for the stellar frames. The IPs derived from these frames are later used for deconvolving the stellar frames to make DSST. The IPs in the B star + iodine frames might be different from the ones in the stellar frames because of, for example, changes in the telescope’s PSF, changes in the spectrograph IPs due to the addition of the iodine cell in the light path, and so on. This could introduce errors in the DSST. The wavelength solution of these two types of frames can also differ for similar reasons, and although the wavelength zero point for each spectral chunk have a large tolerance for errors, any errors in wavelength dispersion can translate into RV errors.
3. Using the averaged IPs derived from the B star + iodine frames, the stacked stellar spectrum is deconvolved. Its wavelength solution is determined by the averaged wavelength solution of the B star frames. Deconvolution is a ill-posed problem and

does not have unique solutions, and therefore errors in DSST can arise from the deconvolution algorithm employed.

See Section 4.3 for more on the problems related to DSSTs.

2.2.2 The Functional Forms of the Instrumental Profile

The IP $\mathcal{P}(x)$ can take many functional forms, and for Keck/HIRES, an IP of sum of gaussians works exceptionally well (Valenti et al. 1995; $\chi^2 \sim 1$ for pure iodine absorption line fit). The mathematical form for it is:

$$\mathcal{P}_{\text{gaus}}(x) = \sum A_n \exp \left[\left(\frac{x - \mu_n}{\sigma_n} \right)^2 \right]. \quad (2.2)$$

A_n stands for the amplitude for each gaussian component. A_n 's are floated parameters for the fitter to optimize while μ_n and σ_n (i.e., positions and widths of the gaussians) have empirically-optimized fixed values, depending on the instrument setting of Keck/HIRES (e.g., slit width). For Keck/HIRES precise-RV mode (B5 decker, $\sim 60,000$ resolution, with iodine cell in light path), the IP contains 12 free parameters, A_1, A_2, \dots, A_{12} , while A_0 is fixed to 1 (the big central gaussian) and $\mu_n, n = 0, \dots, 12$ and $\sigma_n, n = 0, \dots, 12$ also have fixed values.

Another frequently used IP is the Gauss-Hermite (GH) function, which is composed of gaussians multiplied by Hermite polynomials H_n :

$$\mathcal{P}_{\text{GH}}(x) = \sum A_n u_n(x) = \sum A_n \left(\frac{2}{\pi w^2} \right)^{1/4} \frac{1}{\sqrt{n!2^n}} H_n \left(\frac{\sqrt{2}x}{w} \right) \exp \left[- \left(\frac{x}{w} \right)^2 \right]. \quad (2.3)$$

Mathematically, any sum of gaussians can be decomposed into orthogonal GH terms², and therefore, in principle, the GH IP should present a generic and flexible option for IP choices. However, in reality, the least- χ^2 solver is extremely sensitive to the choices of initial guesses, even for orthogonal bases. As a result, GH IP normally does not outperform sum of gaussians (e.g., see work by Vanderburg et al. 2013). The GH IP is what we use for extracting RVs from HET/HRS data. See Chapter 3 Section 3.3 for more.

2.3 Code Structure and Work Flow

This section documents the structure of the CPS Doppler code, with the main goal to help any reader who wishes to adopt this code for their own work.³ Figure 2.3 illustrates the code's calling sequence.

The most top-level routine for running Doppler reduction is the IDL procedure `dop_driver.pro`, which takes in the name of the star and then automatically locates

²See <http://math.stackexchange.com/questions/28719/how-to-decompose-displaced-hermite-gaussian-function-into-higher-order-hgs> for an illustration, retrieved on March 18 2016.

³Although the code is not publicly available, the readers can request a copy of the code to one of the main authors/maintainers (e.g., John Johnson, Andrew Howard, Paul Butler, and so on).

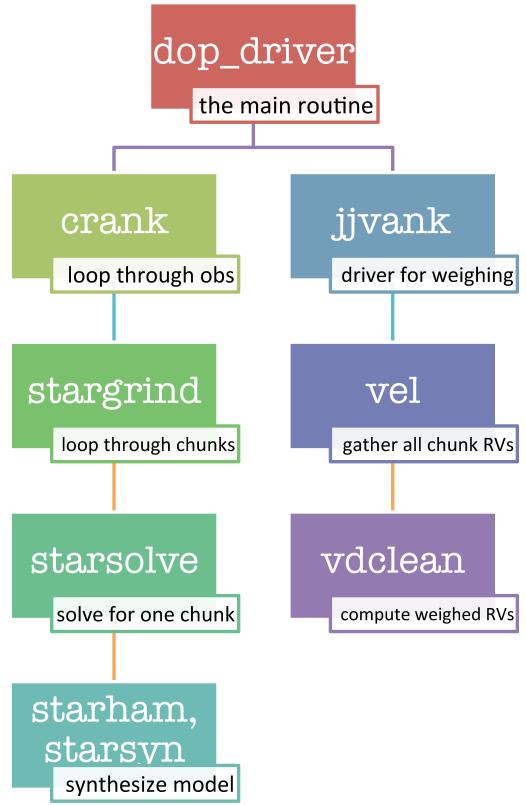


Figure 2.3 Calling sequence and main functionality for core IDL routines in the CPS Doppler code.

input files such as the extracted 1-D spectra, the proper DSST file, the iodine atlas, and the files storing initial guesses for parameters (which are called the `vdiod` files). The code is also very flexible with inputs and the user can specify things like a specific DSST file, choice for a specific IP model, explicit initial guesses for parameters, and so on. It drives the Doppler analysis and outputs a `vd` file for each observed spectrum, which is an IDL save file that contains best-fit parameters and other information for all the spectral chunks. At the end, `dop_driver.pro` calls `jjvank.pro`, which combines the RVs from all the chunks in all the observations for this star, and evaluates numerical weights for each chunk and each observation (see Section 2.4 for more details on weighting). Finally, `jjvank.pro` computes the weighted RV for each observation and its RV uncertainty and outputs the information in `vst` files. The most useful variable is the `cf3` structure in the `vst` file, which contains, for example, the Julian Date (JD) of the observation `cf3.jd`, the BC `cf3.bc`, the weighted RV `cf3.mnvel`, and its uncertainty `cf3.errvel`. The algorithm for `jjvank.pro`, or “vanking”, is described in the next section. In the end, `dop_driver.pro` will produce N_{obs} `vd` files and one `vst` file for each star. If a new observation is taken, a new `vd` file is created, and vanking re-evaluates the chunk and observation weights and outputs a new `vst` file.

While `dop_driver.pro` is the routine to call for Doppler analysis, most of its actual codes simply deal with logistical work such as locating the right files. The real driver behind the scenes is `crank.pro`, which contains the loop through all observations and “fills out” the `vd` files. In a standard CPS reduction routine, `crank.pro` is called three times. In the case of Keck/HIRES standard RV reduction, for example, the first time `crank.pro` is called, it fits for all 15 free parameters (z , w_0 , w_d and 12 IP parameters; see Table 2.1) for each chunk in each observation. The LM fitter takes in the initial guesses for IP and wavelength solution parameters from the nearest B star + iodine solutions and optimizes the 15-parameter model. Due to the complexity of this multi-modal and multi-dimensional problem, very often the best fit for this “first pass” does not yield the global minimum on the χ^2 surface. Therefore a second pass of `crank.pro` is called, with fixed IP parameters and thus only three free parameters. The IP parameters are not simply fixed to last round’s best-fit values, but instead, they were fixed to averaged IP parameters over a region on the CCD chip (for example, using the mean values of IP parameters in fits for the closest eight neighboring chunks and also the central chunk itself). After the second pass, a third pass of `crank.pro` is called with fixed wavelength dispersion w_d and only floating z and w_0 , in search for a deeper minimum or lower χ^2 . There is no evidence that this procedure finds the global minimum, but it certainly suffice for delivering RVs at a precision of 2 m/s (or \sim 20-50 m/s for a single spectral chunk).

Within `crank.pro`, the code calls `stargrind.pro` in a loop of all observations. `stargrind.pro` drives the model fitting for a single observation, where it loops through all spectral chunks, calling `starsolve.pro`, which contains the LM fitter, for each chunk. Inside `starsolve.pro`, the spectral model for each LM fitter iteration is computed using `starham.pro`, which is really just a wrapper around `starsyn.pro`. The core algorithm for model construction is all in `starsyn.pro`, which takes in the observed 1-D spectral chunk, the DSST, the iodine atlas, and the model parameters, and outputs a model spectrum for this chunk, using formulae described in the previous section.

2.4 Adjusting Offsets and Computing Weights for Chunk RVs

Now all observations and all chunks have their best-fit model parameters computed by `crank.pro`, and all RVs are barycentric corrected. What’s next?

Two extremely important things need to happen at this point before we can have an RV time series in hand, and they are accomplished by `jjvank.pro` and most importantly, `vel.pro` and `vdclean.pro` (see Figure 2.3 for the calling sequence).

First, the chunk RVs need to be adjusted so that they have the same “zero point”. Ideally, measured RVs from all chunks in one observation are good estimates for the true RV of this epoch, so the mean of all chunk RVs provides an unbiased estimate for the true RV and their scatter provides a sense for the RV uncertainty. However, in reality, some effects may cause the chunk RV to be biased and have a constant offset from the true RV. One of the leading culprits is the error in the wavelength solution of a DSST. As mentioned in the previous section, the measured RV for each chunk at a certain epoch T_{obs} is a relative RV against the DSST taken at epoch T_0 , $v_{\star,T_{\text{obs}}} - v_{\star,T_0}$. This is because the stellar lines and their Doppler shift in each chunk are modeled by redshifting the DSST: $F'_{\star}(\lambda) = F_{\star}(\lambda \cdot (1 + z))$. The wavelength solution for the DSST implies its absolute RV v_{\star,T_0} at T_0 . If the wavelength solution for all DSST chunks corresponds to the exact same value of v_{\star,T_0} , then measured RVs for all chunks in the observed stellar iodine observation would have the same “zero point”. Consequently, any biases or relative errors in the DSST wavelength solution would result in a shifted zero point for that chunk with respect to other chunks. For example, one source of error comes from the fact that the wavelength solution for DSST is derived from neighboring B star + iodine observations, which assumes that the wavelength solution for the orders and pixels remain the same between the B star and the DSST observations. Or, the wavelength solution derived from B star observations may be imperfect. There are also other reasons why the RV zero point of a chunk deviates from the other chunks. Things like persisting CCD effect (such as CCD stitching) and certain errors in iodine atlas or DSST can cause biases in RV estimates that contain a constant shift component. In the end, the inconsistent RV zero points will translate into RV scatter as we take the average of all chunk RVs to estimate the RV for one observation. Thus, it is very important to determine and correct for the chunk RV zero point offsets.

Second, not all spectral chunks are equal in terms of the quality of their reported RVs, meaning that we need to take a weighted average. There are several reasons why one chunk would consistently have a larger RV scatter than the other. The most obvious one is difference in the amount of Doppler information contained in the chunks. Some chunks have more and/or deeper stellar/iodine lines, which make them more powerful in accessing the RV information of the star. Some chunks land on the peak of the echelle blaze function, which constantly give them more SNR over the chunks down near the bottom. Some chunks may contain stellar lines that have more sensitive response to stellar activity, which would manifest as RV scatter or “jitter”. This is not an exhaustive list, but the important thing is that there could be many reasons which we know or do not know, and even for some of the things we do know, we could have no way of estimate how much extra RV scatter it would introduce to the chunk. Therefore, it is most sensible

to derive the chunk weights empirically, rather than using any a priori ones.

Overall, the algorithm for vanking works like this:

- It rejects chunks (and observations) with poor performances indicated by, for example, high χ_{ν}^2 values.
- It brings all chunks to having the same RV zero point.
- It computes the weighted average RV for each observation based on the chunk RVs, where the weight for each chunk in each observation is evaluated by how well this chunk behaves in long term (by comparing it with other chunks across all observations). Then the weight is also scaled by how good each observation is, which is evaluated by comparing the behavior of a typical chunk in one observation with the other observations.

To be specific, most of the heavy lifting is done in `vel.pro` and `vdclean.pro`. First `vel.pro` calls `vdcube.pro`, which combines all `vd` structures stored in the `vd` files for all the good observations. A good observation is defined by: (1) median photon counts for all chunks is within a user-defined range (e.g., above SNR 50 and well below CCD saturation); and (2) median χ_{ν}^2 values of all chunks is lower than the user-defined threshold (e.g., 10 for Keck/HIRES default). Otherwise `vdcube.pro` will throw out the bad observation and print out warning messages. Second, the `vd` cube put together by `vdcube.pro` gets passed on to `vdclean.pro`, and `vdclean.pro` performs a series tasks including quality checks, outlier rejections, RV zero point offset adjustment, and finally the computation of chunk weights, i.e., it

1. Throws out chunks with bad DSST or containing no Doppler information (meaning it has a weight of 0 as calculated following the method described in Butler et al. 1996a).
2. Rejects the chunks where the fitter always fails to converge (indicated by setting χ_{ν}^2 to 0 or 100 in the code).
3. Rejects the bottom 1% (or other user-defined threshold) of the chunks which have the highest photon-limited RV errors (calculated following Butler et al. 1996a).
4. Rejects the bottom 1% (or other user-defined threshold) of the chunks which have the highest χ_{ν}^2 values.
5. Computes the RV zero point offsets for all chunks and adjust all chunks to have the same zero points. Mathematically, this is done for each chunk by subtracting the offset velocity, which is estimated by the mean velocity of each chunk in all observations, i.e.:

$$\text{offset for chunk } i = \sum_{j=1}^N v_{i,j}/N \quad (2.4)$$

where $v_{i,j}$ means the reported RV for a chunk with index i in observation j , and there are N observations and M chunks in total, so $v_{i,j}$ is a $M \times N$ matrix. This is

basically requiring that the mean RV reported by any chunk over all observations is set to zero (or any arbitrary value, since we only care about the RV variation of the star instead of its absolute velocity).

6. Computes chunk weights as the inverse of the estimated RV variance of each chunk i in each observation j , i.e., $w_{i,j} = 1/\sigma_{i,j}^2$, and

$$\sigma_{i,j} = r_j \cdot \sigma_i, \quad (2.5)$$

where σ_i is defined for each chunk and r_j is defined for each observation as:

$$\sigma_i = std\{\Delta_{i,j}, j = 1, \dots, N\} \quad (2.6)$$

$$r_j = median\{|\Delta_{j,i}| \cdot \sigma_i, i = 1, \dots, M\} \quad (2.7)$$

where *std* stands for standard deviation, and $\Delta_{i,j}$ is the matrix of velocity differences between $v_{i,j}$ and the median velocity of all chunks in each corresponding observation j :

$$\Delta_{i,j} = v_{i,j} - \tilde{v}_j \quad (2.8)$$

$$\tilde{v}_j = median\{v_{i,j}, i = 1, \dots, M\}, \quad (2.9)$$

7. Rejects top 1% of the chunks which have the highest σ_i .

Then `vdclean.pro` passes back the $v_{i,j}$ and the $w_{i,j}$ matrices to `vel.pro`, which rejects the top 1% chunks that have the largest $v_{i,j} - \tilde{v}_j$ for each observation. Then `vel.pro` computes the weighted mean RV for each observation,

$$v_j = \sum_{i=1}^M (v_{i,j} \cdot w_{i,j}) / \sum_{i=1}^M w_{i,j}. \quad (2.10)$$

It is worth noting that because $w_{i,j} = \sigma_j \cdot r_j$, this is equivalent to

$$v_j = \sum_{i=1}^M (v_{i,j} \cdot \sigma_i) / \sum_{i=1}^M \sigma_i \quad (2.11)$$

However, r_j enters the picture when calculating the corresponding RV uncertainty,

$$\sigma_{v_j} = \left(\sqrt{\sum_{i=1}^M w_{i,j}} \right)^{-1}. \quad (2.12)$$

v_j nad σ_{v_j} are then passed back to `jjvank.pro` and stored in the `cf3` structure in the `vst` file for the star. This formula assumes Gaussian errors for the RVs, which is probably false and thus it probably underestimates the uncertainty.

The algorithm of vanking does not have the most rigorous statistical justification. It was more or less conjured up intuitively and tweaked until it worked, and it works

extremely well: the RV scatter is decreased by typically a factor of two or more before and after vanking. It also efficiently eliminates problematic chunks such as the ones with telluric contamination, and as a result, it mitigates the adverse effects caused by the contamination (Section 4.2). Vanking certainly has room for improvement, just like the rest of the code, which is a topic that will be touched on in Chapter 7.

Chapter 3

Improving the Radial Velocity Precision of HET/HRS

3.1 Introduction and Background

The Hobby-Eberly Telescope (HET) is the world's fourth largest telescope (as of 2016) with a diameter of 9.2 meter (11 meter by 9.8 meter). Its primary mirror is tilted at a fixed angle of 55° , which made the telescope extremely cost efficient. The telescope is located at Mount Fowlkes in West Texas and operated by the McDonald Observatory. It is a joint project of the University of Texas at Austin, the Pennsylvania State University, Ludwig Maximilians Universität München, and Georg-August Universität Göttingen.¹

Penn State has about 20% share for the time on HET as of 2016. The telescope is queue scheduled (Shetrone et al. 2007), which makes it highly efficient in pursuing observations in time series such as RV observation on exoplanet systems. Its precise Doppler capability is delivered by the High Resolution Spectrograph (HRS; Tull 1998), which is an echelle spectrograph with an iodine cell as the wavelength calibrator. The typical set up for precise RV purposes has a resolution of $R = 60,000$ (sometimes 120,000). The spectral image is captured by 2 CCDs, covering a spectral range of 4200-11000Å. It is a fiber fed spectrograph, with choices of science and calibration fibers and slits with various widths placed at the fiber exit to provide higher resolutions.

The first planet discovered by HET/HRS is HD 37605b (Cochran et al. 2004), and since then it has contributed to several detections of exoplanet systems (e.g., Cochran et al. 2007; Wittenmyer et al. 2009; Niedzielski et al. 2016) and performed *Kepler* follow ups (e.g., Endl et al. 2014). It has an RV precision of 3-5 m/s (Baluev 2009), although in principle, it should have performed at the same level of Keck/HIRES at 1-2 m/s and perhaps even better because of the additional stability provided by the fiber feed in comparison to the slit-fed Keck/HIRES. What is behind the high RV systematic errors of HET/HRS?

Temperature stability in the spectrograph room of HET/HRS was identified early on as one of the contributing factors to the RV systematic errors, and this issue was resolved since the installation of a fine temperature control system in March 2008 (J. Bean,

¹Stanford University left the HET consortium in 2011.

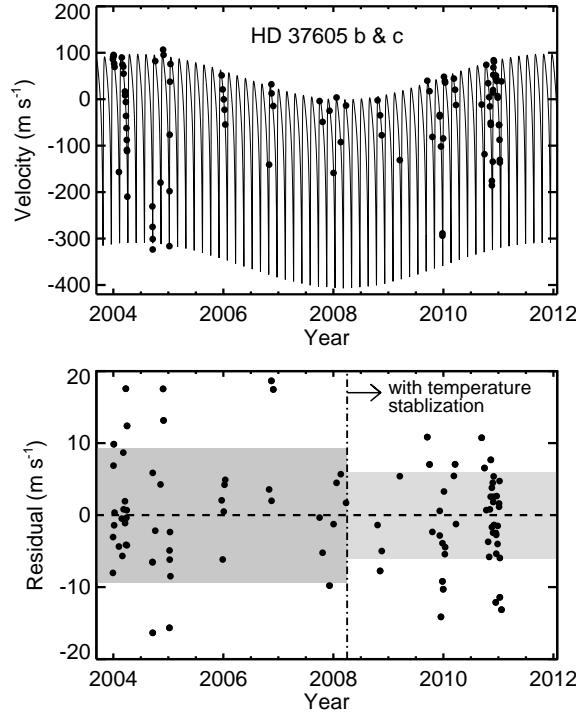


Figure 3.1 Illustration of the effects of temperature stabilization on HET/HRS RV precision, using data on the HD 37605 system (see Chapter 6). The RMS of the RV residuals (bottom panel) against the best-fit two-planet Keplerian solution (solid line in top panel) has dropped from 9 m/s to 6 m/s (gray areas in the bottom plot) after implementation of fine temperature control in the spectrograph room in March 2008. The pre-2009 HET/HRS data are provided by the UT Austin group.

L. Ramsey, P. McQueen private communications). The temperature fluctuation is now controlled down to $\ll 1$ K from 10 K. We confirmed the improvement in RV precision as a result of this upgrade in our analysis with the HD 37605 data, which is illustrated in the lower panel of Figure 3.1. The RMS of HET/HRS velocities with respect to the best Keplerian fit of the HD 37605 system is 9 m/s for data before March 2008, and it is reduced to 6 m/s for data afterwards (see Chapter 6 for more on the HD 37605 system). Such improvement is encouraging, and a closer look at the data and the intermediate products of the Doppler pipeline reveals even more potential contributors to the RV instability of HET/HRS, which is the theme of this chapter.

We set out to first construct a data reduction and RV extraction pipeline for HET/HRS (Section 3.2). Then we tried to improve the precision by seeking out for a better instrumental profile (IP) model, i.e., a functional form that can bring good fit to HET/HRS data (Section 3.3). Section 3.4 describes our efforts in validating the iodine cell Fourier Transform Spectrometer (FTS) scans for HET/HRS and our investigation on the temperature change of the iodine cell.

3.2 Adoption of REDUCE and the CPS Doppler Code for HET/HRS

We built a data reduction pipeline for HET/HRS that performs automatic raw spectral data reduction and RV extraction for any target star or HET observing program.

The first part of our data pipeline is the raw reduction code, which reduces the raw spectral images with calibration frames and produces FITS files containing the extracted 1-D spectrum for each spectral order. We adopted the REDUCE reduction package by Piskunov & Valenti (2002). Sarah Gettel and Stephen Bongiorno did the initial adoption of the REDUCE package for reducing HET/HRS data. It involved much parameter tuning with the original package to make this package, which was designed for slit-fed spectrometers like Keck/HIRES and Lick/Hamilton, to work with the fiber-fed HRS data. The most important modification we did was to automate the routine for tracing the echelle orders on the spectral image, which used to require human-intervention (i.e., clicking on the echelle orders to help the code recognize the traces). Capitalize on the fact that the shape of the order traces do not change significantly across time, we use a set of previously fit polynomials that describe the trace positions and shift them to match the traces on the raw image via cross correlation. We also wrote wrapper routines to automate bulk reduction for HET/HRS observation programs or any single target star.

The second part of our data pipeline is the code that extracts RVs from the 1-D spectral data. We adopted the CPS Doppler code, which is documented in Chapter 2. The adoption process mostly involved bookkeeping and code debugging works to make sure it could run on Penn State Astronomy network machines, to train the code to recognize HET/HRS images, to work with our local file structures, and to be able to produce DSST locally using HET/HRS data (instead of using Keck/HIRES DSST like we did initially, which does not produce RVs as precise as using HET/HRS DSSTs).

In the spirit of making it easier for any future CPS code adopters, we document the requirements and procedures for making such an adoption below. Many of these procedures would not make sense to someone who has not handled the CPS Doppler code before, but they are tremendously time-saving for anyone who wants to use the CPS code to extract RVs from their own spectra.

Required files for Doppler reduction:

- Reduced, extracted 1-D spectra with both iodine and stellar lines. The 2-D matrix holding the 1-D spectra in each order should have the format `matrix[pixel, order]`.
- A log file which lists: observation ID, target of the observation, JD, barycentric velocity correction for the target for the observation, and nature of the observation (iodine cell in light path?). The default log file should be in the `/bary/` folder but it can be customized via keywords of `dop_driver.pro`.
- A `vd` example file (an IDL save file) which contains an example `vd` structure for the spectrograph (see Chapter 2 for more details). The `vd` structure also contains

definition for spectral chunks, i.e. their orders and starting and ending pixels. It also contains first guesses for the parameters for each chunk (most importantly, a rough wavelength solution which must be good within 20 pixels). This should be in the `/files/` folder.

- An IPGUESS file (IDL save file) with the structure `ipguess`, which contains the JD and file names for the `vd` files to be used as first guesses of parameters. The `ipguess` structure must contain at least one `vd` file entry, and very often it has many entries so that the code can pick the most recent one. This should be in the same folder as the Doppler code.
- A DSST file (IDL save file) with the `DSST` structure, which contains the DSST spectrum in chunks, and the definition of chunks must be consistent with the example `vd` structure. This should be in the `/files/` folder.

List of routines to modify:

- `dop_driver.pro`
 - Add a keyword to distinguish jobs for your instrument, e.g., “het=het” for HET/HRS reductions.
 - Add codes to define necessary data paths near line 124.
 - Add a line near line 406 so that the code recognize the first two characters in your observation name, which will be used as an identifier for your instrument. For example, the HET/HRS observations are named as “20160101.00001”, so the identifier for HET/HRS data is “20” (it is “rj” for post-detector upgrade Keck/HIRES data).
 - If you want the sum of Gaussians IP for your instrument, add this option near line 75.
- `crank.pro` Near line 164, add instrument setups such as data paths.
- `rdsi.pro` Add option (`tp eq '20'` for HET/HRS) for reading in the spectral data for your instrument.
- `chip.pro` Add gain value for your detector, even if your reduced spectra are gain-corrected already (then just have `gain = 1`).
- `vdcube.pro` Define the `highcts` near line 120 for your detector, which defines the saturation threshold for your CCD and such observations will be disregarded during vanking. Also add your customized `/files/` path in the code near line 168.
- `rdhbcvel.pro` At the beginning, make sure this code knows that it is handling data specific to your instrument (using, for example, the name of your log file as identifier). Then near line 80, you want to define a full file name for your reduced 1-D spectra based on the observation ID.

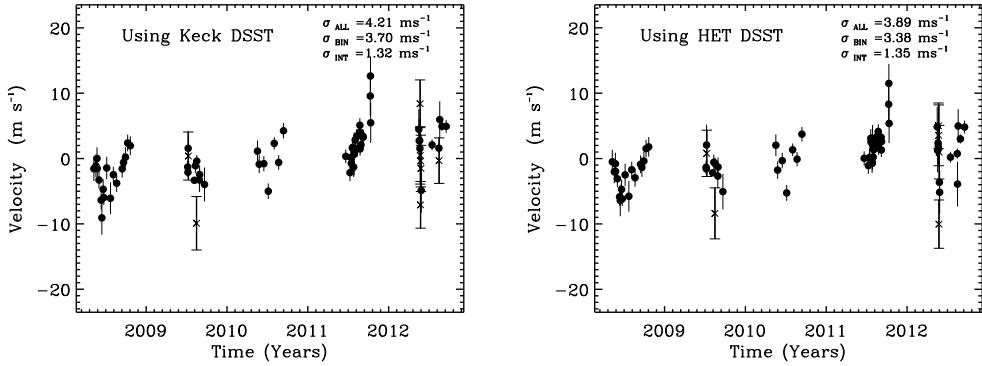


Figure 3.2 RV time series on HD 185144 as observed by HET/HRS and reduced by our data pipeline. The left panel shows RVs extracted using the Keck HD 185144 DSST, and the right panel shows RVs with DSST derived using HET/HRS data and the CPS code we adopted.

- Making your own DSST: `/rv/deconv/deconvolve_stellar_template.pro` contains the code to make DSST using HET/HRS data. A quick walk through `make_dsst_keck_psu.pro` and compare it with `make_dsst.pro` will also tell you how to modify `make_dsst.pro` to make it work for your template observations.

There could be other ad hoc places that need to be modified to accommodate your customized 1-D spectrum format. For example, around line 523 in `stargrind.pro`, the code defines a lower boundary for wavelength at 5000Å, which is completely arbitrary, and the user may need to modify this to accommodate any orders bluer than 5000Å. Or, if you are using sum of Gaussians as your IP model, you want to modify `getpsf.pro` for customized positions and widths for your satellite Gaussians.

A typical call to run the Doppler code looks like this:

```
dop_driver, 'target.star.name', 'your.short.tag', dsstname='file name of your
DSST', dsstobnm='observation name of your DSST', /gh, /vank, /noprint, /het,
tag='your.long.tag'
```

The previous chapter is a general documentation of how the CPS Doppler code works. Other useful documentation for the code are in the `/doc/` folder.

After we completed our data pipeline, we immediately reduced the data on HD 185144, which is a favorite RV standard at both Keck and HET and also many other precise RV facilities. The results, plotted in Figure 3.2, were very encouraging – a 3-4 m/s precision out of the box, without any tuning of the Doppler code. It is also found out later that the RVs produced using the HET/HRS DSST are better than the ones with Keck/HIRES DSST, possibly because the errors in the deconvolution and the forward modeling (convolution) would cancel out if the DSST and star+iodine data come from the same telescope, which is commonly agreed among the iodine precise RV artists.

The typical Keck/HIRES RV RMS on HD 185144 is about 2.57 m/s. Our mission is to find out if and how we can bring the RV precision of HET/HRS to the level of Keck/HIRES.

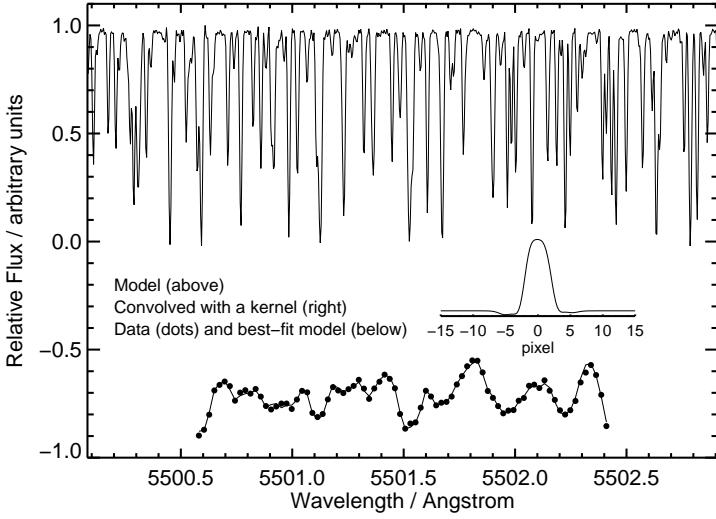


Figure 3.3 Illustration of convolving the iodine atlas (sharp solid lines) with a kernel (middle right insert) to fit the observed iodine lines (black dots near the bottom, with best-fit model plotted in solid line).

3.3 The Search for a Better Instrumental Profile

Finding a good customized function for the instrumental profile (IP) for a precise RV spectrograph is undoubtedly one of the most important and difficult tasks in achieving a < 3 m/s RV precision. IP modeling is a crucial part of the precise RV work with iodine calibration, as it affects directly several key procedures in the Doppler pipeline, such as the creation of stellar spectrum template and the forward-modeling of the observed stellar+iodine spectrum. The heroic efforts of early Keck/HIRES users pinned down its IP to a 12-parameter sum-of-Gaussians profile, with two sets of 11 pre-determined positions and width of the satellite Gaussians (Valenti et al. 1995). It is the product of careful studies and numerous trials with IP modeling.

It is very easy to imagine that imprecise and inaccurate IP modeling could be the bottleneck for improving HET/HRS's RV precision. The ability to capture the shape of the IP and its changes largely determines how well the RV spectra are fitted, and thus how precisely the Doppler shift is measured. The 3-4 m/s precision on HD 185144 we obtained (Figure 3.2) was the results of an “out of the box” reduction – we had only modified the CPS Doppler code to the point where it could process HET/HRS data, but we had not yet put in any efforts to tune it for HET/HRS. The first step of such tuning would be to find a better IP.

How well the IP is being modeled can be tested by fitting a pure iodine spectrum taken by the spectrograph (Figure 3.3), with the continuum source being either a lamp or a (mostly) line-free B star. Such spectra are often referred to as the iodine spectra (or frames or flats), or the B star iodine spectra. The typical χ^2_ν value that we obtain for fitting HET/HRS iodine spectra with a generic IP model (Gauss-Hermite polynomials)

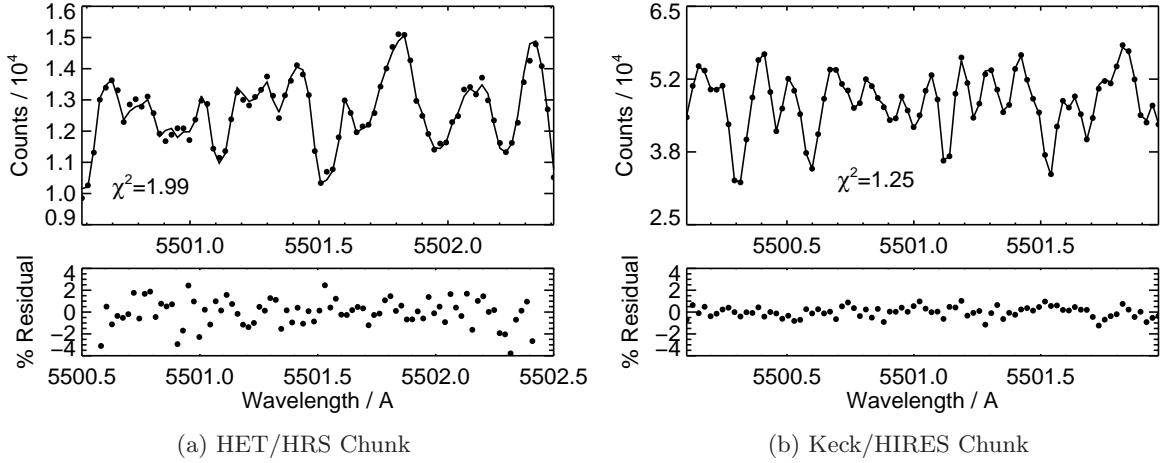


Figure 3.4 Comparison between fits for a typical iodine-only chunk using HET/HRS data (left panel) and Keck/HIRES data (right panel). Bottom panels are showing the residuals against best-fit models, plotted on the same y -axis scale. HET/HRS fit is significantly worse than Keck/HIRES, which we believe is one of the major drivers behind HET/HRS’s poorer RV precision.

is about 2-5, while for Keck/HIRES, the χ^2_ν value is typically around 1 (Figure 3.4). The origin of this difference in χ^2_ν does not seem to rise from a signal-to-noise difference in the spectra (Figure 3.5).

The current “go-to” IP model for HET/HRS is the very versatile, orthogonal, 11-parameter Gauss-Hermite polynomial (GH), which was described in Chapter 2. Another customized IP for HET/HRS was tried out by CPS using the sum of Gaussians, the same as the one used for Keck/HIRES but having the wings at different locations with different default widths. The two IPs basically perform at a similar level, with GH being slightly better (Figure 3.6). We have also tried several other functional forms such as GH convolved with a top hat function with a varying or fixed width, Lorentzian-Hermite (replacing the Gaussian in GH with a Lorentzian), which all performed marginally worse than GH, just like the sum of Gaussians. Or, more precisely, these IPs all seem to be “equally bad”.

We then looked for clues in the Fourier space: Figure 3.7 plots the Fourier transform power spectrum of the HET/HRS data (using the *fft* procedure in IDL) for the entire $\sim 1000\text{\AA}$ 1-D extracted spectrum used for precise-RV purposes; with Keck/HIRES data also plotted for comparison). At high frequency in Fourier space, or shorter periods in pixel space, i.e. on small scales, the power spectrum is dominated by the signature of the IP. A “null” in the power spectrum at 4.3 pixel is clearly visible, which suggests some sort of sharp feature, and indeed, it exactly corresponds to the slit width of HET/HRS projected onto the detector at a resolution of $R = 60,000$. This feature is a direct result of the fact that HRS has the slit in front of a round fiber, creating somewhat of a sharp feature in its IP, unlike Keck/HIRES, whose slit projects onto the (usually) smaller image of the star on the sky.

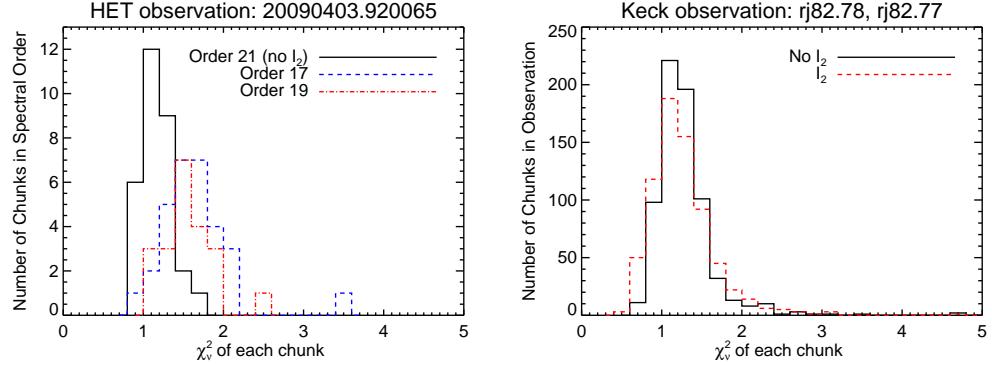


Figure 3.5 Comparison of fits for iodine-free region/observation (fitting a straight line for each chunk) and for iodine region/chunks. The left panel is for HET/HRS data in one observation, but using spectral orders with or without iodine lines. The right panel is for Keck/HIRES data in two different observations with and without iodine cell in place. The fits for iodine-free spectral chunks turn out to be consistent what is expected with photon noise for both HET/HRS and Keck/HIRES. This eliminates additional noise as a suspect in contributing to the bad fits to HET/HRS iodine spectra.

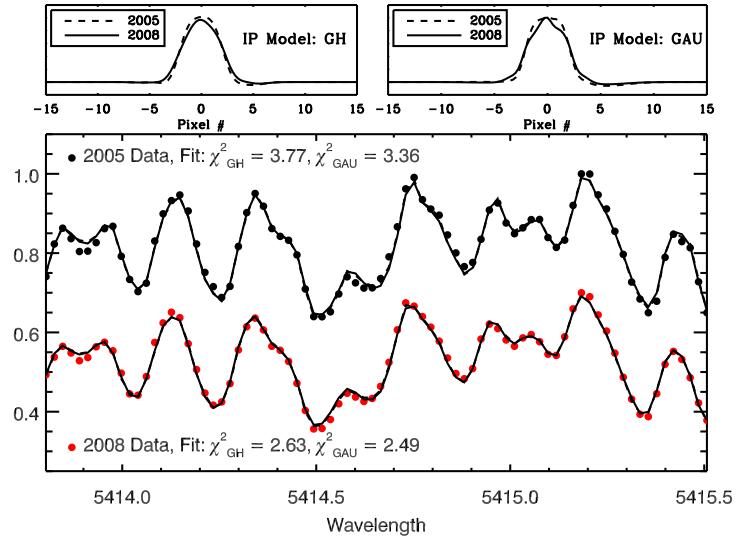


Figure 3.6 Illustration of fitting the iodine-only data (bottom panel) using different IPs (top panels), in this case, using GH and sum of Gaussians (GAU). These two IPs are practically “equally bad”, having similarly large reduced χ^2 but neither produces a satisfactory fit. It is also interesting to see how “stable” the best-fit IP can be across the years (i.e., in 2005 vs. 2008) and its smoothness, hinting that the best IP may take a simple, slowly-varying form.

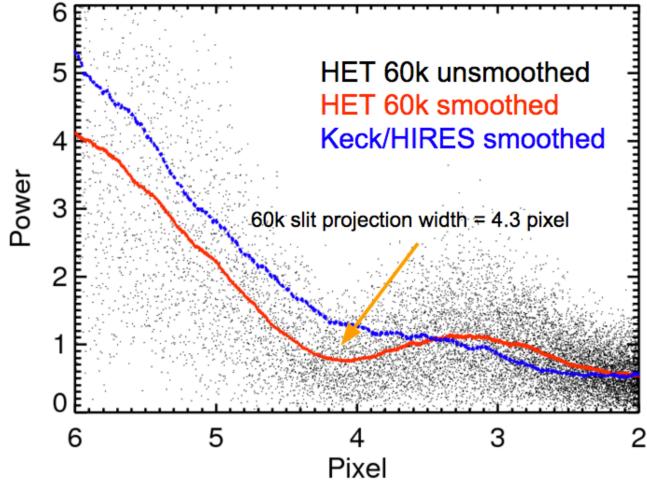


Figure 3.7 Fourier transform or power spectrum of a HET/HRS iodine-only spectrum (black dots) and its smoothed version (blue line). There is a clear signature of the HET/HRS slit at 4.3 pixel (corresponding to slit width for resolution $R = 60k$). For comparison, the red curve is for Keck/HIRES data, which shows no clear signature of a slit, because Keck/HIRES is not fiber-fed and the PSF of the star falls mostly within its slit.

Upon seeing the Fourier transform of the HET/HRS data, we tried out another IP using GH multiplying a triangle (with a half width of 2.4 pixel and a height of 1), whose Fourier transform has a null at 4.3 pixel, and it produced the best fit among all IP models we have ventured. Figure 3.8 illustrates this new fit in comparison with the GH IP fit, although it was perhaps still equally bad.

The little difference between different IPs made us suspect that the “ground truth” for the iodine lines, the iodine atlas, may be problematic. It would not be possible to derive a correct form for the IP using a wrong iodine atlas, and thus we shift our priority towards validating the iodine atlas and investigating possible changes in the cell, which is described in the next section.

Besides the problem with the iodine atlas, which fundamentally prevents us from finding a precise IP, we know for sure that the GH function does not work very well mostly because of the fitter fails to converge for this complicated function. We have two lines of evidence supporting this statement. The first one is that the GH IP performs terribly on Keck/HIRES data because the L-M least- χ^2_ν fitter has trouble converging (unless fine-tuned and informed from previous fits using sum of Gaussians; Vanderburg et al. 2013). The second piece of evidence is that we tried to fit GH to unsaturated ThAr lines in HET/HRS calibration frames, and it often fails to converge onto a good fit to the ThAr line.

To end this section with a somewhat positive note, we present a promising lead for a better IP function for HET/HRS, the modified Moffat function:

$$[1 + (x/\theta)^2]^{-\beta \cdot (x/\delta)^2}, \quad (3.1)$$

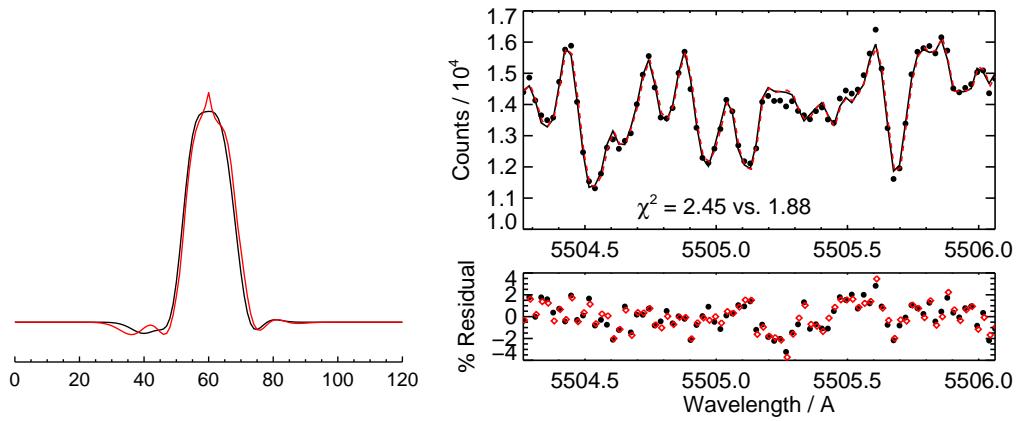


Figure 3.8 Introducing a sharp feature into the HET/HRS IP model, a triangle on top of the GH IP (red curve in the left panel), produces a better fit, somewhat to our surprise. The black in the left panel is the best-fit GH IP. GH+triangle is the IP model that produces the least χ^2_ν among all of our IP models. However, as shown by the right panel, the two fits barely have any visible difference (red curve for GH+triangle IP and black for GH; bottom panel plots the residuals). Such a sharp feature in the IP is nonphysical, and we interpret this results as a hint for an unreliable iodine atlas (the sharp peak at the center is perhaps the IP model trying to “stretch” the iodine lines deeper; see Section 3.4 for more details).

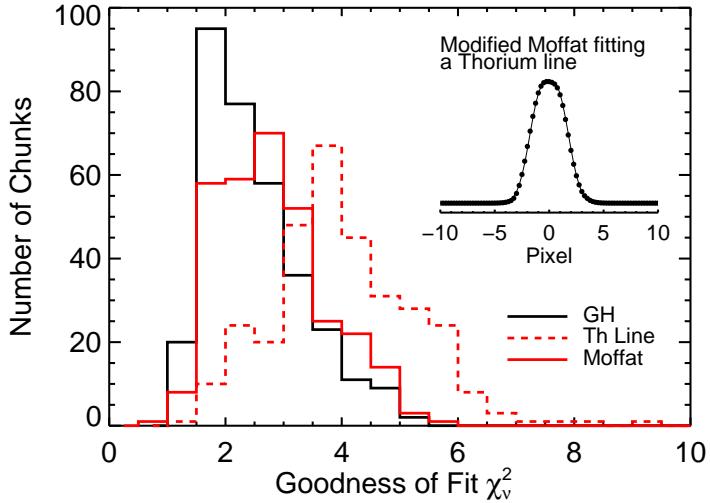


Figure 3.9 Histogram of goodness of fit, χ^2_ν , values for spectral chunks of an iodine spectrum. The modified Moffat function (red) performs almost equally well while having only 3 parameters, compared with the complicated 11-parameter GH function (black solid). Red dashed histogram is for fits using a ThAr line profile as IP. The insert is showing the modified Moffat function can fit a ThAr line quite well.

where x is the pixel coordinate, θ determines the characteristic width of the profile (typically ~ 0.6), and β (typically ~ 6) and δ (typically ~ 8.6) determine the shape of the wings. It is called the “modified” Moffat function because the original Moffat function does not have the $(x/\delta)^2$ term. We added this term to add flexibility at the wings to enable change of characteristic IP width while preserving wing profile. Figure 3.9 illustrates the results using the modified Moffat fitting a ThAr line (insert), and also the χ^2_ν distribution of all spectral chunks for this new IP compared with the GH IP. Unfortunately, the modified Moffat function does not always fit a ThAr line (starting with uninformed initial guesses), so it faces the same problem as GH. However, it only has four parameters and they are mostly physically meaningful. For example, one can imagine that the θ parameter describes some characteristic width. This makes this function easier to work with than GH.

One can imagine getting a better fit by adding small perturbation terms to the modified Moffat IP to account for asymmetries and subtle wings due to scattered light. Moreover, the modified Moffat function is potentially applicable to other fiber-fed instruments, since such instruments are likely to have IPs with the same characteristic flat top and sharp wings. We hope to continue this effort after the iodine atlas problem is resolved (see next section) and carry on this knowledge to other projects such as the new HRS and MINERVA (Chapter 7).

3.4 Investigation on the Iodine Cell and the Iodine Atlases

As discussed in the previous section, successful modeling of the iodine observations (B star spectra taken through the iodine cell) is a good indication of a working radial velocity (RV) pipeline. We have explored one of the two model components for modeling iodine observations, the choice of IP. In this section, we examine the other model component, the iodine atlas, that is, the “ground truth” spectrum for the iodine absorption lines unique to the HET/HRS iodine cell. A “ground truth” iodine atlas is crucial for the precise iodine radial velocimetry. It is used for modeling the observed iodine lines in the stellar+iodine RV observation to anchor the absolute wavelengths and the spectrograph response function.

3.4.1 Why did we suspect the iodine atlas?

An iodine atlas is normally obtained using a Fourier Transform Spectrometer (FTS; whose mechanism is just like a Michelson interferometer). The scan and its subsequent data reduction provides very high resolution spectrum (translated from Fourier space into real space) with typically $R > 200,000\text{--}500,000$.²

The existing iodine atlas for the HET/HRS cell is from an FTS scan taken at the National Solar Observatory at KPNO using the Babar FTS (nicknamed for its large size; this machine has been decommissioned) in 1993. The main reason why we took a new scan is that the FTS scan was taken almost two decades ago, and during this time the cell may have gone through changes (such as temperature, leaking or condensation, etc., though unlikely, since the cell was designed to be stable). This would mean that the FTS scan is out of date and inaccurate, and it could explain the bad fits to the iodine observation.

Therefore, we arranged the HET/HRS cell to be sent to the National Institute of Standards and Technology (NIST) facility in Gaithersburg, MD and obtained a new FTS scan in 2011. The work was done by Drs. Gillian Nave and Stephen Redman, in coordination with Drs. William Cochran, Jason Wright, and Jeff Valenti. A close comparison between this new scan from NIST and the old scan at KPNO reveals that they have many differences:

- The overall line depths are very different — the NIST scan has deeper lines.
- The absolute wavelength solutions are different, and the drifting of wavelength solution or the dispersion scales at different wavelength are also different.
- Even after we adjust the normalization level of the NIST scan (assuming the FTS data has normalization issues or low frequency noise/offset), the line ratios of the two scans still exhibits differences.

Figure 3.10 shows the comparison between the two scans in a selected 2Å region. As the two scans also differ in resolution (the NIST scan has a higher resolution), the

²It is worth noting here that although FTS normally provides wavelength solutions (from the registered arm lengths), but because of the inaccuracy of the default reported wavelengths, the final wavelength solution for the iodine atlas is usually derived from a theoretically computed iodine line list (e.g., Knöckel et al. 2004).

middle panel is a more direct comparison: the NIST scan has been convolved down to the same resolution with the KPNO scan; it is also shifted in wavelength space so that the two scans match in absolute wavelength solution; and it is adjusted to a different normalization level to match with the KPNO scan as much as possible in order to compare their relative line ratios.

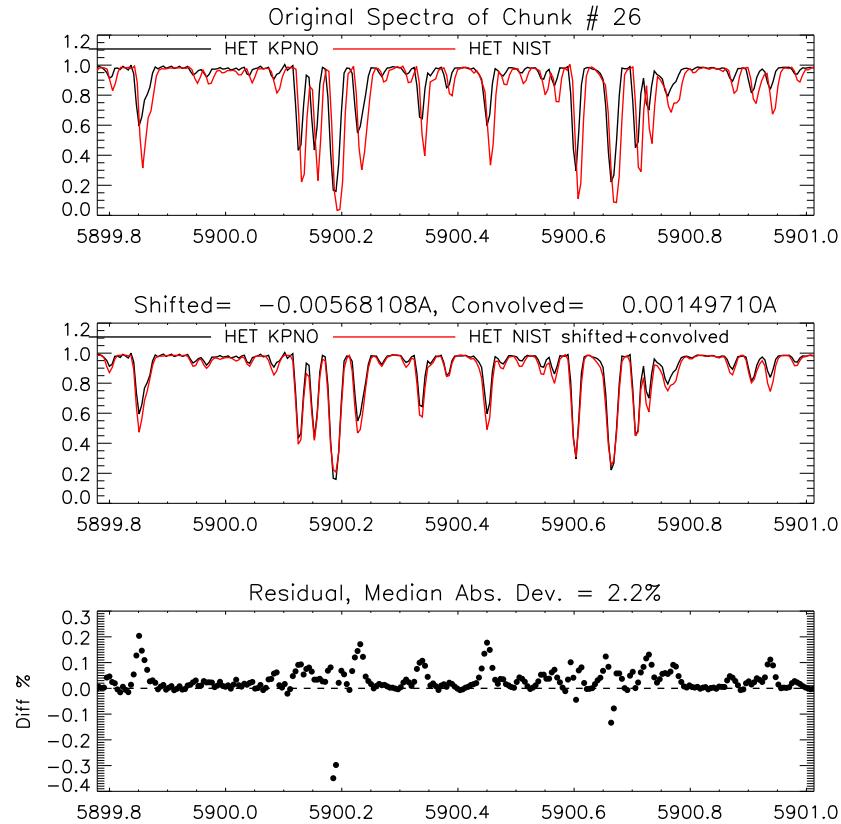


Figure 3.10 Comparison of the KPNO FTS scan (black) and the NIST FTS scan (red) for the HET/HRS iodine cell for a selected 1.5 \AA chunk. **Top:** Two scans at their native resolution and original wavelength solution. **Middle:** Comparison of the two scans after adjusting the normalization, shifting, and convolution for the NIST scan to match the KPNO scan for a more direct comparison of line depths/ratios. **Bottom:** Residuals of the middle panel, NIST spectrum minus the KPNO spectrum. The median absolute deviation between the two spectra is 0.02 (2%), though at many places, especially at line centers, the two can differ by up to 5–10%.

We initially suspected that the NIST scan was problematic. The reason is illustrated in the left panel of Figure 3.11, where it shows the histogram of χ^2_ν values for fitting an selected iodine observation using the two scans, respectively. Each χ^2_ν value is for a 2 \AA chunk in this selected iodine observation. It is clear that the NIST scan provides worse fits.

Since the direct comparison between the KPNO scan and the NIST scan has hinted

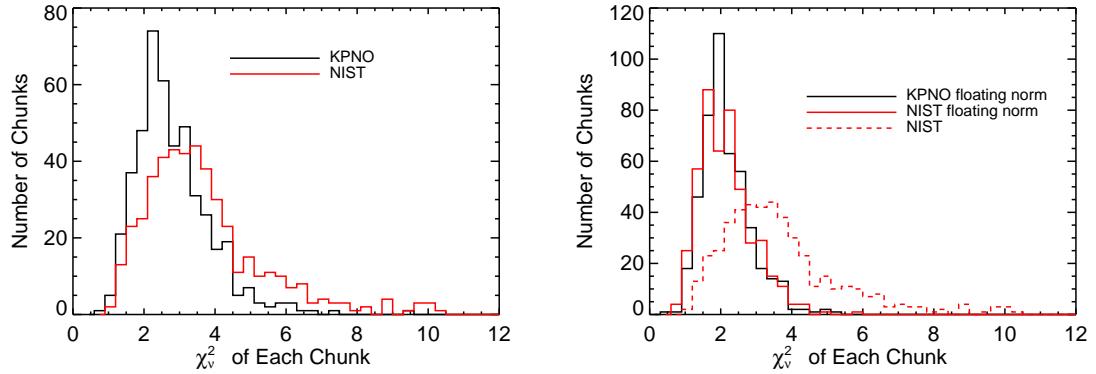


Figure 3.11 Both plots are histograms of χ^2_ν values of a single iodine observation. Each χ^2_ν value in the histogram represents the χ^2_ν goodness of fit for a $\sim 2\text{\AA}$ spectral chunk in this iodine observation (each iodine observation is chopped into several hundred of chunks and is fitted independently). **Left:** χ^2_ν histograms for the fit of the iodine observation using the KPNO (black) and NIST (red) scan as iodine templates, respectively. The KPNO scan obviously performs better. **Right:** χ^2_ν histograms for the two scans, but both with the normalization as a free parameter for each chunk (as we suspect the NIST scan has problems in normalization). The two scans now perform at essentially the same level. Dashed red line is the same red histogram as plotted in the left panel. Notably, the KPNO scan also performs better when we float the normalization parameter.

that the normalization of the NIST scan might be problematic, we decide to add a free parameter to account for this normalization error when fitting the iodine observation. The right panel of Figure 3.11 shows the χ^2_ν histograms for the same iodine observation using the two scans, but adding a free parameter as the normalization when fitting each chunk (note: the normalization parameter is a free parameter for each chunk, not a global single parameter). The two scans now perform at essentially the same level.

This is both encouraging and worrisome at the same time. It is encouraging because it seems that we have found the problem with the NIST scan, and also have a solution for it. It is very worrisome because this reveals that:

- Even the KPNO cell performs visibly better when we float the normalization parameter. This may suggest that there are normalization issues or low frequency errors/noise in the KPNO scan as well.
- Obtaining high-quality, reliable FTS scans of iodine cell is very difficult, and the FTS scans cannot be naively trusted as the “ground truth” super accurate templates of the complicated iodine spectrum.
- The reason why adding a floating normalization fits the data better might be because it accounts for optical depths difference between the atlas and the actual observations, which may be a result of changes in cell temperature or iodine column density in the cell.

- The pipeline (when floating normalization as a free parameter) cannot distinguish which scan is more accurate (by χ^2_ν) even when two scans differ as much as $\sim 5\text{--}10\%$ at places and also have obvious line ratio differences (see comparison in bottom panel of Figure 3.10). However, this level of difference in FTS may affect the RV precision, and not knowing which atlas is the correct one definitely affects our ability to search for a better IP and improve the RV precision of HET/HRS.

Perhaps even more alarmingly and more puzzling, when we use the KPNO scan for the iodine cell used at Keck/HIRES to fit an HET/HRS iodine observation, it yields smaller χ^2_ν values than using any of the other two scans (Figure 3.12). The HET/HRS cell KPNO scan was taken at the same time using the same FTS machine as the Keck/HIRES cell scan. However, the set-temperatures of these two cells are very different: the Keck/HIRES cell is designed to work at 50°C, while the HET/HRS cell is designed to work at 70°C. A closer look reveals that the HET KPNO scan and the Keck KPNO scan have very similar line depth, with Keck having slightly deeper lines (thus higher iodine molecule column density).

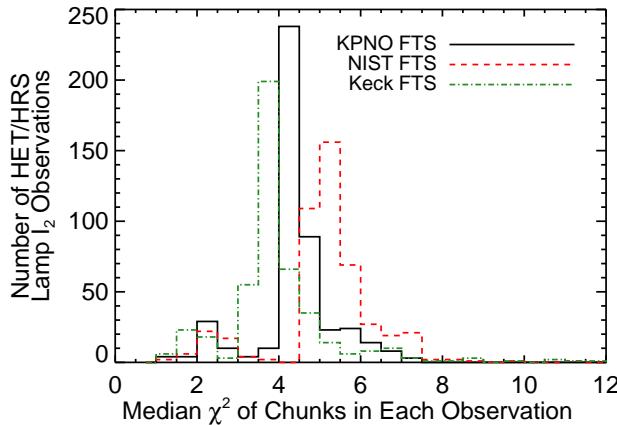


Figure 3.12 Comparison of the median χ^2_ν values for fits of iodine observations using the HET/HRS cell KPNO scan (black solid line), the NIST scan (red dashed), and the Keck/HIRES cell KPNO scan (green dotted-dashed). Each data point represents the median χ^2_ν value for all the chunks in a single iodine observation (these are all lamp-illuminated – no B star observation). Results of 550 HET/HRS observations are plotted here to illustrate the statistically significance. The Keck/HIRES cell KPNO scan provides a better fit than the both HET/HRS scans when fitting HET/HRS iodine observations.

All of the facts above prompted us to seek a relatively independent way to perform quality checks for any FTS scan — not just comparing their relative qualities or performances. One natural choice is to obtain spectra taken with high-resolution echelle spectrographs, which are measurements of the iodine spectrum directly in the real wavelength space instead of in the Fourier space, and thus they serve as good reference spectra as they suffer from different types of error compared to FTS. Since FTS scans

are usually at a very high spectral resolution (200,000–500,000), this limits our choice to essentially only one spectrograph — the TS12 setting of the Tull Spectrograph at the 2.7m Telescope at McDonald.

Our goal is to answer the question of **which FTS scan better describes the HET/HRS iodine cell: the KPNO one, or the NIST one? Why?** While adding an additional normalization parameter can provide better fits for the iodine frames, we cannot afford to add such an additional parameter when extracting RVs from star+iodine data – this normalization parameter is highly degenerate with Doppler shift and wavelength solution parameters, and it will undoubtedly decrease the RV precision. Another key question is **why the FTS scan for the Keck/HIRES cell works the best.**

3.4.2 Ultra-High Resolution Echelle Spectra of Iodine Cells

To break the tie between various FTS scans, we used the TS12 setting of the Tull Spectrograph at the 2.7m Telescope at McDonald Observatory, which provided a resolution of 500,000 (based on ThAr line measurements done by David Doss at McDonald).

We had two rounds of TS12 runs, both during day time and when the telescope was scheduled for Cassegrain instrument and the Tull Spectrograph room was free for use. For the first run (from September 7 to September 9, 2013; done by Ming Zhao and the author) we measured the iodine absorption spectrum for the iodine cell at the Sandiford (2.1m) Telescope, because the HET/HRS cell was unavailable (it was still under active use for planet search programs not long before we did this test). The main purpose of the first run was to validate the quality of the TS12 spectrum and whether we can use it to cross-validate the FTS scans. The Sandiford cell also has a KPNO FTS scan which was taken together with the KPNO scan of the HET cell in 1993, so it also serves the purpose of testing the overall quality of the KPNO scans.

In the second TS12 run, we measured the spectrum for the HET/HRS iodine cell in its new enclosure and temperature controller (along with the MINERVA iodine cell and the iodine cell on McDonald Harlan J. Smith 2.7m telescope; from October 13 through October 16, 2014; carried out by Ming Zhao, Kim Star, and Joey Schmitt). Our TS12 runs are enabled by Anita Cochran, Bill Cochran, Phillip MacQueen, and the astronomers who used the Cassegrain instruments at night (VIRUS-W for the first run and IGRINS for the second), with great help and excellent engineering support from David Doss and Coyne Gibson at McDonald Observatory.

The hardware settings and data reduction methods are the same for both runs, which are described below.

Hardware Settings: We used the TS12 arm of the Tull Spectrograph, and the specific instrument choices are listed in Table 3.1. Slit #23 is chosen to maximize SNR while maintaining sufficient resolution — it is among the longest slit and is also the second narrowest slit. The Sandiford cell was kept at a temperature of 49.9–50.1°C, the same as its working temperature for RV work and its temperature when the KPNO was taken (50°C). The HET/HRS cell was measured at four different temperatures: room temperature, 50°C, 60°C, and 70°C (its working temperature).

Observation: A single exposure frame for the iodine spectrum covers about 1.9Å (Figure 3.13). The dispersion direction runs vertically along the chip with increasing

Table 3.1. Hardware Settings for
TS12 Iodine Spectrum Test

Parameter	Unit and Meaning
Tull Spectrograph, TS12, Coude107	
Echelle	E1
Cross Disperser	c
CCD	TK4, 1024×1056
On-chip Binning	1×1
Slit	#23 ($L \times W = 30'' \times 0''.32$)

wavelength when increasing the y -axis pixel. The dispersion scale is about 0.002\AA per pixel (~ 7 pixels per resolution element). We immediately preceded or followed each exposure with a flat fielding frame. The exposure times for the iodine and flat frames are both 45 seconds (90 seconds for HET/HRS cell) to achieve a signal-to-noise ratio (SNR) of 160 per pixel (higher for HET/HRS cell). Neighboring frames differ by about 1\AA in absolute wavelength. If prominent Solar or ThAr line was predicted within the wavelength coverage of a frame, then we also took a Solar or ThAr frame to verify the rough wavelength solution (the exposure time varied — typically a couple minutes to up to 10 minutes). We took dark frames (45s each, about 10 frames) in the morning at the beginning of each day.

Reduction: We combined and averaged all available dark frames and created a master dark frame. Then we subtracted the master dark from all flat and iodine frames. After outlier rejection (cosmic rays, chip defects, etc.), we modeled the scattered light for each row of pixels by using the region outside the slit image. We stacked 160 neighboring rows and fitted a third order polynomial along the column, and then interpolated for the amount of scattered light within the slit image region and subtracted it. Both the flat and iodine frames have scattered light removed. Then we normalized the flat frames and divided each iodine frame by its associated normalized flat (for the slit image regions only).

Extraction: As the slit does not lie perfectly along the x -axis direction on the chip, we corrected for this by cutting columns along the dispersion direction and cross-correlating the columns. Then we interpolated and shifted the columns to create an aligned image, which we stacked along the x -axis direction and obtained the reduced, extracted spectrum. Each spectrum is then normalized by dividing the estimated continuum (top 5% counts). Due to lower quality of scattered light removal near the edge of the chip, we discarded the top 80 and bottom 80 rows of pixels. Thus the extracted spectrum from each frame is about 1.6\AA across (instead of 1.9\AA). The reduced frames are then stitched together by finding the overlapping region through cross correlation for each pair of neighboring frames and taking into account the changes and differences of dispersion scales across frames.

Mapping onto FTS: To compare with the FTS scans, we chopped the TS12 spec-

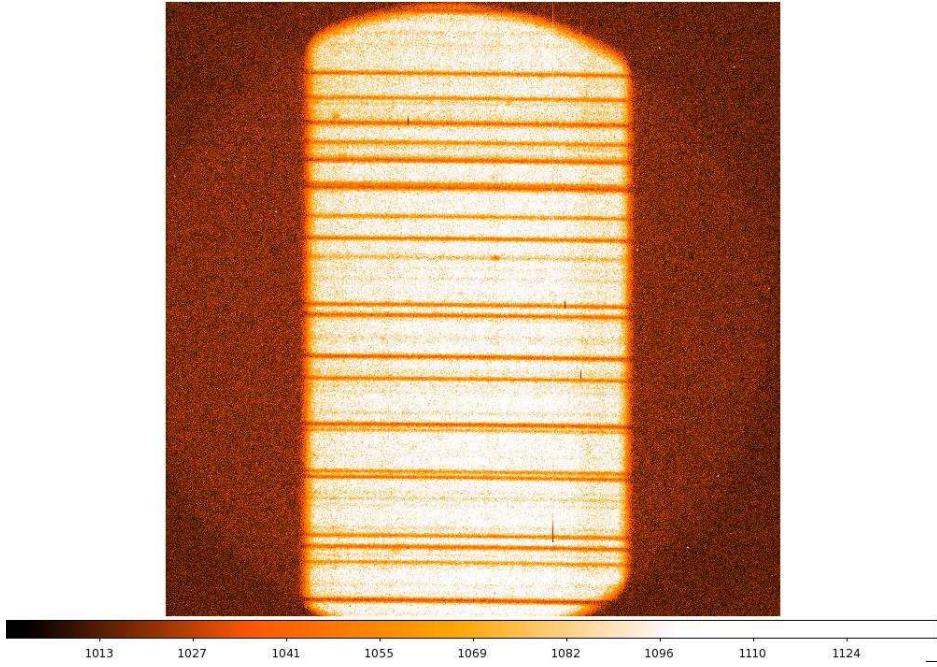


Figure 3.13 One raw image frame taken using the TS12 setting of Tull Spectrograph. It contains about 1.9\AA of iodine absorption spectrum.

trum into 2\AA chunks and project each chunk onto the FTS spectrum by cross correlation. In this way we obtained the absolute wavelength solution and dispersion scale (as set by the wavelength solution of the FTS scan) for the TS12 spectrum.

The results from our first TS12 run using the Sandiford cell demonstrated that an iodine cell spectrum taken with TS12 has the same quality as an FTS scan to serve as the ‘true solution’ of the iodine spectrum. The left panel of Figure 3.14 shows a direct comparison of the reduced TS12 spectrum (a random 2\AA chunk) with the KPNO FTS scan, at their native resolutions.³

To make a more direct comparison and also to see the differences of the two spectra (if any) would make a significant impact when fitting a 60,000 resolution iodine observation, we degraded the resolution of both spectra to 60,000 by convolving them with a Gaussian of a proper width. The right panel of Figure 3.14 illustrates the comparison of the two spectra at $R \sim 60,000$, with residuals of the TS12 spectrum minus the KPNO FTS spectrum plotted in the bottom panel. The two spectra differ by a median absolute deviation of 0.3% (0.4% for the entire $\sim 30\text{\AA}$ spectrum available as shown in Figure 3.15).⁴ As the TS12 spectrum has a SNR of about 160 and we have con-

³Note that the TS12 spectrum appears to have a higher resolution than the FTS scan. According to the header of the FTS scan, its resolution is about 491,000. An FFT analysis on the TS12 spectrum (to see where the high-frequency signal cuts off and becomes indistinguishable from the noise) shows that its resolution is about 455,000 and maybe even higher.

⁴For comparison: when fitting the HET/HRS Iodine observation used for creating Figure 3.11 (median SNR for a typical chunk is ~ 150 , or 0.65% shot noise), for a typical chunk, the median absolute deviation between the observation and the best-fit model is 0.73% (the RMS value is 1%, thus χ^2_ν is $\sim 2\text{--}3$).

volved the comparison spectrum down to $R \sim 60,000$, the expected shot noise should be $\sim 1/160 \times \sqrt{450,000/60,000} = 0.23\%$. The additional $\sim 0.1\%-0.2\%$ of noise may come from flat fielding, scattered light removal, cosmic ray removal and interpolation between pixels, stitching of spectra, projection onto the FTS spectrum and interpolation for comparison purposes, and so on.

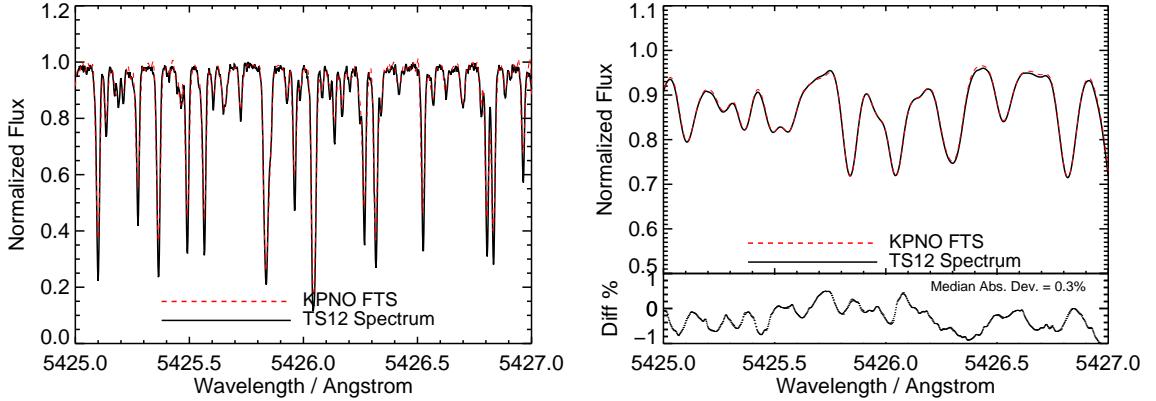


Figure 3.14 Comparison of the Sandiford iodine cell KPNO FTS spectrum and the spectrum taken with TS12. **Left:** Comparison of the two spectra in their native resolutions (both about 400,000–500,000). **Right:** Comparison of the two spectra convolved down to about 60,000 resolution, which is the resolution of typical iodine observations or radial velocity observations (star+iodine). Bottom panel shows the residuals in percentage of the TS12 spectrum minus the KPNO spectrum, with a median absolute deviation of 0.3%.

After demonstrating that we can use TS12 spectrum to validate FTS scans, we performed our second TS12 run with the HET/HRS cell (and the 2.7m cell and the MINERVA cell). For the 2.7m cell, its TS12 spectrum matches very well to its FTS atlas, again (together with the 2.1m cell data from 2013) proving that TS12 is an appropriate tool for validating FTS atlases. For the HET/HRS iodine cell, the results are very informative:

(1) Assuming that the HET/HRS cell temperature control was reliable during our TS12 run (Phillip MacQueen from McDonald Observatory, who built the cell and its enclosure, was at the TS12 run to set it up), then temperature change on the order of 5–10°C in iodine cell should induce a visible change in the absorption spectrum (right panel of Figure 3.16). On the other hand, the temperature of the iodine gas in the cell was not at its set temperature of 70°C during the NIST scan (left panel of Figure 3.16). This could explain the difference between HET/HRS cell’s NIST scan and the KPNO scan. The NIST scan appears to have stayed at higher than 70°C the entire time (one hypothesis is that the gas on the light path was heated up by the ultra-luminous continuum lamp, while the glass container, which the temperature probe was monitoring, stayed cool). The Keck/HIRES cell was also scanned at three different temperatures (50, 60, and 70°C) at KPNO in 1993, and there are also visible differences between these three sets of scans.

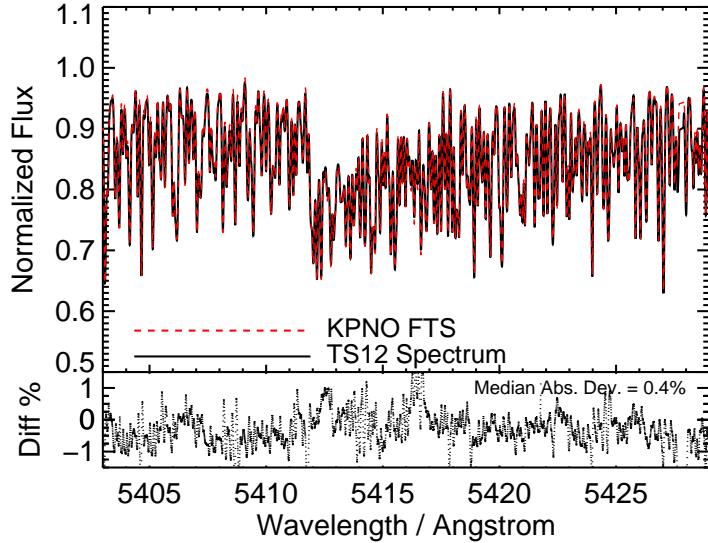


Figure 3.15 The same as the right panel of Figure 3.14, the KPNO spectrum and the TS12 spectrum for the Sandiford cell both at 60,000, but for the entire $\sim 30\text{\AA}$ TS12 spectrum available.

(2) The TS12 spectrum at 70°C matches better with the more recent but potentially problematic NIST FTS atlas (Figure 3.17). This is completely unexpected and contradicts with the assumptions and/or conclusions we lay out in finding (1). The TS12 spectrum also does not match the FTS scan well enough given the high SNR nature of both. Perhaps neither TS12 or the NIST scan was at 70°C . Or, perhaps the amount of iodine vapor was somehow different.

These results prompted us to resolve for a second route to try to break the tie: using synthetic iodine absorption spectra, which is described in the next subsection.

3.4.3 Measuring iodine cell temperatures using synthetic Iodine spectra

Using the TS12 spectra, we found that the temperature of the iodine gas in the cell might not be the same one reported by the temperature controller. However, we still could not break the tie between the KPNO scan and the NIST scan for HET/HRS cell: the KPNO scan provides a better fit to real observed data, but the TS12 spectrum shows us that the NIST scan looks closer to the “truth” as defined by TS12. Nor did we understand why the KPNO scan for the Keck/HIRES cell works the best for HET/HRS data.

To answer these questions, we have found⁵ a second venue that provides reliable, ultra-high resolution, and wavelength calibrated iodine atlas – a theoretical code that computes synthetic iodine transmission spectrum (at any specified temperature) based on both physics and empirical calibrations (IodineSpec5; Knöckel et al. 2004). We emailed the authors and obtained the code, which only runs on Windows machines. The direct

⁵The author is deeply grateful for Iouli Gordon for introducing IodineSpec5 during her visit at CfA.

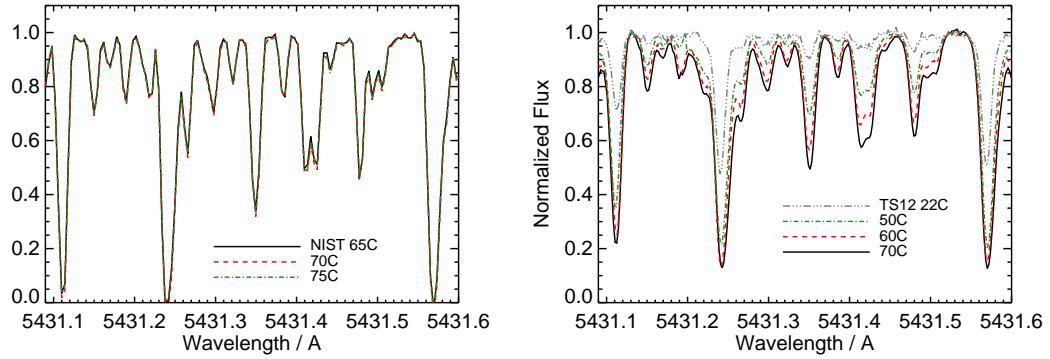


Figure 3.16 **Left:** HET/HRS cell NIST FTS scan at three different temperatures. **Right:** HET/HRS cell TS12 spectra at four different temperatures for the same wavelength region, which, unlike the NIST scans, shows significant difference when the temperature of the cell changes by 10°C.

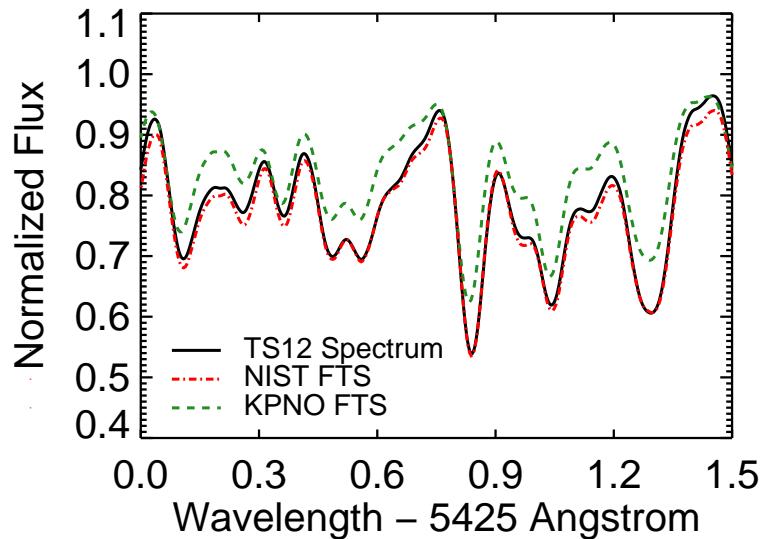


Figure 3.17 TS12 spectrum (black solid line) vs. NIST FTS (red dotted-dashed) vs. KPNO FTS (green dashed) for the HET/HRS iodine cell at 70°C, all convolved down to a resolution of $R = 60k$ (the same as a typical HET/HRS observation) for comparison purposes. The TS12 spectrum matches the NIST FTS better, having deeper lines compared to the original KPNO FTS. The remaining difference between NIST FTS and the TS12 spectrum might be due to differences in cell temperatures or other changes with the cell.

output of the code contains arrays of wave numbers and effectively, opacity (α) for user-specified iodine isotope mix (for our purposes we only add $^{127}I_2$), temperature, wave number range, and line broadening kernel parameter (we chose thermal/Gaussian). To use the output of IodineSpec5 to fit an actual iodine absorption spectrum, there are two parameters: gas temperature and a constant which scales with iodine molecule column density, which we simply refer to as the column density hereafter.

Quickly comparing the NIST scan with the IodineSpec5 models reveals that the NIST scan seems to be around 110°C, mostly based on visually examining the line ratios (Figure 3.18). However, the synthetic iodine spectrum and the FTS scans have different broadening kernels. In order to “fit” the NIST scan with the synthetic spectra at various temperatures, we convolved the NIST scan with a single Gaussian kernel with $\sigma = 0.0078$ (roughly at $R = 200,000$ to try to match the Keck/HIRES cell KPNO scan for comparison and illustration purposes). Except for the Keck/HIRES scan, we did the same to lower the resolution of other FTS scans or TS12 spectra when using IodineSpec5 to find out their temperatures. There are four parameters in our fit: temperature, column density, resolution (σ for the single Gaussian kernel to convolve the synthetic spectrum with), and a wavelength shift. We first optimized the column density, σ , and the wavelength shift while fixing the temperature (using L-M least- χ^2 fitter using `mpfitfun` package in IDL), and then we compare the goodness of fit of each model at different temperatures to determine which temperature best describes the FTS scan or TS12 spectrum. The reason for this two-step optimization is because we have to generate the IodineSpec5 model spectra on a discrete temperature grid.

We first fitted the Keck/HIRES cell KPNO FTS scan, whose temperature is known (50°C) and reliable. We also know that this FTS scan is probably true to its reported temperature since Keck/HIRES iodine atlas fits the data very well, as described in the previous subsection. Choosing a region where it contains temperature-sensitive lines, we found the best-fit temperature for the Keck/HIRES KPNO scan is 55°C, although synthetic spectra ranging from 40-70°C all fit the data almost equally well and they are hard to distinguish (column density and temperature are highly degenerate parameters). We thus conclude that IodineSpec5 is reliable for estimating temperatures for iodine FTS scans, at least to an accuracy of 5-10°C.

We then fitted the NIST scan, which has the highest SNR and resolution among all FTS scans or TS12 spectra (we also have a rough idea about its temperature). The results are shown in Figure 3.19, where the top panel shows the best-fit models at different temperatures, and the bottom panel compares the NIST scan with its the best-fit model at 110°C, the KPNO scan, and the Keck/HIRES cell KPNO scan.

When we tried to fit the HET/HRS KPNO scan, the high degeneracy between column density and temperature hindered us from getting an accurate estimate for the temperature. Models in 40-80°C appear to fit equally well with varying column densities. However, only the fit at 70°C has the same column density as the best-fit value derived from the NIST fit. We thus fixed the column density in our fits for HET/HRS KPNO scan, and (unsurprisingly) the best-fit temperature came out to be 70°C. Therefore, we conclude that the NIST scan was indeed at a different gas temperature, and the old KPNO scan seems to be at the right temperature.

But how about the TS12 spectrum? Again, using the best-fit column density derived

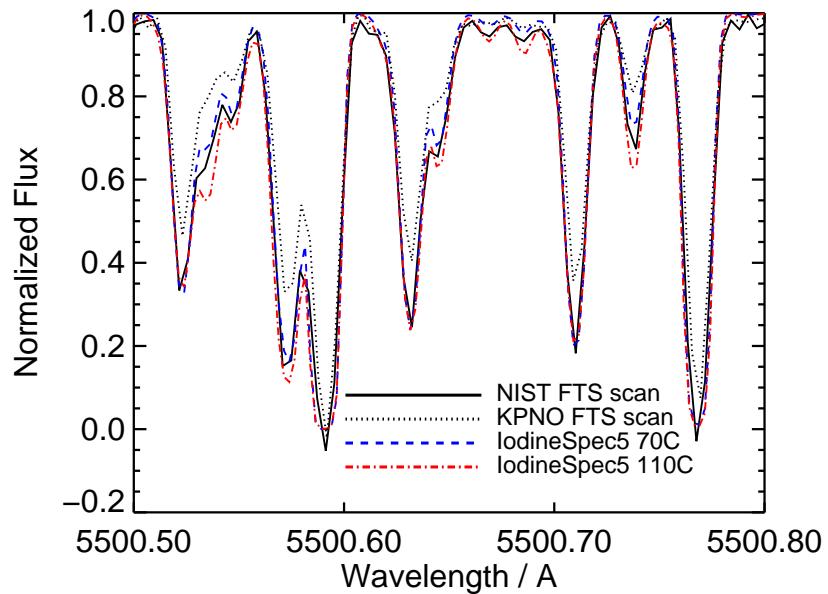


Figure 3.18 NIST FTS (black solid lines) and KPNO FTS (black dotted lines) compared with theoretically computed iodine lines at 70°C (blue dashed) and 150°C (red dotted-dashed). All spectra are at their original resolution. There are two free parameters for the theoretical lines: temperature and iodine column density. For this plot, we optimized the iodine column density for the theoretical lines at both temperatures to try to match the NIST FTS. As illustrated, neither temperature can produce a good match, and the best temperature is around 110°C. Note that the theoretical lines and the NIST FTS have different broadening kernels. The NIST and KPNO FTS scans probably differ in both optical depth and cell temperature.

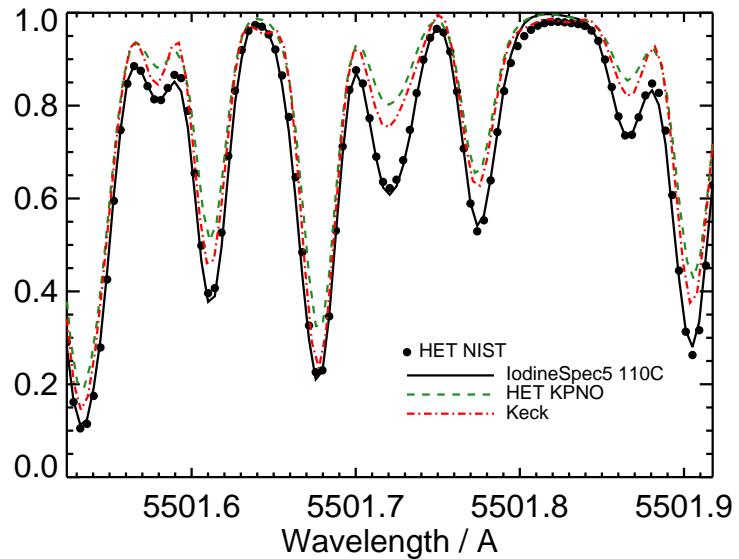
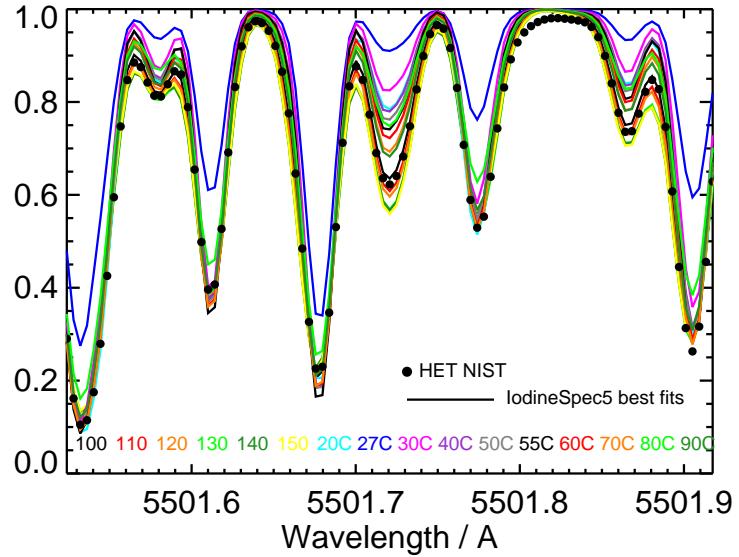


Figure 3.19 Top: Fitting HET/HRS NIST scan (black dots; temperature set at 70°C) with IodineSpec5 synthetic iodine lines at various temperatures and column densities. **Bottom:** The HET/HRS NIST scan (black dots) overplotted with the best-fit Iodine-Spec5 model (black solid line; at 110°C) and the HET/HRS cell and Keck/HIRES cell KPNO scans. All spectra in both panels are convolved down to a resolution of 200,000 (roughly Keck/HIRES KPNO FTS scan resolution) to wash out the intrinsic IP difference between FTS scan (sinc function IP) and the synthetic spectrum (only natural broadening IP models).

from our NIST and KPNO fits, we estimated the temperatures for the TS12 spectra (supposedly) at 50, 60, and 70°C. The best-fit temperature turns out to be 55°C for claimed 50°C TS12 spectrum, 80°C for the 60°C one, and 100°C for the 70°C one. The results fitting for the “70°C” TS12 spectrum are in Figure 3.20.

These findings on HET/HRS cell temperatures could explain the fits to real data. If the HET/HRS cell was kept at a higher temperature (e.g., $\sim 100^\circ\text{C}$) instead of 70°C when it was in active use for precise RV calibration (despite what the temperature control reported), or if the actual temperature of the gas in the cell changes over time, then the KPNO scan, which was done at 70°C, certainly cannot fit the observed data and provide precise calibrations to measure RVs to the level of Keck/HIRES, which has a precise iodine atlas. Looking at the bottom panel of Figure 3.20, it is not hard to imagine why the Keck/HIRES cell scan provides the best fits to HET/HRS data (Figure 3.12) – if the gas temperature is between 70 and 110°C during actual observations (and perhaps more often closer to 70°C), then an iodine atlas which has line depths in between the KPNO and the NIST scans would provide a better fit, which the Keck/HIRES scan happens to satisfy.

Why was the temperature of the iodine gas in the cell higher than 70°C? We explore potential cause for the rise of cell temperature in the following subsection.

3.4.4 Cause for a Higher Iodine Gas Temperature

Throughout the FTS scans at NIST and the TS12 experiment, the readings of the temperature probe on the HET/HRS iodine cell stayed at 70°C. However, as we demonstrated above, the temperature of the iodine gas in the cell appears to be higher: 110°C for the NIST FTS scan and 100°C for the TS12 spectrum with the 70°C setting.

One potential explanation is that the temperature probe was malfunctioning. It does not report a true temperature but is biased towards a lower reading than the actual temperature. This has to be true for both the old and the new temperature probes to explain the estimated gas temperatures in both the NIST scans (measured using the old probe) and the TS12 spectra (using the new probe).

An alternative explanation is that the temperature probe was working, but either there was a temperature gradient in the cell, or the gas was at a higher temperature while the glass cell stayed cool. These two alternative scenarios seem plausible, especially as an explanation for the high temperature of the gas for the NIST scans. As recalled by Jason Wright, the halogen lamp used for the NIST scan, which provided the continuum emission shining through the iodine cell, was exceptionally bright. It was a lamp with a power of 1 kW and the room was considerably warmed up by this lamp. We discuss the possibility of these two scenarios below through order-of-magnitude estimates. We use the iodine cell properties for the Keck/HIRES cell from Butler et al. (1996a) for these estimates, because we do not have the data for the HET/HRS cell. The Keck/HIRES cell is a Pyrex glass cylinder that is 10 cm long, and 5 cm in diameter on the windows, with a gas pressure of 0.01 atm at 50°C. We do know that the Keck/HIRES and HET/HRS cells are very similar in size and gas column density, so it should suffice for our order-of-magnitude calculations.

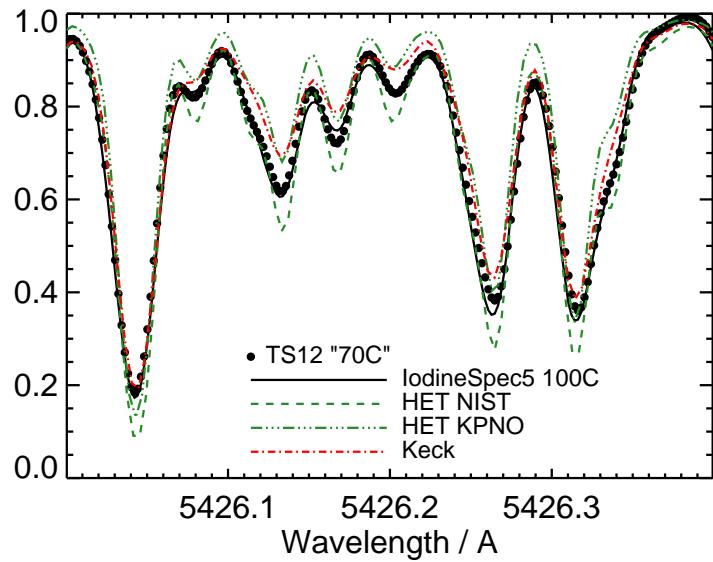
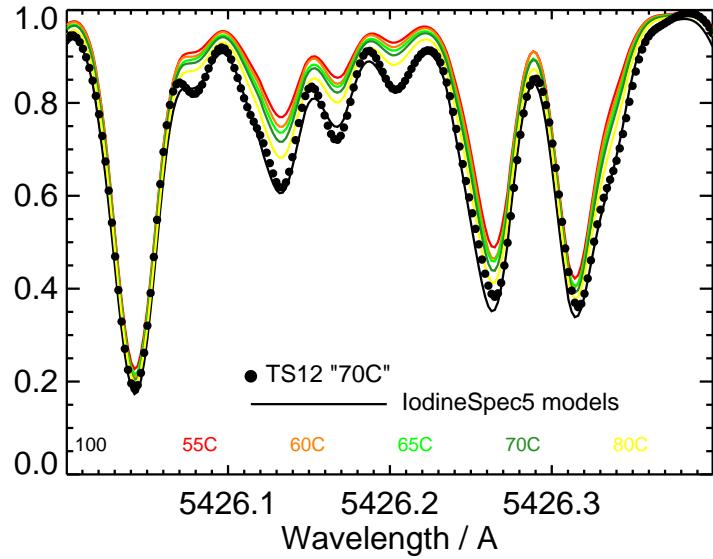


Figure 3.20 Top: Fitting the TS12 spectrum (temperature set at 70°C; black dots) with IodineSpec5 models, with fixed column density derived from best fits using HET/HRS KPNO and NIST scans. **Bottom:** The best-fit temperature for the TS12 spectrum is about 100°C (black solid). It is clearly at a lower temperature than the NIST scan (green dashed) but at a higher one than the KPNO one (green dotted-dashed). Again, all spectra in both panels are convolved down to $R = 200,000$.

3.4.4.1 Temperature Gradient in the Glass

We first see if shining such a bright light at the cell would induce a large enough temperature gradient in the cell glass (made of Pyrex) to explain the difference between the gas temperature (110°C) and the temperature probe reading (70°C, assuming correct). The HET/HRS cell only has one temperature probe, and this is true for the old enclosure used at NIST and the new one at TS12 (the Keck/HIRES cell has two probes: one in the middle and one at the end of the cell). A temperature gradient in the glass cell could mean that the heated glass on one side can bring the gas inside to a higher temperature, while the glass on the other side with the temperature probe stays at 70°C. To sustain a temperature gradient, there needs to be a continuous heat input, which would be the halogen lamp shining on the cell. According to the law of heat conduction, or Fourier's law (in its 1-D differential form), the temperature gradient can be estimated via:

$$q = -k \frac{dT}{dx}, \quad (3.2)$$

where q is the heat flux density in W/m^2 , k is the thermal conductivity of the material in $W/m \cdot K$, and dT/dx is the temperature gradient in K/m .

The heat flux density coming from the halogen lamp can be derived by

$$q = P_{\text{lamp}} \cdot f_{\text{beaming}} \cdot f_{\text{absorption}} / 4\pi D^2, \quad (3.3)$$

where $P_{\text{lamp}} = 1 \text{ kW}$ is the power of the lamp, f_{beaming} is the beaming factor of the lamp (as it is not isotropically radiating towards all directions), $f_{\text{absorption}}$ is the fraction of radiation absorbed by the glass, and D is the distance from the lamp to the cell. For a rough estimate, we assume $f_{\text{beaming}} = 3$ (meaning the light from the lamp is filling a solid angle of $4\pi/3$) and $D = 1 \text{ m}$. To estimate $f_{\text{absorption}}$, we need the emission spectrum of the halogen lamp and the absorption spectrum of Pyrex. We assume the halogen lamp spectrum is a black-body spectrum with peak around $1\mu\text{m}$, which is typical for this type of lamps. The transmission spectrum of Pyrex can be found in this website:

<http://www.me.mtu.edu/microweb/GRAPH/Laser/GLASS.JPG> (or can be easily found at several other places documenting the properties of Pyrex). Pyrex is largely transparent in the optical, but it has an absorption feature near $2.7 \mu\text{m}$ and it heavily absorbs UV light short of 300 nm and IR light longer than $5 \mu\text{m}$. Considering that there was an UV filter between the lamp and the iodine cell at the NIST scan, $f_{\text{absorption}}$ mostly comes from IR. This fraction can be estimated by numerically integrating the Planck function for the intervals $[2.5, 3.0] \mu\text{m}$ and $[5.0, \text{inf}] \mu\text{m}$, and $f_{\text{absorption}}$ turns out to be about 6% of the total lamp emission. From these numbers, q is estimated to be 14.3 W/m^2 .

The thermal conductivity (k) of Pyrex is a well known number, as it is a widely used and manufactured material, and it is about $k = 1.1 \text{ W/m} \cdot K$.⁶

Plugging in q , k , and using $dx = 10 \text{ cm}$ or 0.1 m (the length of the iodine cell), dT turns out to be about 1.4 K . Considering the fact that the light is shining on the window of the glass cell and then being conducted down the thin cylinder shell, q may be boosted by a factor of a few (assuming the thickness of the glass is a couple mm).

⁶From <http://www.azom.com/article.aspx?ArticleID=4765>, for example.

This means that the temperature gradient is unlikely to be as large as 40°C, though not completely ruled out given our crude method of estimation. However, this amount of heat may not be large enough to heat the glass from 70°C to 110°C, or even sustain the high temperature of the glass because of its cooling rate.

First, this amount of energy from the lamp is probably too small to cause a significant temperature rise in the glass. Using the specific heat of Pyrex, $C_P = 0.75 \text{ J/g} \cdot \text{K}$,⁷ assuming the weight of the glass cell to be $M_{\text{glass}} = 100 \text{ g}$, and using an input power of $q \times \pi r^2$, where $r = 0.025 \text{ m}$ is the radius of the iodine cell window, the rate of the temperature rise in Pyrex from heating of the lamp is:

$$dT/dt = q \cdot \pi r^2 / C_P M_{\text{glass}} = 3.8 \times 10^{-4} \text{ K/s.} \quad (3.4)$$

Even assuming the glass does not cool, this means the temperature would only rise for about 4 K in an hour.

Second, the thermal radiation from the glass alone would cool the glass off fast enough. At $T = 100^\circ\text{C} = 375.15\text{K}$, the thermal radiation has a power of $\sigma T^4 = 1123 \text{ W/m}^2$ – much larger than the input energy from the halogen lamp. Even considering that the glass is wrapped in thermal insulators and only two windows are exposed to the air, this should still be enough to cool the glass down fast enough and prevent continuous heating (which would be too slow anyway, as argued above).

Therefore, if the temperature probe is working, it is likely that the entire glass stayed at its reported temperature instead of having a temperature gradient as large as 40°C during the NIST FTS scans.

3.4.4.2 Heated Gas within a Cooler Glass

Next, we look at the temperature of the iodine gas in the cell. If the gas is heated up to a higher temperature without considerably heating up the glass, then perhaps the temperature probe was working but it was only reporting the temperature of the glass but not the gas. First, the small mean free path of the iodine molecule in the cell suggests that they primarily thermalizes with themselves and then with the glass. The mean free path of iodine molecules is estimated by using ideal gas assumption:

$$\lambda = (\sqrt{2}\sigma \cdot n)^{-1} = \frac{k_B T}{\sqrt{2}\pi d^2 P}, \quad (3.5)$$

where σ is the interaction cross section of the molecule, n is the number density of molecules, k_B is the Boltzmann constant, d is the effective diameter of the molecule, and P is the pressure of the gas. The iodine gas in the cell has a pressure of about 0.01 atm or 1 kPa (Butler et al. 1996a), at a temperature of 110°C or 385 K, with an effective diameter of about 5 Å (Juhola 1975; Topley 1926). Using these numbers, the mean free path is estimated to be 4.8 μm, much smaller than the size of the glass cell. This means that the gas is not tenuous enough so that their thermalization is completely dominated by the glass enclosure. The small mean free paths of the iodine molecules does not rule out the possibility that the gas stays at a higher temperature while the glass stays cooler,

⁷From http://www.engineeringtoolbox.com/specific-heat-solids-d_154.html

since the gas molecules primarily thermalize within themselves. But more evidence is needed to support this hypothesis.

First, what is the dominating heating source for the iodine gas? There are three heating sources for the gas during the NIST scan: the heat from the temperature controller heater conducted via the glass, the heat from the halogen lamp absorbed by the glass then conducted to the gas, and the direct radiation from the halogen lamp. According to the calculation above, the glass cell probably stays at 70°C, so it is not heating up the gas to a higher temperature.

Second, for the additional heat received by the glass from the lamp, we have worked out that the amount is 14.3 W/m^2 . This is too small to heat up the glass, but could this small amount of heat be conducted to the iodine gas inside and heat it up? There are four outlets for this received energy: radiative cooling of the glass, thermal conduction to the rest of the glass cell, thermal conduction to the air, and thermal conduction to the iodine gas within. The thermal conductivity of the Pyrex glass is $1.14 \text{ W/m}\cdot\text{K}$, as mentioned before, and the thermal conductivity of the air is about $2 \times 10^{-2} \text{ W/m}\cdot\text{K}$. Both are much larger than the thermal conductivity of the iodine gas, which is on the order of $10^{-4}\text{--}10^{-3} \text{ W/m}\cdot\text{K}$.⁸ This means that the glass will primarily conduct along the glass cell itself, and then a smaller fraction to the air, and then an even smaller fraction to the iodine gas. Considering the factor of 100 – 1000 difference in thermal conductivity between glass/air and iodine gas, probably only a < 1% of the heat input $q = 14.3 \text{ W/m}^2$ is being conducted to the iodine gas, i.e., < 0.14 W/m^2 . This number would be even smaller if thermal radiation of the glass is considered.

Third, another heating source for the gas is the halogen lamp, which is contributing a power of $q \sim 1.7 \text{ W/m}^2$ (or $3.3 \times 10^{-3} \text{ W}$, given the radius of the cell window is 2.5 cm) to the iodine gas. This number is estimated by using the same equation for q as above, but substituting $f_{\text{absorption}}$ with 0.3×0.024 , 0.3 being the approximate fraction of light iodine molecules absorbs in the optical window [500, 650] nm (estimated from normalized iodine spectrum) and 0.024 is the fraction of energy emitted the halogen lamp in this wavelength range. If the gas was heated to beyond 70°C, this would be the primary source of heating. However, this amount of power is probably too small. Again, the thermal radiation of a 100°C material is 1123 W/m^2 , as calculated above, much larger than the 1.7 W/m^2 input power. Plus the iodine gas also cools via band emission (Waser & Wieland 1947). Therefore, it is very unlikely that the iodine gas was actually heated up to 110°C because of the halogen lamp.

Additionally, although the thermalization between the glass and the gas might be slow, it would not be at a longer time scale than the FTS scan (about a few hours). The specific heat of Pyrex is $C_P = 0.75 \text{ J/g}\cdot\text{K}$, and assuming the iodine gas is at $T = 100^\circ\text{C}$ and the glass is at $T = 70^\circ\text{C}$:

$$dT/dt = \frac{(\sigma T_{\text{gas}}^4 - \sigma T_{\text{glass}}^4) \cdot A}{C_P M_{\text{glass}}} = 0.05 \text{ K/s}, \quad (3.6)$$

⁸The estimates for the thermal conductivity of iodine gas comes from Page 36 of Vargaftik (1993), which gives the value of $\sim 4 \times 10^{-3} \text{ W/m}\cdot\text{K}$ for a gas pressure of about 0.2–0.5 atm. The estimate of $10^{-4} \text{ W/m}\cdot\text{K}$ comes from theoretical values assuming ideal gas, $k = n < v > \lambda C_V/3N_A = k_B \sqrt{8RT/\pi M}/\pi d^2 \simeq 10^{-4} \text{ W/m}\cdot\text{K}$, where M is the molar mass of iodine molecule.

where A is the surface area of the iodine cell, on the order of 10^{-2} m^2 . This is probably a lower limit, because the iodine gas would also emit band emission which will be absorbed by the glass (in UV) and the glass would not cool very efficiently as it was wrapped in thermal insulators. Therefore, if the gas was heated up to 100°C , then the glass would be heated up by the gas fast enough to change the reading on the temperature probe.

To summarize the findings above, it is highly unlikely that the glass underneath the temperature probe was at 70°C while the iodine gas inside was at 110°C . This means that a biased temperature probe or a malfunctioning temperature controller is the most likely explanation for the NIST FTS scans and the TS12 spectra. One possible scenarios is that, if the heater of the temperature controller was turned on the entire time, the temperature of the cell and the gas would have kept rising until they had reached an equilibrium, for example, with its thermal radiation. This means the power of the temperature controller heater was $1123 \times 10^{-2} \text{ W} = 11 \text{ W}$, which sounds plausible. This would explain why all three NIST scans, supposedly taken at 65, 70, and 75°C , all look like they were at 110°C .

We believe that the difference between the iodine atlas and reality as a result of gas temperature changes or differences is the most likely reason behind HET/HRS's under-performance in RV precision compared with Keck/HIRES. We outline potential solutions to improve the RV precision of HET/HRS archival data and our recommendations for the new HRS in the conclusion section of this Chapter.

3.5 Conclusions and Future Work

To summarize, we think that two of the major drivers behind HET/HRS's RV systematic errors are temperature issues with the iodine cell (or inaccurate iodine atlas which fails to capture such changes) and inaccurate modeling of the IPs.

Overall, the HET/HRS spectra exhibit more changes than the Keck/HIRES spectra, which probably makes it harder to model and leads to its large RV systematic error (Figure 3.21). These variations may be due to temperature changes in the cell or spectrograph changes such as IP or focus changes. HET/HRS is perhaps more prone to spectrograph focus change due to temperature variation than Keck/HIRES because of the differences in their camera design (the HIRES camera focus was athermalized by the mirror and its spacers in the mount; Vogt et al. 1994; Tull 1998). Precise Doppler spectroscopy requires great care to minimize any changes in the iodine cell and spectrograph performance. For example, standard CPS Keck/HIRES observations are preceded with a series of procedures which ensures that (1) the spectrograph is very well focused and (2) the wavelength solution for the CCD chip stays roughly the same, i.e., the ThAr lines always land on the same pixels. HET/HRS does not (and probably cannot, due to a lack of actuators on optical parts) perform these procedures, to the best of our knowledge. The upgraded HET/HRS will have actuators to enable adoption of the same spectrometer adjustment routine as Keck/HIRES.

Our adventure in improving the RV precision of HET/HRS, though incomplete and sometimes inconclusive so far, have taught us several important lessons for iodine-calibrated precise RV work:

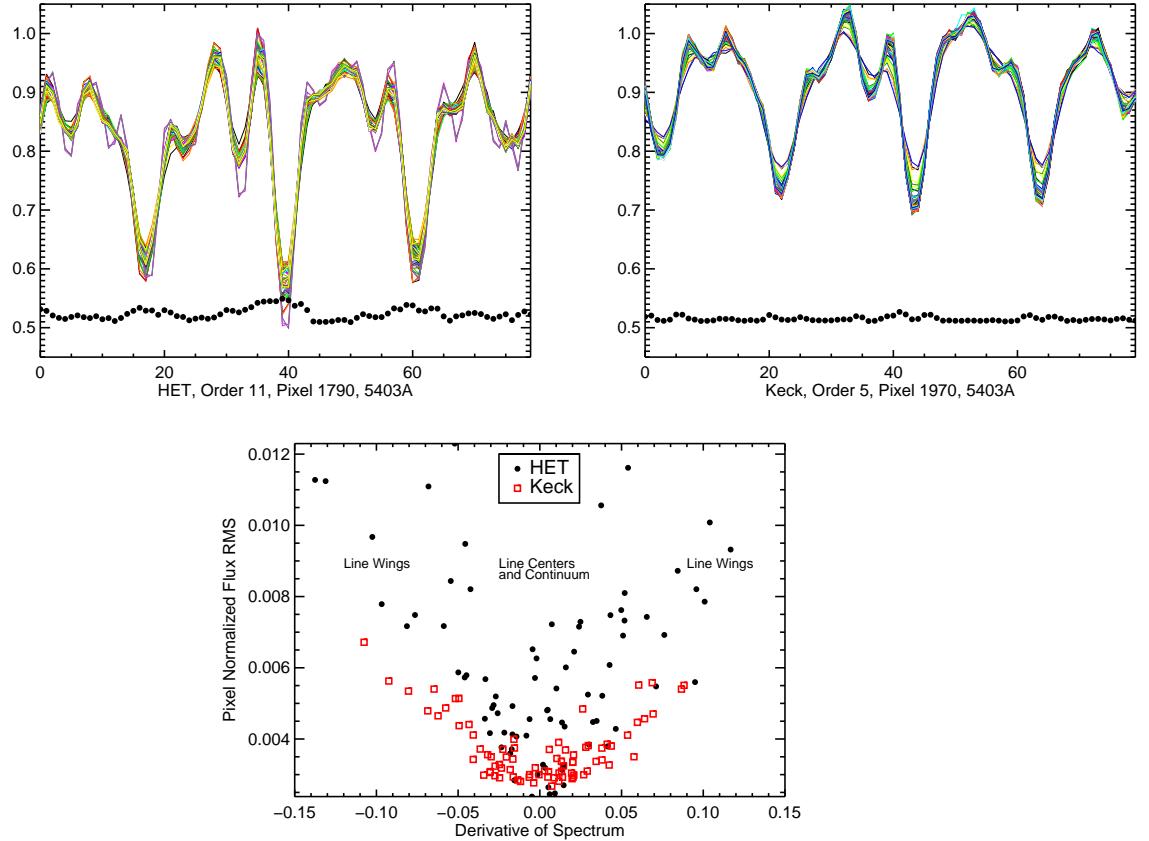


Figure 3.21 Comparison between the amounts of variation in a spectral chunk in ~ 500 HET/HRS and Keck/HIRES lamp iodine observations, respectively. The top two panels show the overplotted spectra in this chunk, shifted and normalized so that the spectral lines match (using cross correlation and interpolation; normalized using total counts in this chunk). The black dots in each of the two top panels are the RMS values of the normalized counts in each pixel, multiplied by 4 and added by 0.5 for illustration purposes. The HET/HRS spectra exhibit significantly more variation than the Keck/HIRES spectra, and we discuss potential reasons in Section 3.5. The bottom panel is plotting the pixel counts RMS as a function of the derivative of the spectrum, which shows that the HET/HRS spectra have more variation in both line centers/continuum and line wings than Keck/HIRES.

- The IP functional forms for fiber-fed spectrographs probably differ quite significantly from slit-fed ones, and it is important to find a good IP function which not only describes the IP well but also has good convergence properties for the forward modeling fitter. Initial guesses on IP parameters could play an important role.
- The iodine atlas provided by FTS scans should not be taken for granted as the “ground truth”. There are challenges with FTS measurements and data reduction for measuring absorption spectrum with such a big spectral span. The cell may also change over time, making older FTS scans obsolete. Echelle spectrograph with hyper resolution like the TS12 arm on the Tull Spectrograph is very helpful for validating iodine atlas or checking cell quality.
- It is vital to stabilize the cell to a desired temperature precisely and accurately, probably at least to $\pm 10^{\circ}\text{C}$.⁹ Theoretical code computing iodine lines such as IodineSpec5 is helpful for diagnosing problems with iodine cells. One can imagine that this can also be helpful in identifying permanent cell change like the one experienced by the Lick/Hamilton cell, which compromised the usefulness of the spectrograph for precise RV work fatally(Fischer et al. 2014).

Looking forward, there are several things which can help solving the mystery and hopefully eventually leading to an improvement in the RV precision of HET/HRS:

(1) Determining cell temperature before extracting RVs for each observation.

It is probably valuable to obtain FTS scans of the HET/HRS cell at various temperatures and interpolate between them to make a finer model grids of iodine atlases at different temperatures. Floating the temperature as a free parameter in RV extraction is also an option, but the degeneracy it introduces will be detrimental to the RV precision. Therefore, a better route is to use the temperature-sensitive regions among the iodine spectrum to determine the cell temperature first, then use the corresponding atlas in the forward modeling process to determine RVs. Validating the iodine atlas and temperature control stability/reliability¹⁰ is crucial for the upgraded HET/HRS.

(2) Finding a better IP function: This can be first done through fitting ThAr frames, and then tested via fitting B star + iodine frames, which would be a much easier task for observations with pre-determined iodine cell temperature. We plan to continue to explore possibilities with the modified Moffat function by adding perturbations and providing good initial guesses.

(3) Examining the spectral data for modal noise or raw reduction errors:

We have taken day time engineering data before the HET shutdown/upgrade to test if modal noise is significant in HET/HRS data in very high SNR regime. Although we have calculated that the number of modes in the HET/HRS fiber is large enough that modal noise should not be a concern, it would still be valuable to reduce and analyze these data to validate this calculation.

⁹The importance of stabilizing iodine cell temperature is demonstrated in a poster presented at the 2016 June SPIE conference lead by U. Seeman et al. from Georg-August Universität Göttingen.

¹⁰The cell enclosure and temperature controller for the HET/HRS used in our TS12 observations are the ones for the new HET/HRS, which is troubling.

More importantly, the upgraded HET/HRS has great promise in RV precision, and carrying on these lessons is crucial for ensuring good RV performance (~ 1 m/s) – especially the lesson on iodine atlas and cell temperatures. Chapter 7 has more on the upgraded HET/HRS and related future work.

Chapter 4

Improving the Radial Velocity Precision of Keck/HIRES

4.1 Introduction and Background

The Keck telescopes are two 10-m telescopes on the mountain of Mauna Kea in Hawaii at the Keck Observatory, which is operated by the University of California and financially supported by the W. M. Keck Foundation. Precise Doppler observations are made through the HIRES spectrometer (Vogt et al. 1994) on the Keck I telescope, which has a long and glorious history in exoplanet discoveries (see Chapter 1). Just like HET/HRS, the precise RV observations are calibrated by an iodine cell. The typical resolution for RV frames is chosen to be $R = 55,000$ with the B5 decker on the slit plate, which is also tall enough to provide good sky subtraction (for more details, see Howard et al. 2009 and Valenti et al. 2009). Keck/HIRES is known as the most precise RV instrument on a 10-meter class telescope, with a stunning long-term RV precision of 1-2 m/s.

Penn State does not have any private share of Keck/HIRES time, but we are part of the California Planet Survey group which regularly conducts RV surveys using Keck/HIRES for about 30 nights a year or more (especially for *Kepler* follow-up programs). As a result, we have made joint discoveries using Keck/HIRES and HET/HRS RVs (e.g., Chapter 6) and have access to all the data and extracted RVs on the standard stars such as HD 185144 and HD 10700.

My work on Keck/HIRES started by accident. As we were investigating the causes behind HET/HRS’s RV systematics, we regularly compared Keck/HIRES performance on HD 185144 or iodine frames with HET/HRS. When we were looking into the quality of HET/HRS’s iodine atlas and studying the potential effects of its “wrong” wavelength solution, I spotted a strong correlation between the RVs reported by a spectral chunk and its wavelength in the Keck/HIRES data. The chunk RVs and their wavelengths seem to have a linear correlation whose slopes varied from observation to observation but with a zero point all anchored at around 5200Å, with the redder region near 6000Å exhibiting the largest scatter. My thesis advisor, Jason Wright, quickly suggested that this could be due to telluric lines in the spectrum, because there are more telluric lines in the red and the 5200Å region is relatively free of tellurics. Indeed, we saw strong correlations

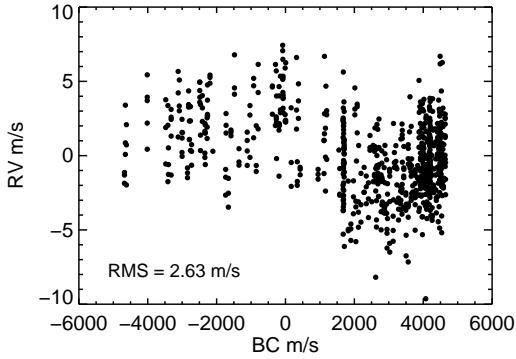


Figure 4.1 Keck/HIRES RV vs. BC for HD 185144, an RV standard star. The strong correlation between velocities and the barycentric velocity of the star indicates RV systematic errors that are not intrinsic to the star. This is because the barycentric motion of the Earth is the dominant source of redshifts of stellar lines as Earth moves at a large speed of up to 30 km/s relative to the star. As the stellar lines move back and forth annually through the iodine lines, any systematic errors will manifest as correlations between RV and BC.

between RVs and the barycentric velocity correction (BCs) – a smoking gun signature for RV systematics – for telluric contaminated chunks. Perhaps this is the culprit behind the long-known RV-BC correlation in the data of HD 185144 (Figure 4.1)? Hoping that eliminating this (more or less) long ignored spectral contamination source would bring visible improvements to Keck/HIRES RV precision, we began our investigation on the effects of telluric contamination in Keck/HIRES data, which subsequently lead to the other works in this Chapter.¹

The data we used for this Chapter are kindly provided by the CPS group, in particular, Andrew Howard and Howard Isaacson.

4.2 Effects of Telluric Contamination and Remedies

4.2.1 Introduction

Traditionally, telluric contamination is not considered to be damaging to the RV precision in the optical. It is certainly a severe source of spectral contamination and a bottleneck for achieving higher RV precision in the near infra-red (NIR) region (e.g., Bean et al.

¹As it later turned out, the correlation between the chunk RVs and wavelengths are neither due to telluric contamination nor the wavelength errors in the iodine atlas (though it could have such an effect, but hard to imagine at the magnitude we saw). It is purely because that I had chosen to perform a weighted least- χ^2 fit for the linear correlation between RVs and wavelengths, but I had used the wrong weights. Instead of using the weights empirically derived by vanking, I was using the analytical weights calculated following Butler et al. (1996a). Because there is no such real correlation between RVs and wavelengths, the “correlation” I saw or fitted for was merely reflecting the facts that there are more stellar lines in some region of the spectrum than others, and the region near 5200Å receives the highest weight.

2010), where a large number of deep water and methane lines reside. However, there is only a small wavelength range in the optical that has deep telluric lines, and typically such regions are simply thrown out for the purpose of precise RV analysis, either by giving them zero weights in the cross correlation masks (for ThAr calibrated spectra, e.g., Pepe et al. 2002a) or flagging them as bad pixels (for iodine calibrated spectra, e.g., for Keck/HIRES).

Recently, the works by Artigau et al. (2014) and Cunha et al. (2014) have characterized and mitigated the effects of telluric contamination in the precise RV data taken by the ThAr-calibrated HARPS-S. Cunha et al. (2014) focuses on the issues with “micro-telluric” lines (shallow telluric absorption lines with < 1-3% depths; Figure 4.2), which are recognized for the first time there as being important to precise RV measurements. Cunha et al. (2014) fit and then divide out the telluric lines in the observed spectra using synthetic telluric spectra generated by the LBLRTM package (Line-By-Line Radiative Transfer Model, Clough et al. 1992; with line lists from HIgh-resolution TRANsmision molecular absorption database, or HITRAN, Rothman et al. 2013) and also TAPAS (Bertaux et al. 2014), which is a more user-friendly but less flexible package wrapper using LBLRTM. They concluded that the micro-tellurics have an impact (defined as the root-mean-square, RMS, of difference between RVs before and after micro-telluric removal) of \sim 10-20 cm/s for G stars observed with low to moderate air masses, but the impact can be substantial in some cases to up to \sim 0.5-1 m/s.

Artigau et al. (2014) use principal component analysis (PCA) to empirically correct for telluric lines in HARPS-S data (both micro-tellurics and the deep lines in the \sim 630 nm region), and combined PCA with rejection masking, they reduced the RV RMS by \sim 20 cm/s (and more significantly for the \sim 630 nm region). More recently, Sithajan et al. (2016) characterized the effects of telluric contamination and effectiveness of some typical remedies (masking and modeling) for emission line-calibrated spectra for the optical, broad optical (300-900 nm), and NIR. Their conclusion for the optical region is similar to the results in Artigau et al. (2014) and Cunha et al. (2014).

This section characterizes and corrects for the adverse effects of telluric contamination under the context of iodine-calibrated precise RV, especially for the micro-telluric lines. We first quantify the effects of tellurics in RV precision and accuracy through simulations (Section 4.2.2), and then we talk about remedies and their effectiveness to eliminate the adverse effects in real observed data (Section 4.2.3). We summarize our recommendations for treating telluric contamination in iodine-calibrated RV data in Section 4.2.4.

4.2.2 Characterization of the Impacts of Micro-tellurics on RV Precision via Simulation

To evaluate the impacts of micro-tellurics (referred to often simply as “tellurics” below), we performed end-to-end simulation of Keck/HIRES data and analysis process on RV standard stars in order to isolate error sources. We use Keck/HIRES data to for our study because Keck has the highest RV precision among all iodine-calibrated spectrometers, and it also has long observing baselines on a number of RV standard stars. RV standard stars are bright and quiet stars which do not host known planets, and thus exhibit the smallest RV variation in both short term and long term. Their data are often good

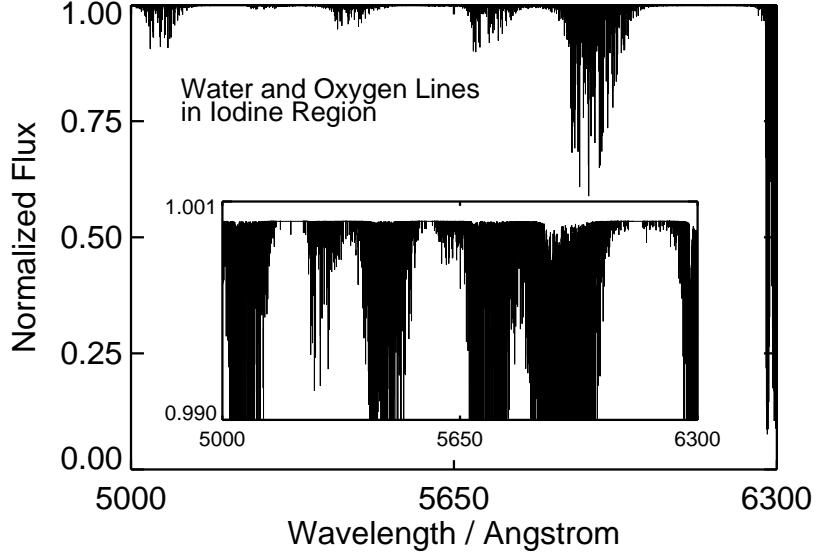


Figure 4.2 Telluric lines in the iodine region are mostly shallow water lines, with some moderately deep water lines near 5900Å and very deep oxygen lines near 6300Å. The insert plot is showing the pervasiveness of micro-telluric lines, i.e. $\leq 1\text{--}3\%$ in depths.

diagnostic tools for identifying RV systematics. For our study, we used and simulated Keck/HIRES RV spectra on two standard stars, σ Draconis (HD 185144) and τ Ceti (HD 10700), which are benchmark classics in precise RV work.

HD 185144 (spectral type G9V, per Simbad) has 712 Keck/HIRES observations, with RV RMS = 2.57 m/s, and it has a relatively small barycentric velocity (often referred to as the barycentric velocity correction, or BC; see Chapter 2) span, $[-4.7, 4.6]$ km/s, because it is near the north ecliptic pole. HD 10700 (spectral type G8.5V) has 623 observations, with RMS = 3.05 m/s, and its BC span is $[-27.8, 26.8]$ km/s. The RV RMS numbers quoted here come from reductions using our version of CPS Doppler pipeline, and they are larger than the RMS values from the most up-to-date CPS pipeline due to some recent improvements in the CPS version. The most recent CPS inventory (as of April 2016) also has a few new observations on these two stars.

We describe how we simulated our data and used them to characterize the impacts of tellurics in Section 4.2.2.1, and then lay out our results in Section 4.2.2.2.

4.2.2.1 Methodology of the Simulation

The overall strategy is to bring out the effects of telluric contamination by comparing the extracted RVs from pairs of simulated observations: one with telluric absorption, and the other one in the pair without. For example, for HD 185144, we simulated two spectra for each of the 712 real observed spectrum. We then had two sets of simulated observations, one free of tellurics and one contaminated with tellurics. Next we ran the Doppler code on each set of 712 spectra to extract RVs from them, and for the set with tellurics

we used telluric-contaminated DSST and a telluric-free DSST for the telluric-free spectra.

Below is our recipe for simulating Keck/HIRES spectra and DSSTs:

1. Generate a synthetic spectrum for the target star using Spectroscopy Made Easy (SME; Valenti & Piskunov 1996; Valenti & Fischer 2005; the SME spectra are kindly provided by Jason Curtis). We refer to this spectrum as the input synthetic spectrum or the SME spectrum.
2. Based on the SME spectrum, generate a deconvolved stellar spectral template (DSST) for the target star. First, we redshift the SME spectrum so that it has the same barycentric velocity as the real DSST of the target star. Then we break the SME spectrum into chunks as defined by the real DSST (an IDL structure) of this target star.
3. Run TERRASPEC (Bender et al. 2012), a versatile software for synthesizing telluric spectrum (it is a wrapper around the HITRAN line list (Rothman et al. 2013) and the LBLRTM package (Clough et al. 1992), developed by Chad Bender, who provided crucial help in this work for using TERRASPEC), to generate a telluric absorption spectrum using a typical Mauna Kea atmospheric condition in the summer, with precipitable water vapor (pwv) 1 mm (a little bit more humid than Keck’s usual condition), and oxygen column density that is consistent with the altitude of the real DSST observation. We multiply this synthetic telluric spectrum with the SME spectrum to make a telluric-contaminated DSST.
4. Loop through each real observed spectrum of the target star, and simulate a pair of observations following this recipe:
 - (a) Shift the SME spectrum to have the same barycentric velocity as the star at the epoch of the real observation (i.e., assuming intrinsic stellar velocity is zero).
 - (b) Generate a continuous wavelength solution for each echelle order based on the best-fit wavelength solution for the real spectrum.
 - (c) Re-sample the SME spectrum and the iodine atlas onto a wavelength grid that is four times finer than the wavelength solution grid. Multiply the stellar spectrum with the iodine atlas.
 - (d) To add telluric contamination: multiply the simulated spectrum with a telluric absorption spectrum generated with TERRASPEC, again for typical Mauna Kea summer, with pwv = 1 mm, and oxygen column density that is consistent with the altitude of the real observation.
 - (e) Determine the IP for each echelle order with two options: (1) a single Gaussian with $\sigma = 1.7$ pixel, matching the Keck/HIRES resolution; (2) a sum-of-Gaussian IP derived using the average IPs for each echelle order. We will explain these two options later in the text.
 - (f) Convolve each order of the stellar×iodine spectrum with the IP.

- (g) Fit for the blaze function for each order of the real observed spectrum with a third-order polynomial using the top 1% of the flux, and multiply the simulated spectrum so that they have matching “blaze functions”. This also ensures that the simulated spectrum has the same photon counts as the real one.
 - (h) Add photon noise according to Poisson statistics, when desired.
5. Generate a separate data log file for each set of simulated observations under the same conditions (i.e., telluric choice, IP choice, photon-noise choice).

we first simulated spectra free of photon noise and with a simple single Gaussian IP for the spectrograph ($\sigma = 1.7$ pixels to match the Keck/HIRES resolution). When extracting RVs for such sets of simulated spectra, we fixed the IPs to the same as input and only fitted for three free parameters. This removes the errors induced by photon noise and uncertainties in IP fitting (which, after all, involves 12 additional free parameters). We then added the Poisson noise to the spectra, but still using simple and fixed IPs, to see how the effects will change in cases with limited-SNR (as opposed to infinite SNR). Next, we used complex IPs, the same ones as the best-fit sum-of-Gaussian IPs found for real spectra, and float the IP parameters when extracting RVs. For the simulated spectra with complex IPs, we extracted their RVs following exactly the same procedure for real observations, e.g., using IP initial guesses from solutions for real neighboring B star + iodine observations. The conditions and results of these three pairs of simulations are listed in Table 4.1, and also illustrated in Figure 4.3. These results show the impacts of telluric contamination on Keck/HIRES data, which is discussed in the following subsection.

4.2.2.2 Results of the Simulation

Figure 4.3 illustrates the effects of telluric contamination under different conditions. The top panels are RV differences between the simulated spectrum pairs with infinite SNR and perfectly known IPs, which might somewhat represent the capability of future precise RV instruments. These two plots highlight the net effects of telluric contamination: they cause large biases in RV estimates, which manifest as strong RV-BC correlations or trends. Clearly, the current statistical weighting or “vanking” procedure in the Doppler code is unable to fully correct these biases.

Interestingly, when we added in photon noise (middle panels of Figure 4.3), the effects of telluric contamination are washed out significantly. This is probably because photon noise induces a much bigger RV scatter than the micro-telluric lines. This is not surprising, given that the depths of these micro-telluric lines are often comparable to noise level. We are still investigating the exact reason of this “wash out”.

When using complex IPs and fitting for IPs in RV extraction (bottom panels of Figure 4.3), the effects of telluric contamination are further washed out, as we are adding in more errors that are perhaps larger than telluric-induced ones. We can barely see any RV-BC correlation in the case of HD 185144 and only some in HD 10700.

The RV-BC correlations shown in Figure 4.3 will translate into spurious peaks in the periodograms of the target star, and such spurious peaks will have periods around

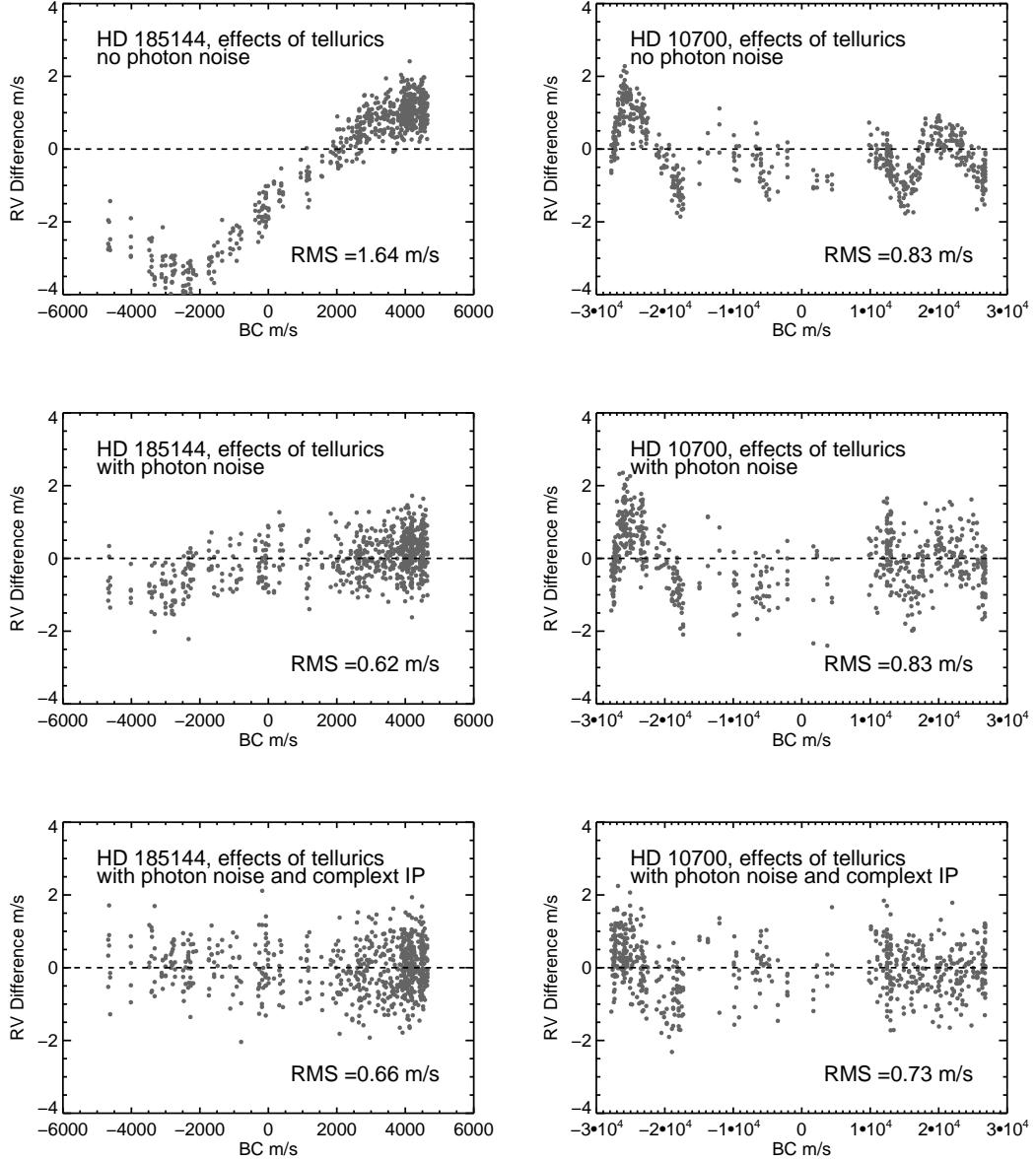


Figure 4.3 Effects of telluric lines manifested as correlation between RV and BC. Each point represents the difference in RV estimates for a pair of simulated spectra: one without telluric absorption, and one with telluric absorption on top of the stellar and iodine spectra. **Top 2 panels:** To isolate the effects of telluric lines, the simulated spectra used for this plot do not have Poisson noise added, and they have simple one-component Gaussian IPs which have fixed width and thus the IP parameters are all fixed to the true values in the RV extraction. **Middle 2 panels:** same as the top panels, but for simulated spectra with Poisson noise (same noise for the telluric and non-telluric spectrum pairs; and still the same simple IPs). **Bottom 2 panels:** same as above, but for simulated spectra with Poisson noise and complex IPs that are similar to the ones in actual observations. IP parameters are not fixed in this case, so the code is fitting 12 additional parameters for the IP on top of the 3 for wavelength solution and Doppler shift (see Chapter 2 for more details on the code).

Table 4.1. List of Simulations

Simulation Conditions	Tellurics?	185144 RMS	10700 RMS	More Details in
No photon noise, simple, fixed IPs	No	0.62 m/s	1.95 m/s ^c	Section 4.2.2.2
	Yes	1.21 m/s	2.42 m/s ^c	top panels, Figure 4.3
Poisson noise, simple, fixed IPs ^a	No	2.11 m/s ^c	2.54 m/s ^c	Section 4.2.2.2
	Yes	2.12 m/s ^c	2.81 m/s ^c	middle panels, Figure 4.3
Poisson noise, complex, floating IPs, ^a	No	1.22 m/s	1.35 m/s	Section 4.2.2.2
	Yes	1.29 m/s	1.44 m/s	bottom panels, Figure 4.3
Poisson noise, complex, floating IPs, masking the telluric pixels ^a	No	1.27 m/s	1.43 m/s	Section 4.2.3.1
	Yes	1.27 m/s	1.43 m/s	Figure 4.4 Table 4.2
Poisson noise, complex, floating IPs, modeling tellurics ^b	pwv = 0.5, 0.9 pwv = 1.0, 1.1	1.28, 1.22 m/s 1.25, 1.23 m/s	1.32, 1.30 m/s 1.33, 1.31 m/s	Section 4.2.3.2 Figure 4.5

^aWith pairs of simulated spectra with or without tellurics and each pair has the same Poisson noise added.

^bTelluric lines added with pwv = 1 mm, and we extracted RVs using pwv = 0.5, 0.9, 1.0, 1.1 mm to access how accurately/precisely one needs to model the tellurics.

^cThese simulations suffer from severe numerical and algorithmic errors for reasons not yet fully understood.

a sidereal year and its harmonics. This is very damaging for the search of exoplanets in the Habitable Zone around Sun-like stars, which would have periods around 360 days. Besides these spurious signals, telluric contamination also causes increases in RV RMS. In all cases of telluric-free and telluric-contaminated simulation pairs, the RV RMS is larger for the telluric-contaminated case (Table 4.1). The RMS numbers written in Figure 4.3 represent the amount of RV RMS added (in quadrature) because of telluric contamination (again, the spectrum pairs are same-noise pairs – they only differ in terms of having or not having telluric lines). These numbers also represent an RV RMS floor set by the telluric contamination: for example, if we ignore tellurics in RV reduction, one would not achieve an RV precision of <0.6-0.7 m/s on G type stars using iodine cells as calibrators. This would compromise our ability to detect Earth or super-Earth exoplanets greatly (again, Earth 2.0 would have an RV amplitude of 0.08 m/s). In the next section, we discuss strategies to eliminate these adverse effects of telluric contamination for iodine-calibrated precise RV data.

4.2.3 Remedies and Effectiveness

There are several ways to remedy the adverse effects of telluric lines on RV precision and accuracy: masking, modeling, or a combination of both. Like the rest of this section, we focus our efforts on iodine-calibrated data. For ThAr-calibrated data, the current official HARPS pipeline masks deep telluric lines and abandons spectral orders that are heavily contaminated (Xavier Dumusque, private communication; Artigau et al. 2014). Sithajan et al. (2016) also discusses remedies and their effectiveness for ThAr-calibrated

Table 4.2. Comparison of RV RMS between Simulations with or without Masking

HD 185144	HD 10700	Simulation Conditions
1.22 m/s	1.35 m/s	No tellurics
1.29 m/s	1.44 m/s	With tellurics
1.27 m/s	1.43 m/s	No tellurics, but masking telluric pixels
1.27 m/s	1.43 m/s	With tellurics, and masking telluric pixels

data in more band options including the near infrared.

4.2.3.1 Masking is an ineffective solution

The simplest solution is to mask out telluric lines in the spectrum, which means, in practice, locate the telluric-contaminated pixels and flag them as bad pixels in the observed spectrum so that the least- χ^2 fitter will ignore them. For Keck/HIRES or any iodine-calibrated RV reduction, this also means masking out the regions corresponding to locations of telluric lines in the deconvolved stellar reference spectrum – because the stellar reference spectrum was taken at a different BC, the telluric lines therein are shifted with respect to the ones in the epoch observation as we try to match up the stellar lines in observed and reference spectra. This double masking procedure is illustrated in Figure 4.4. This is done dynamically in the fitting process, in the sense that, for each iteration in the least- χ^2 minimization process, the contaminated pixels are located according to the current wavelength solution parameters in this fitting iteration. The wavelength solution changes from iteration to iteration, and thus the masked pixels can change too.

To investigate the effectiveness of masking, we performed RV extraction on simulated spectra with or without telluric lines injected and with or without masking (all with Poisson noise and complex IP to mimic real observations as much as possible). For stellar reference spectrum, we used the synthetic spectrum with telluric lines. The results are already listed in Table 4.1, but are listed again in Table 4.2 for clear comparisons. In terms of improving RV precision or reducing RV RMS, masking is very ineffective. The additional errors it introduces diminish its merits. On the other hand, masking does improve the accuracy to some degree: for example, masking does remove the downward RV trend seen in HD 10700 data on the bottom right plot of Figure 4.3 in the BC range $[-3 \times 10^4, -2 \times 10^4]$ m/s. However, masking is an ineffective way to mitigate the effects of telluric contamination overall, especially since the RV errors and RV-BC trends are dominated by photon noise and algorithmic errors (and other types of errors too in real observations).

masking does not work? First of all, it complicates the χ^2 surface and “breaks” the L-M fitter. Due to the dynamic nature of the mask mentioned above, the degrees of freedom for fitting could change, because some telluric lines may shift in and out of this spectral chunk as the wavelength solution changes. This would make the fitter harder

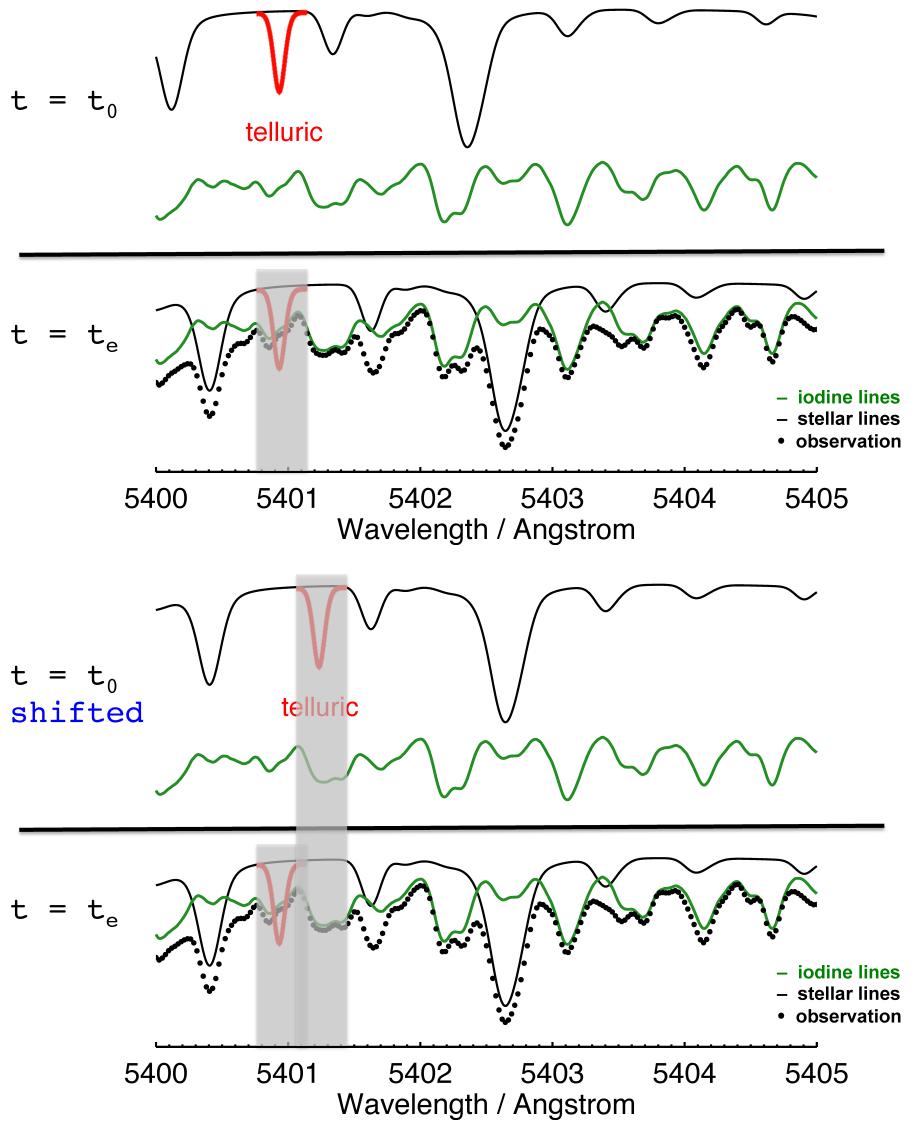


Figure 4.4 Illustration for how we mask telluric contaminated pixels. The top panel shows how we mask the telluric lines (red solid lines) in the epoch observation taken at epoch $t = t_e$. The bottom panel shows why we also need to mask pixels associated with telluric lines in the deconvolved stellar reference spectrum taken at epoch $t = t_0$ and being shifted in order to model the observation.

to converge or may create more local minima in which the fitter can get stuck, causing additional errors. A better fitter might solve the convergence problem, but it is hard to justify throwing away useful iodine and stellar content embedded in the masked pixels. Finally, to mask the telluric lines out, one needs to pick a flux threshold for the masks. This threshold must maintain a balance between masking too much (throwing away too much iodine and stellar information) and too little (leaving shallow telluric lines and line wings untreated). In our study, we have chosen a flux threshold of 0.3%, which means any pixel with telluric absorption deeper than 0.3% will be masked (reference telluric spectrum is generated by TERRASPEC at an zenith angel of 70° , meaning deep oxygen lines, and with $\text{pwv} = 0.8 \text{ mm}$, a little more than typical Keck/HIRES humidity). This masks 11% of the spectral domain, which is quite substantial and is very damaging to the RV precision, but is almost the minimal amount of masking required to achieve some RV accuracy improvement as we found out by testing various masking depths.

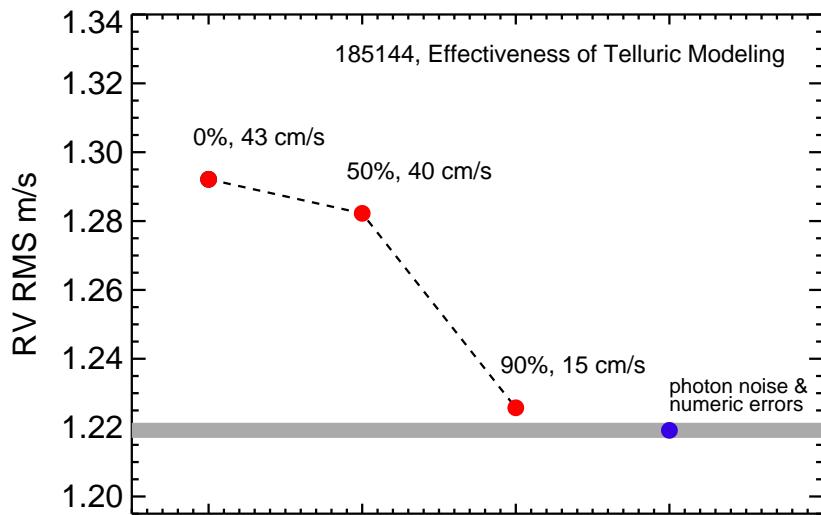
We also applied telluric masking in RV reduction for real observations, and saw no improvement over RV precision or accuracy. This is because other effects dominate rather than tellurics, as mentioned above, such as photon and algorithmic errors and especially deconvolution errors in stellar reference spectrum, which are discussed in Section 4.3 and 4.4.

To summarize, masking sounds like a simple solution to the problem of micro-telluric contamination, but it is actually complicated to implement (for iodine-calibrated RV reduction) and it is ineffective in terms of improving RV precision. We do not recommend masking as a remedy for treating micro-telluric lines in iodine-calibrated RV work. We believe the most effective way is to forward model telluric lines, and combine that with some “masking” for deep or troublesome telluric lines, which we discuss in the next subsection.

4.2.3.2 How precisely does one need to model the tellurics?

The other way is to incorporate telluric lines as part of the iodine RV forward modeling process, where water column density can either be from a priori knowledge or an additional free parameter. In principle, the oxygen column density can also be a free parameter, not because the amount of oxygen varies on a noticeable level, but just to allow some compensation for errors in atmospheric temperature and pressure profile and so on, in order to model the oxygen lines more precisely. We do not fit for oxygen in our simulation or treatment for real observations in this work for simplicity, and also because the chunks contaminated with oxygen lines are in the reddest part near 6300\AA , where the amount of iodine and stellar contents are minimal anyway, and these chunks tend to be thrown away or heavily de-weighted in the final RV weighting process.

Modeling telluric absorption lines to high precision (below 1–2% RMS residual) can be a challenging task. There are several reasons for this: lab measurements of a large number of water lines are inaccurate, in terms of line depths, line positions, and line shapes; and these line properties can also be uncertain due to change or a lack of knowledge of the atmospheric conditions, such as wind, high line-of-sight variations (e.g., water vapor), and mixing uncertainties. For a summary of the state of the problem and paths forward recommended by the RV community, see Section 4.6 in Fischer et al. (2016). However,



Varying Levels of Telluric Correction

Figure 4.5 Improvements in RV RMS for different “level” of telluric modeling/removal. For example, the mid point labeled with “50%, 40 cm/s” means that if you model your telluric absorption lines to 50% of their original depths, the effects of the residual telluric absorption will add 40 cm/s in quadrature to your final RV RMS. The blue point marks the RV RMS for simulations with Poisson noise and complex IP on HD 185144, which represents the photon-limited RV precision (subject to additional numeric or algorithmic errors; see Chapter 7 for more on the limitation of the Doppler code).

the goal here is not to model or remove the telluric lines perfectly, but to mitigate their impact on RV precision and accuracy as much as possible. A central question is: how well do we need to model telluric lines to reach a certain RV precision (Fischer et al. 2016)?

To answer this question under the context of iodine-calibrated RV, we performed RV extractions on the simulated HD 185144 data with telluric absorption (all with $\text{pwv} = 1.0 \text{ mm}$, as described in Section 4.2.2.1), incorporating forward modeling of telluric lines with different levels of accuracy and using a stellar reference spectrum free of tellurics. The results are illustrated in Figure 4.5. All three simulations were run with simulated spectra of HD 185144 with $\text{pwv} 1.0 \text{ mm}$, but the one labeled “0%” has no telluric modeling in the RV extraction, while the one labeled with “50%” has synthetic telluric lines with $\text{pwv} 0.5 \text{ mm}$ in the forward modeling process, and “90%” meaning using telluric model with $\text{pwv} 0.9 \text{ mm}$. In addition, we also used telluric model with $\text{pwv} 1.0 \text{ mm}$ and 1.1 mm in the forward modeling, which produced similar RV RMS as the “90%” simulation with no visible RV-BC trends or correlations. A $\sim 90\%$ modeling accuracy for the water lines would control the RV RMS contribution from tellurics to below 20 cm/s , which is near the target precision for the next generation RV spectrographs. This level of modeling precision is very easy to achieve in reality. However, it is worrisome that even 100% modeling with $\text{pwv} 1.0 \text{ mm}$ cannot eliminate the increase in RV RMS completely. To improve beyond 20 cm/s , we think that a next-generation Doppler code is needed.

One important point to notice is that the reason why the damage of 10% telluric modeling residual is controlled down to $\leq 20 \text{ cm/s}$ is the additional masking and weighting process in the Doppler code, i.e., “vanking” (Chapter 2). A combination of modeling (even only to 90% precision) and statistical weighting can effectively control the RV RMS introduced by tellurics to $\leq 20 \text{ cm/s}$. Weighting plays a role in telluric contamination remedy because it is essentially performing some masking on the chunks that are badly contaminated by tellurics and/or have large modeling residuals, such as the ones near 6300\AA with deep oxygen lines and little stellar or iodine content. Chunks with deep and numerous oxygen lines are normally thrown out completely, and other contaminated chunks which suffer from low precision will receive lower weights and thus cast a lower impact on the final precision and accuracy. In reality, we are using a combination of modeling and masking or weighting to tackle problem of telluric contamination, which we believe is the optimal solution for iodine-calibrated RVs.

4.2.3.3 Treating telluric contamination in real observations

The situation is much more complicated for real observations, because other noise sources enter the picture, some uncontrollable and/or unknown. It turns out that telluric contamination is not the dominant sources of RV systematic errors in Keck/HIRES data. We later found out that the major culprit behind the RV-BC correlation patterns is probably errors in the DSSTs, which are discussed in Section 4.3. This fact makes it somewhat difficult to assess the effectiveness of telluric treatment on real data, but the simulations have demonstrated that the best and most effective strategy is to have a clean DSST and to forward model the telluric lines (and also to mask out deep lines and de-weight chunks with hard-to-clean tellurics, which are already taken care of by

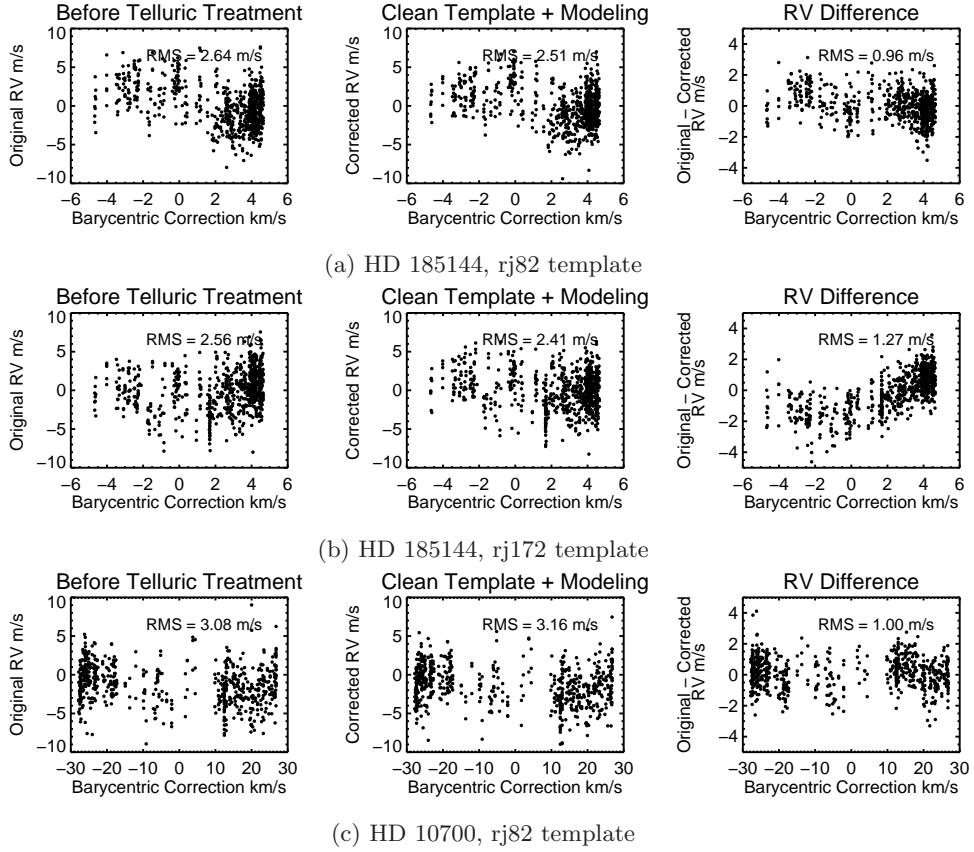


Figure 4.6 Effects of using clean DSST and preliminary telluric modeling on RV precision and accuracy, for HD 185144 (top two rows) and HD 10700 (bottom row). There are two sets of data for HD 185144 because we ran the Doppler code with two DSSTs derived from spectra taken at two different epochs (i.e., at different BCs). The left and middle panels are showing RVs vs. BCs for RV extractions with and without telluric treatment, respectively. The right panels show the RV difference between the two panels.

vanking).

Figure 4.6 represents our best efforts so far and their effects. Our treatment plan includes using “cleaned” DSSTs (we describe how we cleaned the DSSTs later in this subsection) and preliminary modeling of telluric lines in the Doppler code, where we used $\text{pwv} = 1.0 \text{ mm}$ for all observations instead of fitting pwv for each one. In the case of HD 185144, the RV RMS values decreased after we treated the tellurics, but it is not the case for HD 10700 – we suspect that the difference comes from the choice of the pwv value of 1 mm, which is more typical for HD 185144 observations but perhaps not so for HD 10700 (these two stars are at very different Declinations). We plan to model the water absorption in each observation using best-fit or best-estimated pwv in the near future.

How about telluric lines in the DSST? For simulations, we have the privilege of using the synthetic spectrum which is naturally telluric-free. For real observations, telluric contamination enters every observed spectrum associated with the making of DSST (for a detailed description of how the DSST is made, see Section 2.2.1). To remove errors in the DSST caused by the telluric contamination, we formulated a recipe to “clean up” the DSST:

1. When fitting the bracketing B star + iodine observations, we incorporate synthetic telluric lines generated by TERRASPEC as one of the input reference spectra (just like the iodine atlas). The water column density, or pwv , is determined by fitting two rich water bands among the spectrum taken by the red chip data of Keck/HIRES (7000-8000Å). Although we use Mauna Kea’s typical atmospheric condition and adjust oxygen column density according to observation altitude, we still fit for oxygen column density to allow some room for errors in the model (the adjustment needed is typically <0.5%).
2. Using the IPs and wavelength solution derived in step (1), we run the deconvolution algorithm and generate a DSST, which would have telluric lines in it because telluric lines exist in the stellar (iodine-free) spectra.
3. We divide out the telluric lines in the DSST in two steps:
 - (a) Fitting two parameters to make sure the telluric lines match the ones in DSST: broadening parameter (width of a Gaussian kernel) and wavelength shift. The broadening parameter is necessary because the synthetic telluric lines are at an extremely high resolution, while the telluric lines in the DSST are deconvolved but are typically lower. The wavelength shift is needed because the wavelengths of the telluric lines, provided by the wavelength solution of the DSST, could differ from rest-frame telluric line wavelengths because of errors in DSST or changing atmospheric conditions such as wind.
 - (b) Fitting these two parameters is not a straightforward least- χ^2 problem, because stellar lines are also present in the spectrum and we do not have an input model for them. Thus, instead of minimizing residuals between model (broadened and shifted telluric lines) and data (deconvolved telluric lines plus stellar lines), we minimize the flux above the continuum level to optimize the

parameters and the telluric model. In the future, we hope to incorporate this into the code for making DSSTs (see Chapter 2), i.e., fitting for telluric lines and stellar lines at the same time.

- (c) We divide out the best-fit telluric model from the DSST to obtain a cleaned DSST.

Naturally, this process does not clean up the DSST perfectly, especially for deep and moderately deep lines, where the line profiles are hard to match. We hope to improve this cleaning process, as well as the entire process for making DSSTs, which we discuss in Section 4.5.

4.2.4 Summary of Recommendations on Treating Telluric Contamination

As argued and tested in the previous subsections, to effectively eliminate the adverse effects of telluric contamination in iodine-calibrated precise RV data, we recommend the following strategies:

- Masking deep and saturated lines and wings liberally, or deserting such spectral regions completely.
- Creating DSST following the recipe described in Section 4.2.3.3, i.e. modeling “out” telluric lines in every step.
- Incorporating forward modeling of telluric lines in the RV extraction process.
- Assigning low statistical weights to RVs reported by telluric-contaminated spectral regions which suffer from low RV precision and accuracy.

We outline potential improvements and future works at the end of this chapter in Section 4.5.

4.3 Errors Induced by Imperfect Stellar Reference Spectra

For a while, we believed that telluric contamination was the major culprit behind the Keck/HIRES RV-BC anomaly. However, the simulations in previous section have revealed that tellurics probably only contribute a small amount, mostly buried underneath photon noise and errors induced by the data analysis code (algorithmic errors). We quickly focused our suspicion to DSST, because we saw the differences in DSSTs before and after telluric cleaning (described in Section 4.2.3.3), which are often larger than the micro-telluric lines and could easily manifest as trends in the RV-BC plane.

Any errors in the DSST, i.e., differences between the true stellar spectrum and our assumed knowledge of truth (the DSST), are just like persisting spectral contamination in the star’s frame (instead of the Earth’s frame like the telluric contamination). Therefore, it beats against the iodine lines as the stellar lines move back and forth through the forest of iodine lines due to the Earth barycentric motion and the star’s intrinsic RV variation.

As a result, it manifests as anomalous RV-BC trends and adds bias and scatter to the final RVs.

We do know for sure that there are errors in the DSST, which could arise from every step of its making (See Chapter 2.2.1 for details on how DSSTs are made and potential sources of errors associated with each steps). However, the question is how much does a DSST differ from the real stellar spectrum, and more importantly, how do these errors translate into RV errors and how much. Is the RV-BC trend with ~ 1 m/s amplitude dominated by such errors?

To answer these questions, we performed simulations using synthetic data and the Doppler code. In the previous section, when simulating the effects of telluric lines, we used “perfect DSSTs” that are simply the input synthetic stellar spectrum (no deconvolution involved, of course), and naturally they do not have errors in them (except for a small amount of interpolation errors). To study the effects of errors in DSST, we need to bring back the deconvolution process and simulate the entire process of the making of a DSST. In the following texts, the true stellar spectrum, i.e., the synthetic input stellar spectrum, is referred to as the “true spectrum”, while the DSST made based on simulated data is referred to as the “simulated DSST”, as opposed to a real DSST which is derived from real data.

Here is how we made a simulated DSST:

1. Simulating the relevant observed data, using the same method as described in Section 4.2.2.1, which include: iodine-free stellar frames (normally 4–5 frames with high SNR) and a couple of B star + iodine frames before and after the stellar frames. The B star + iodine frames are simulated using the best-fit IPs and wavelength solutions derived from their real observation pairs. The IPs and wavelength solutions used in simulating the iodine-free stellar spectrum are from the same source. All of these simulated frames are free of telluric lines.
2. Fitting the simulated B star + iodine frames for wavelength solution and IPs, just as if they are real observed data.
3. Using the fits on B star + iodine frames to deconvolve the iodine-free stellar frames, and the result would be the simulated DSST. The codes used in this step are exactly the same ones in the CPS Doppler package.

We then used the simulated DSST for RV extraction on an ensemble of simulated star+iodine data (again, made in the same way as how we made the telluric-free simulated data as described in Section 4.2.2.1). We compare the RVs between the extraction using the perfect DSST and the simulated DSST in Figure 4.7. Indeed, the errors in DSST are capable of inducing RV-BC trends with an amplitude near 1 m/s. Because we used the same IPs and wavelength solutions for the B star + iodine frames and the stellar frames, the errors in DSST we are probing here are mostly from the deconvolution algorithm and random errors from the photon noise and the CPS Doppler code/algorith. The current CPS code uses an unconventional “deconvolution” algorithm written by John Johnson, which forward models the underlying stellar spectrum by adding Gaussians at places of stellar lines, convolving that with the derived IP from B star frames, and minimizing the χ^2 to find depths and positions of the Gaussian nodes. CPS has used

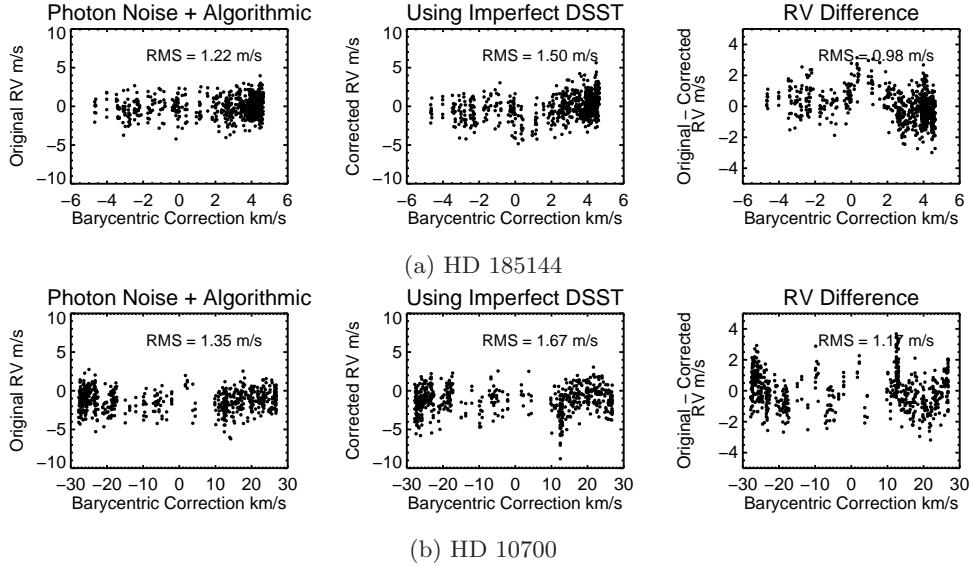


Figure 4.7 Effect of imperfect DSST as revealed by simulated data. The top panels are for HD 185144 and the bottom panels are for HD 10700. The left panels are RVs from the simulation run using the perfect DSST or the true spectrum. The middle panels are RVs from the simulation run using the simulated DSST, which has errors from the manufacture process of DSST such as the deconvolution. The right panels are RV differences between the left and the middle ones, illustrating the effects of DSST errors on RV precision and accuracy, because the only difference between these two runs is the DSST.

other deconvolution algorithms such as the Jansson constrained nonlinear method, which did not out-perform the current recipe using “piston Gaussians” (Ming Zhao, Jason Wright, and John Johnson, private communications).

The RV-BC patterns in the middle panels of Figure 4.7 are different from the real observed RV-BC trends, which we tentatively attribute to: (1) the synthetic spectrum is different from the real stellar spectrum; (2) the noise and other random errors are different, e.g., how IPs used in deconvolution can differ from the real IPs due to the added photon noise, although they have the same input in the simulation; (3) and the fact that there are other sources of RV systematic errors in real data. Note that in reality, the difference in IPs and wavelength solution between the B star and stellar frames can rise from many more places besides photon noise.

We then compared the simulated DSST to the true spectrum (the input synthetic spectrum), which revealed small but non-negligible differences, as illustrated in the left panels of Figure 4.8. The effects on RVs rising from these differences are also illustrated in Figure 4.8, in the right panels. These are cherry-picked chunks which show the same RV-BC patterns in simulation and real observations. Why is it that not all chunks show the same RV-BC patterns in simulation and real observations? The answer to this question will shine some light on the original of the DSST errors and hopefully even a remedy, which will be one of the focuses of future works.

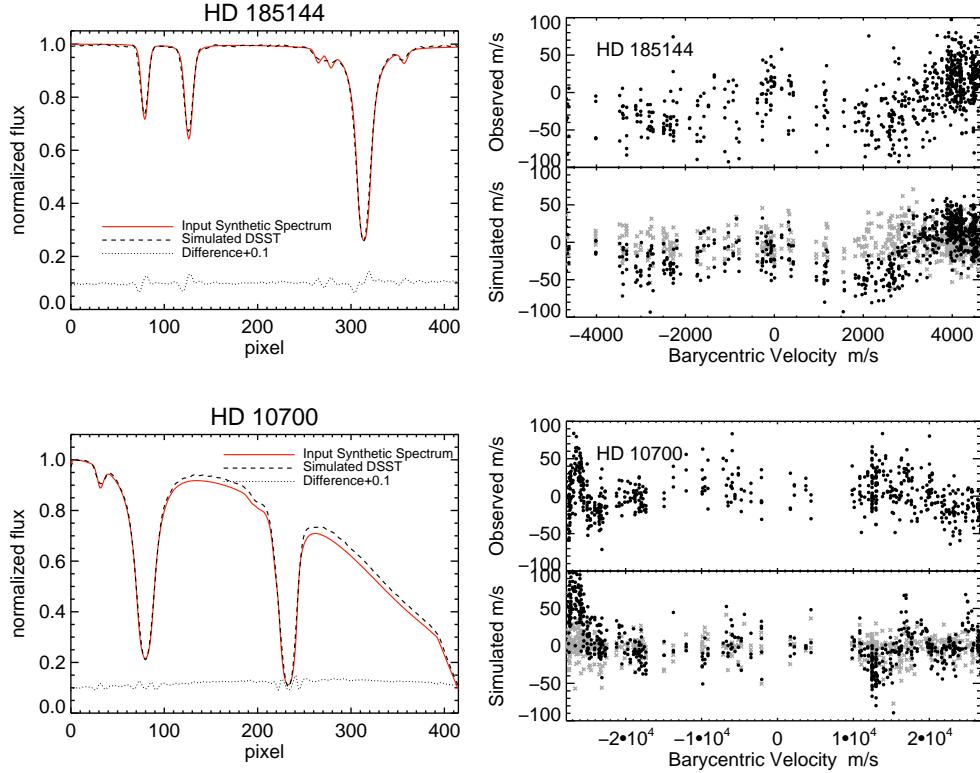


Figure 4.8 Effect of imperfect DSST on simulated data for a single spectral chunk. The top panels are for a chunk near 5160Å for HD 185144, and the bottom panels are for a chunk around 5166Å for HD 10700. The left panels illustrate the differences between the simulated DSST (solid red) and the true spectrum (dashed black), i.e., the errors in DSST. The right panels show the derived RVs for this chunk as a function of BC: the RVs on top, with *y*-axis labeled as “Observed”, are from real Keck/HIRES data, and RVs below, labeled with “Simulated”, are from simulations using the simulated DSST (black dots) and the “perfect” simulation using the true spectrum as the DSST (gray dots, with no apparent RV-BC trends). The 5100–5200Å spectral region tends to receive high weights due to its high density of stellar and iodine lines and a lack of telluric lines.

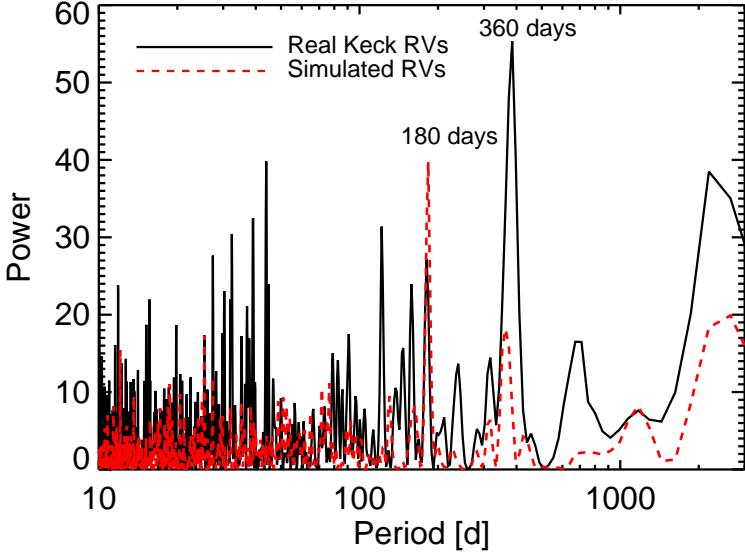


Figure 4.9 Lomb-Scargle periodograms for HD 185144 RVs using real (black line) or simulated (red dashed line) Keck/HIRES data. The spurious peaks near 180 and 360 days are caused by systematic errors which correlates with BCs. In the case of the simulated data, we know that this systematic arises from the errors in the DSST because all other conditions are kept “perfect”. For the real data, it is a combination of DSST errors and other systematics (such as, potentially, telluric contamination).

The RV-BC pattern induced by the errors in the DSST can be very detrimental to planet search, because it manifests as spurious signals at a period of one sidereal year and its harmonics with a large amplitude (m/s in HD 185144 Keck/HIRES data). For example, Vogt et al. (2014) tentatively claims a planet candidate around HD 185144, with a period of ~ 308 days and a mass of $12 M_{\oplus}$. They show, in their Figure 16, a power spectrum/periodogram of the HD 185144 Keck/HIRES RVs, with the dominant peak around 300 days (and a harmonic around 180 days). However, the errors in the DSST can produce a spurious periodic signal just like this, as shown by the red dashed lines in Figure 4.9. Before these errors are eliminated, it would be very hard to argue for the existence of any planet with such orbital periods, which, unfortunately, coincides with the Habitable Zones around Sun-like stars.

4.4 Numerical and Algorithmic Errors

Solving the least- χ^2 problem in a multi-dimensional, multi-modal setting is not easy. Efficient as it might be in terms of computational time, the least- χ^2 algorithm employed by the CPS Doppler code could give a biased solution and comprise the RV precision and accuracy. We performed simulations to probe the magnitude of such errors.

Figure 4.10 illustrates the RV scatter and RV-BC trends induced by algorithmic

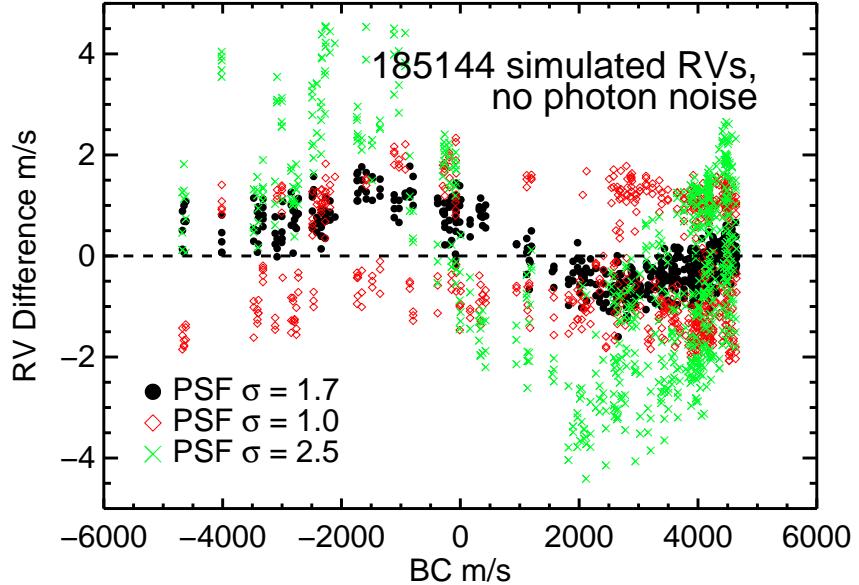


Figure 4.10 RV vs. BC for HD 185144, for simulations with fixed simple IPs and no photon noise added. The data plotted here are from simulated spectral data with different IP widths (or spectral resolution, $\sigma = 1.7$ pixels corresponds to original Keck/HIRES resolution). The origins of the RV scatters and trends in these plots are purely algorithmic.

errors – the RVs plotted are from simulations with no photon noise added, using the exact input spectrum as the DSST, and the exact simple Gaussian IPs as the input. A perfect algorithm should return all zeros for the RV, which is not the case here. To probe the origin of this algorithmic errors a bit further, we have run three sets of simulations at three different spectral resolution (i.e., using different Gaussian IP widths). The amplitude of the RV scatter decreases as the resolution increases, which is as expected because shallower spectral lines probably translate to shallower χ^2 surfaces. However, the signature period or length scale (in BC space) of the RV-BC trend does not change, which is about 8 km/s or 6 pixels on the CCD. We are still investigating the origin of this signature length scale (perhaps its in the blaze normalization, or interpolation or rebinning algorithms, to name a couple). The dichotomy in the red points (the fact that they split into two RV groups) is probably due to the convergence criteria (which are tuned for Keck/HIRES resolution, and not sufficient for significantly higher ones) and the algorithm's sensitivity in initial guesses. One caveat is that because the IPs were fixed in the Doppler code run for these simulations, the fitter only went through one pass (three free parameters) instead of the standard two passes, where in the final pass the wavelength dispersion scale is fixed. We will investigate whether is this the major cause of algorithmic errors shown in Figure 4.10.

Figure 4.11 shows an example chunk with visible algorithmic errors. Again the top

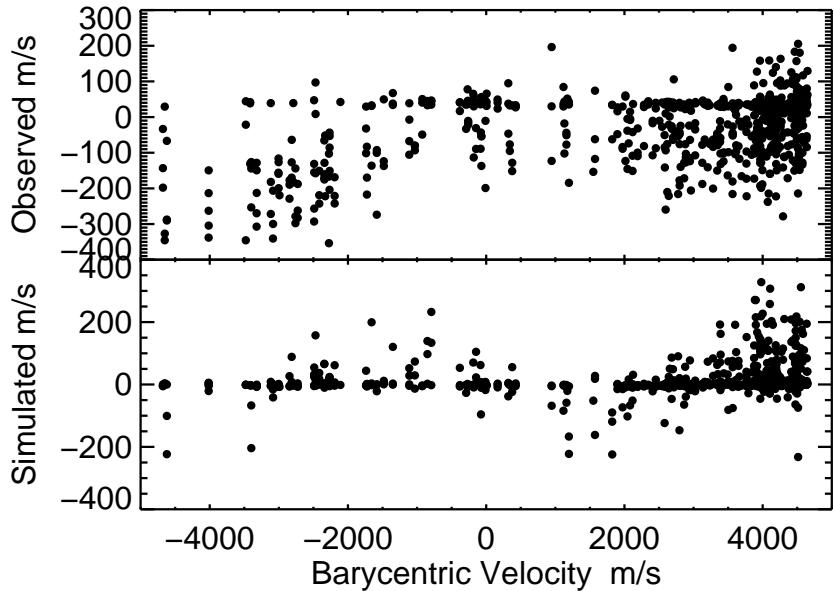


Figure 4.11 Real (top) or simulated (bottom) Keck RVs from a 2Å chunk from HD 185144 spectra showing RV systematic errors caused by algorithmic errors. The RVs in the bottom panel are from the simulations with complex IPs and photon noise.

panel is showing real data, and the bottom one is the simulated data (with photon noise and complex IPs). This is a severe case, because its a chunk near 5900Å which does not contain a large number of stellar or iodine lines, and therefore it is particularly challenging for the algorithm to find a good χ^2 minimum.

Unfortunately it is hard to quantify how much RV RMS the algorithmic errors add to the RV budget, because the major damage comes from the biases instead of the increased scatter (and the “vanking” procedure breaks down because of this, unable to weigh the chunks effectively and improve the precision). Figure 4.10 suggests that it can lead to spurious signals with considerable amplitude (1-3 m/s) for high SNR observations (the simulated observations here essentially have infinitely SNR).

To list a few places in the Doppler code which may cause significant numerical and algorithmic errors (not in any particular order):

- The LM least- χ^2 fitter: it can get stuck in local minima, and it may have reached the crude convergence criteria but not actually converged, and it is extremely sensitive to initial guesses.
- The interpolation algorithm for interpolating or rebinning the model spectra. The current interpolator does not seem to make an effort at conserving flux, for example.
- The simple linear fit to the blaze function in each chunk (each chunk contains 80 pixels).
- “Vanking”, or the final statistical weighting process, only takes into account RV

Table 4.3. Summary of Keck/HIRES RV Errors Found in This Work

Term	RV RMS or Amplitude
precision on an RV standard star	2 m/s
error added by errors in the DSST	1 m/s
error added by telluric contamination	0.6 m/s
projected precision after telluric and DSST correction	1.6 m/s
<hr/>	
spurious signal at period of 1 year and harmonics	
added by errors in the DSST	$K = 1 \text{ m/s}$
added by telluric contamination	$K = 0.2 \text{ m/s}$

scatter of each chunk but not biases.

Overall, it is not surprising at all that this decade-old Doppler code fails to deliver precise RVs to the modern era standard. We discuss the paths forward in Chapter 7.

4.5 Conclusions and Future Work

In this chapter, we described our work on discovering and characterizing three RV systematic error sources for Keck/HIRES: telluric contamination, errors in the DSST, and algorithmic errors. A summary of RV errors is presented in Table 4.3.

Telluric contamination, and in particular, micro-telluric lines, has a small but non-negligible effects on the RV precision and accuracy. For a typical G type star like the ones targeted by CPS, micro-tellurics adds $\sim 0.6 \text{ m/s}$ to the RV error budget (in quadrature) and also a spurious signal with an amplitude on the order of 10-20 cm/s if left untreated. We have summarized the best strategies for treating micro-telluric lines for precise RV using iodine calibrators in Section 4.2, which are masking deep lines, cleaning DSST, forward modeling the telluric lines in RV extraction, and assigning lower statistical weights to telluric-contaminated chunks.

Errors in the DSSTs are one of current major RV systematic sources in Keck/HIRES RV data, and quite likely the largest one. It adds 1 m/s to the RV error budget and creates a spurious signal with amplitude $\sim 1 \text{ m/s}$. RV-BC trends in simulated data match the ones seen in the observed spectra among many spectral chunks, which pinned down this source of error unambiguously.

Algorithmic errors contribute a considerable amount to the current Keck/HIRES RV budget as well, and it is hard to imagine it would deliver sub-m/s precision. After all, it is extremely challenging to find a χ^2 minimum on a complex surface with 15 dimensions (parameters) using spectral data with 80 points (80 pixels per chunk) and highly covariant model parameters.

In terms of future work along these lines:

(1) Continuing the battle with telluric contamination: We plan to run simulations on an M star, HD 95735, also an RV standard. M stars are particularly

interesting because they may be more susceptible to telluric contamination – they have more stellar lines in the red where the telluric lines are denser. We also plan to implement full forward-modeling of telluric lines in real data reduction (as opposed to our toy modeling with fixed water column density for all dates/observations), and find out whether a priori or floating water column density parameter works better. We plan to improve the DSST cleaning process by incorporating forward modeling of telluric lines in the “piston Gaussians deconvolver” that is currently used for making DSST, instead of dividing them out. We hope to implement our telluric correction package into the official CPS pipeline eventually, which will be busily processing Keck/HIRES follow-up data on TESS targets (many M dwarfs, undoubtedly) in the near future.

(2) Searching for a better DSST fabrication method: We will diagnose the origin of the DSST errors while searching for a better deconvolution algorithm. It is possible that the current CPS “deconvolution” algorithm described in Section 4.3 is sufficient, but the problem is in the normalization or some other parts of the process. Nonetheless, it could be rewarding to jump out of the current frame work and try another completely different approach (e.g., constructing DSST from all the star+iodine frames, which has been attempted by Andrew Vanderberg and John Johnson).

(3) Eliminating the numerical and algorithmic errors: The CPS Doppler code can certainly use an upgrade. However, in order to eliminate a large portion of the algorithmic errors we are seeing now, it probably means major structural changes to this lengthy and complex legacy code written for early-1990 computers, regardless whether one would like to stick with the L-M least- χ^2 fitter. Therefore, instead, I am building a new RV extraction code from scratch, using Python and modern algorithms, which is one of the topics in Chapter 7.

Chapter 5

Characterization of Exoplanet Systems Using Radial Velocities

5.1 Introduction and Background

Extracting exoplanet signals from RV data is hard in many ways (see Section 4 of Fischer et al. 2016 for a summary of the status of this problem as of 2015). First, it can be hard to identify and quantify stellar activity signals from planetary signals, and one of the examples is the famous case of the GJ 581 system (Mayor et al. 2009; Vogt et al. 2010; Hatzes 2013; Robertson et al. 2014; Anglada-Escudé & Tuomi 2015; Robertson et al. 2015). Second, even with the definitive knowledge that the RV signal is dominated by planets, it can still be challenging if the planetary system is dynamically active, e.g., for the case of the 55 Cnc system (Nelson et al. 2014b), and characterizing planetary orbits is a numerically and computationally challenging problem (Nelson et al. 2014a). Third, even if the star is RV quiet and the planetary system is dynamically quiet, and all orbits can be described by simple Keplerian orbits, several challenges remain for parameter estimation and model selection: the number of planets in the system can be hard to pin down (Vogt et al. 2015; Motalebi et al. 2015; Johnson et al. 2016); some orbits may not be well constrained and thus raise ambiguities (e.g., circular orbits or eccentric; e.g., Anglada-Escudé et al. 2010; Wittenmyer et al. 2013; Kürster et al. 2015); and it can be computationally demanding.

In 2012, there were very few published codes on performing simple Keplerian fit using RV data under the context of exoplanet detection (i.e., handling multi-planets, data from multiple telescopes, etc.; e.g., the systemic console, Meschiari et al. 2009), especially ones that were easy to use. The RVLIN package by Wright & Howard (2009) addresses the problem of simple Keplerian orbital fitting using least- χ^2 algorithm and exploiting the linear parameters (i.e., K , ω , γ , and RV trend) to speed up the convergence. It handles multi-planet systems and RV data from multiple telescopes, is very easy to use (simple input requirements and easy commands; written in IDL), and its typical time for convergence is within seconds. It is fairly popular and has a large user group beyond the exoplanet community (e.g., for binary stars and systems characterized using astrometry, Koren et al. 2016).

However, RVLIN does not provide estimates for uncertainties on the best-fit parameters. This becomes a much desirable feature especially for the Transit Ephemeris Refinement and Monitoring Survey (TERMS) project (Kane et al. 2009). TERMS follows up RV detected exoplanets with moderate separation from their bright, nearby host stars (semi-major axis > few hundredths of an AU) to search for transiting “warm” planets (as opposed to the Hot Jupiters), which are unique and important targets for atmospheric characterization. TERMS uses RVLIN for orbital parameter estimates, but only having the best-fit Keplerian parameters will not suffice for transit prediction and transit observation planning. A transit search can only be feasible if the transit window is well constrained by the RV data. Otherwise, more RVs need to be collected. As many of the TERMS targets were reported in the literature a while ago, the predicted transit ephemerides might have drifted from the true ones as the predictive power of old RV data faded as time went by. Therefore, estimating uncertainties on transit ephemerides becomes crucial for the project.

With the purpose of calculating transit ephemerides and their uncertainties for TERMS, and also to supplement RVLIN with a tool to estimate uncertainties, I constructed the BOOTTRAN package, which uses bootstrapping to calculate uncertainties for the Keplerian parameters estimated by RVLIN. BOOTTRAN was heavily used in the TERMS project and others (e.g., some of my co-authored publications: Henry et al. 2013; Dragomir et al. 2012; Kane et al. 2011e,a; and Feng et al. 2015). As of March 2016, BOOTTRAN and Wang et al. (2012) have received ~ 30 citations. The statistical justification and algorithm of BOOTTRAN is described in the next section. The final section of this chapter describes my contributing work on characterizing exoplanetary systems in in Feng et al. (2015) and Henry et al. (2013) using other tools besides BOOTTRAN. The next chapter, Chapter 6, describes the planetary system around the star HD 37605, which serves as an example of application of RVLIN and BOOTTRAN(similar with other TERMS publications citing Wang et al. 2012 such as Dragomir et al. 2012).

5.2 BOOTTRAN: Uncertainties for Orbital Parameters Estimated Using Radial Velocities

The following text is originally published in the appendix of Wang et al. (2012) in ApJ, and copyright belongs to IOP Publishing. The original text is written by the author, with light edits from co-authors. They are used in this thesis with permission (with minor modification to fit in this chapter).

We have constructed a package, BOOTTRAN, to calculate uncertainties for Keplerian orbital parameter estimates¹ and transit mid-time T_c via bootstrapping (Freedman 1981; Davison & Hinkley 1997). BOOTTRAN is designed to calculate error bars for transit ephemerides and the Keplerian orbital fit parameters output by the RVLIN package(Wright & Howard 2009), but can also be a stand-alone package. The two packages, RVLIN and BOOTTRAN, are publicly available at <http://exoplanets.org/code/> and the Astrophysical Source Code Library (ASCL.net). Thanks to the simple concept

¹Through out this thesis, we refer to the “estimates of the parameters” (as distinguished from the “true parameters”, which are not known and can only be estimated) simply as the “parameters”.

of bootstrapping, it is computationally very time-efficient and easy to use.

The basic idea of bootstrap is to resample based on original data to create bootstrap samples (multiple data replicates); then for each bootstrap sample, derive orbital parameters or transit parameters through orbital fitting and calculation. The ensemble of parameters obtained in this way yields the approximate sampling distribution for each estimated parameter. The standard deviation of this sampling distribution is the standard error for the estimate.

We caution the readers here that there are regimes in which the “approximate sampling distribution” (a frequentist’s concept) is not an estimate of the posterior probability distribution (a Bayesian concept), and there are regimes (e.g., when limited sampling affects the shape of the χ^2 surface) where there are qualitative differences and the bootstrap method dramatically underestimates uncertainties (e.g., long-period planets when the observations are not yet sufficient to pin down the orbital period; Ford 2005; Bender et al. 2012). In situations with sufficient RV data, good phase coverage, a sufficient time span of observations and a good orbital fit, bootstrap often gives a useful estimate of the parameter uncertainties. For the data considered in Chapter 6 (Wang et al. 2012), it was not obvious that the bootstrap uncertainty estimate would be accurate, as the time span of observations is only slightly longer than the orbital period of planet c . Nevertheless, we find good agreement between the uncertainty estimates derived from bootstrap and MCMC calculations in Wang et al. (2012).

The radial velocity data are denoted as $\{\vec{t}, \vec{v}, \vec{\sigma}\}$, where each t_i, v_i, σ_i represents radial velocity v_i observed at time (BJD) t_i with velocity uncertainty σ_i . Extreme outliers should be rejected in order to preserve the validity of our bootstrap algorithm. We first derive our estimates for the true orbital parameters from the original RV data via orbital fitting, using the RVLIN package:

$$\vec{\beta} = \mu(\vec{t}, \vec{v}, \vec{\sigma}), \quad (5.1)$$

where $\vec{\beta}$ is the best fitted orbital parameters². From $\vec{\beta}$, we derive $\{\vec{t}, \vec{v}_{best}(\vec{\beta})\}$, the best-fit model (here \vec{t} are treated as predictors and thus fixed). Then we can begin resampling to create bootstrap samples.

Our resampling plan is model-based resampling, where we draw from the residuals against the best-fit model. For data that come from the same instrument or telescope, in which case no instrumental offset needs to be taken into account, we simply draw from all residuals, $\{\vec{v} - \vec{v}_{best}\}$, with equal probability for each $(v_i - v_{best,i})$. This new ensemble of residuals, denoted as \vec{r}^* , is then added to the best-fit model \vec{v}_{best} to create one bootstrap sample, \vec{v}^* ³. Associated with \vec{r}^* , the uncertainties $\vec{\sigma}$ are also re-assigned to \vec{v}^* – that is, if $v_j - v_{best,j}$ is drawn as r_k and added to v_k to generate v_k^* , then the uncertainty for v_k^* is set to be σ_j .

For data that come from multiple instruments or multiple telescopes, we incorporate our model-based resampling plan to include stratified sampling. In this case, although

²As described in §6.3.2, this includes the P , T_p , K , e , and ω for each planet, as well as γ , dv/dt (if applicable), and velocity offsets between instruments/telescopes (if applicable) for the system.

³We simply use the raw residual instead of any form of modified residual, because the RV data for any single instrument or telescope are usually close enough to homoscedasticity.

data from each instrument or telescope are close to homoscedastic, the entire set of data are usually highly heteroscedastic due to stratification in instrument/telescope radial velocity precision. Therefore, the resampling process is done by breaking down the data into different groups, $\{\vec{v}_1, \vec{v}_2, \dots\}$, according to instrument and/or telescope, and then resample within each subgroup of data with the algorithm described in last paragraph. The bootstrap sample is then $\vec{v}^* = \{\vec{v}_1^*, \vec{v}_2^*, \dots\}$.

To construct the approximate sampling distribution of the orbital parameter estimates $\vec{\beta}$, we compute

$$\vec{\beta}^* = \mu(\vec{t}, \vec{v}^*, \vec{\sigma}^*) \quad (5.2)$$

for each bootstrap sample, $\{\vec{t}, \vec{v}^*, \vec{\sigma}^*\}$. The sampling distribution for each orbital parameter estimate β_i can be constructed from the multiple sets of $\vec{\beta}^*$ calculated from multiple bootstrap samples ($\vec{\beta}^{*(1)}, \vec{\beta}^{*(2)}, \dots$ from $\vec{v}^{*(1)}, \vec{v}^{*(2)}, \dots$). The standard errors for $\vec{\beta}$ are simply the standard deviations of the sampling distributions⁴.

The sampling distribution of the estimated transit mid-time, T_c , is calculated likewise. Here T_c is the transit time for a certain planet of interest in the system, and is usually specified to be the first transit after a designated time T . However, the situation is complicated by the periodic nature of T_c . Our approach is to first calculate, based on the original RV data, T_{c0} , the estimated mid-time of the first transit after time T_0 (an arbitrary time within the RV observation time window of $[\min(\vec{t}), \max(\vec{t})]$; T_{c0} is also within this window). Then

$$T_c = N \cdot P + T_{c0}, \quad (5.3)$$

where P is the best-estimated period for this planet of interest, and N is the smallest integer that is larger than $(T - T_{c0})/P$. Next we compute T_{c0}^* for each bootstrap sample $\{\vec{t}, \vec{v}^*, \vec{\sigma}^*\}$. Given that within the time window of radial velocity observations ($[\min(\vec{t}), \max(\vec{t})]$), the phase of the planet should be known well enough, it is fair to assume that T_{c0} is an unbiased estimator of the true transit mid-time. Therefore we assert that T_{c0}^* has to be well constrained and within the range of $[T_{c0} - P^*/2, T_{c0} + P^*/2]$, where P^* is the period estimated from this bootstrap sample. If not, then we subtract or add multiple P^* 's until T_{c0}^* falls within the range. Then naturally

$$T_c^* = N \cdot P^* + T_{c0}^*. \quad (5.4)$$

The ensemble of T_c^* 's gives the sampling distribution of T_c and its standard error. Note that T_c^* is not necessarily within the range of $[T_c - P/2, T_c + P/2]$.

Provided with the stellar mass M_\star and its uncertainty, we calculate, for each planet in the system, the standard errors for the semi-major axis a and the *minimum mass* of the planet $M_{p,\min}$ (denoted as $M \sin i$ in the main text as commonly seen in literature, but this is a somewhat imprecise notation). As the first step, the mass function is calculated

⁴The standard deviation of a sampling distribution is estimated in a robust way using the IDL function *robust_sigma*, which is written by H. Fruedenreich based on the principles of robust estimation outlined in Hoaglin et al. (1983).

for the best-fit $\vec{\beta}$ and each bootstrap sample $\vec{\beta}^*$,

$$f(P, K, e) = \frac{PK^3(1-e)^{3/2}}{2\pi G} = \frac{(M_p \cdot \sin i)^3}{(M_\star + M_p)^2}. \quad (5.5)$$

The sampling distribution of $f(P, K, e)$ then gives the standard error of the mass function. The minimum mass of the planet $M_{p,\min}$ is then calculated by assuming $\sin i = 1$ and solving for M_p . Standard error of $M_{p,\min}$ is derived through simple propagation of error, as the covariance between M_\star and $f(P, K, e)$ is probably negligible.

For the semi-major axis a ,

$$a^3 = \frac{P^2 G(M_\star + M_p)}{4\pi^2} \approx \frac{P^2 G(M_\star + M_{p,\min})}{4\pi^2}. \quad (5.6)$$

The standard error of P^2 is calculated from its bootstrap sampling distribution, and via simple propagation of error we obtain the standard error of a (neglecting covariance between P^2 , $M_{p,\min}$, and M_\star).

5.3 Other Applications to Characterize Planetary Systems

This section highlights two published papers, Feng et al. (2015) and Henry et al. (2013), where the author of this thesis is an co-author. The following text first summarizes the paper, and then describe the author's specific contribution. These two papers are selected to represent the author's work in using other tools (instead of BOOTTRAN) to characterize planetary systems.

(1) Feng et al. (2015) reports updates and/or new detections on the planetary systems around eight stars using RV data (primarily taken by Keck/HIRES), including: orbit updates for HD 24040, HD 183263, HD 74156, and HD 187123; a newly detected linear trend in RV for the HD 66428 system; RV coverage for complete orbits for GJ 849c and HD 217107c; detection of HD 145934b. The emphasis of the paper is on detecting and characterizing long-period planets, which can be challenging because some of the planets have orbital periods on a ~ 10 year time scale and thus have relatively poor RV coverage. In such cases, it is important to illustrate good constraints on orbital properties to confirm detection of the planets. A very useful tool is the “Wright diagram” (as referred to by Knutson et al. 2014; see, e.g., Wright et al. 2009b). For example, Figure 5.1 is an Wright diagram for HD 217107c, which shows the χ^2 surface “marginalized” over $M \sin i$ and P (each χ^2 value is for a set of optimized parameters with fixed $M \sin i$ and P ; see Section 6.3.2 for more details). This demonstrates that the mass of HD 217107c is well constrained to justify its planetary nature (nominally $< 13 M_{\text{Jup}}$), and the period is decently constrained even though the RV data barely covers one orbital period for HD 217107c. Figure 11 in Feng et al. (2015) serves the same purpose but for GJ 849c.

(2) Henry et al. (2013) reports the host star properties and transit exclusion for the HD 38529 system. TERMS refined the orbital parameters for HD 38529b and c using new and archival RV data from HET/HRS, Lick/Hamilton, and Keck/HIRES. See Figure 5.2 for the RV plot for this system (orbital fitting and uncertainty estimates from RVLIN and BOOTTRAN). The RV data also enabled us to test the hypothesis of a third planet

in the system speculated by Benedict et al. (2010), using Monte Carlo to estimate the significance of potential periodic signals in the RV residuals to the two-planet Keplerian solution.

The texts below are from Section 3.4 of Henry et al. (2013) and were co-written by me and the second author Stephen R. Kane (with minor revisions for consistency with this chapter).

The results of Benedict et al. (2010) are utilized by the authors to speculate on evidence for a third planet in the system of HD 38529. Thus we also consider this possibility from our analysis since our RV data comprise a substantially larger dataset. As reported by Benedict et al. (2010), a coplanar orbital solution is only stable if the third planet has a period within the window of [33, 445] days and an eccentricity of < 0.3 , or a period larger than their RV data baseline (> 10 years). For this reason, we focused our search for the third planet within the period window of [33, 445] days, and constrained the eccentricity to be < 0.3 .

We first searched for strong periodic signals in the residuals of the two-planet Keplerian solution by fitting sinusoids to the residuals at different periods within [33, 445] days (with 0.4 day step in period). The results are plotted in solid line in Figure 5.3. We then estimated the false positive probability to see if any of the strong peaks are significant enough. We define the false positive probability for a peak with a certain amplitude K' as the probability that a signal with amplitude $\geq K'$ is generated by the residuals just by chance. We generated 1000 sets of simulated residuals by scrambling the true residuals (and their associated errors, with replacements), and then searched for the peak with largest amplitude within the $P = [33, 445]$ day window for each of the 1000 sets. These 1000 amplitudes provide approximately the distribution of amplitudes arising purely from random noise in the residuals. Any peak in Figure 5.3 that has an amplitude smaller than 950 (95%) of these 1000 amplitudes is thus considered having false positive probability of $> 5\%$. This is marked by the top dashed line in Figure 5.3, and similarly for the 10% and 50% lines.

As shown in Figure 5.3, no peak has a false positive probability of less than 5%, and two with less than 10% at 119 days and 164 days. We see no significant peak around 194 days as reported by Benedict et al. (2010). We then performed 3-planet Keplerian fit with our RV data within the $P = [33, 445]$ day window and with the constraint that the eccentricity must be smaller than 0.3. We found that indeed the best-fit is near 164 days, with $e = 0.3$ (also true if we force the third planet to be on circular orbit; best-fit $e = 0.99$ if no constraint on e is required). The χ^2_ν of this fit is 9.58, and an F-test suggests that the 3-planet model provides a better fit though having 5 more parameters. However, the RMS for this fit is 11.92 m s^{-1} , i.e., adding a third planet does not reduce the RMS of the fit. Combining with the fact that this signal at $P = 164$ days does not have lower than 5% false positive probability, we cannot conclude that our data have detected a third planet in the HD 38529 system.

We note here that including or excluding this third planet does not affect our transit exclusion analysis in the following sections, because the changes in the orbital parameters for both HD 38529 b or c , after adding the third planet, are smaller than their error bars.

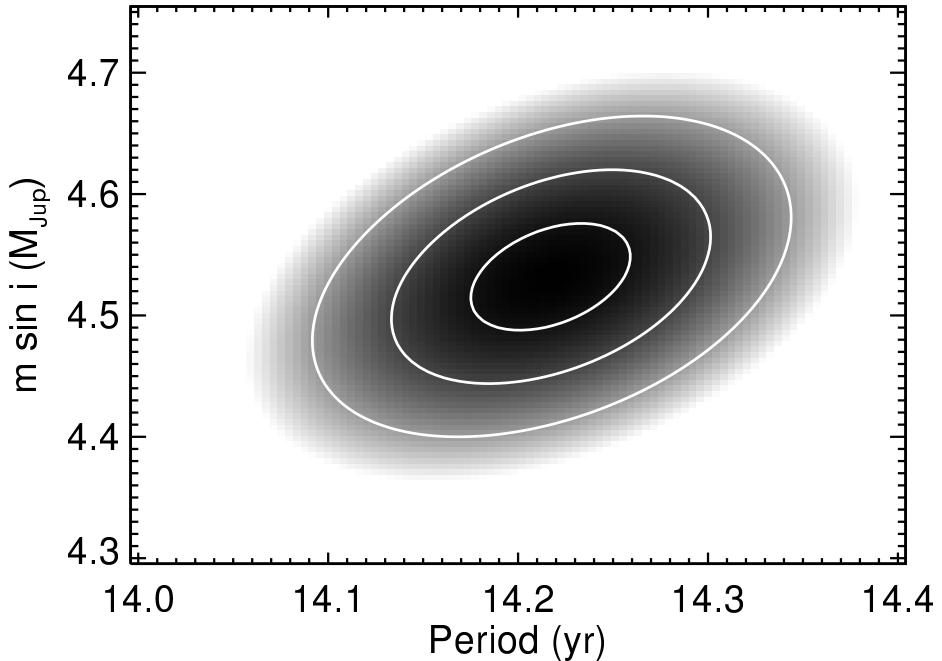


Figure 5.1 Best-fit 100×100 χ^2 map for fixed values of P_c and $M_c \sin i_c$ for HD 217107c. This confirms that the period and mass are well-constrained. We have illustrated the contours of the 1σ , 2σ , and 3σ (defined by $\chi^2 = \chi^2_{\min} + \{2.30, 6.17, 11.8\}$) confidence levels, based on for the number of degrees of freedom in the problem (Press et al. 2002). The center and 1σ limits in both parameters are consistent with the bootstrapping uncertainties for these parameters. This figure is published as Figure 9 in Feng et al. (2015) and was co-produced by the author of this thesis and the leading author of the paper, Y. Katherina Feng.

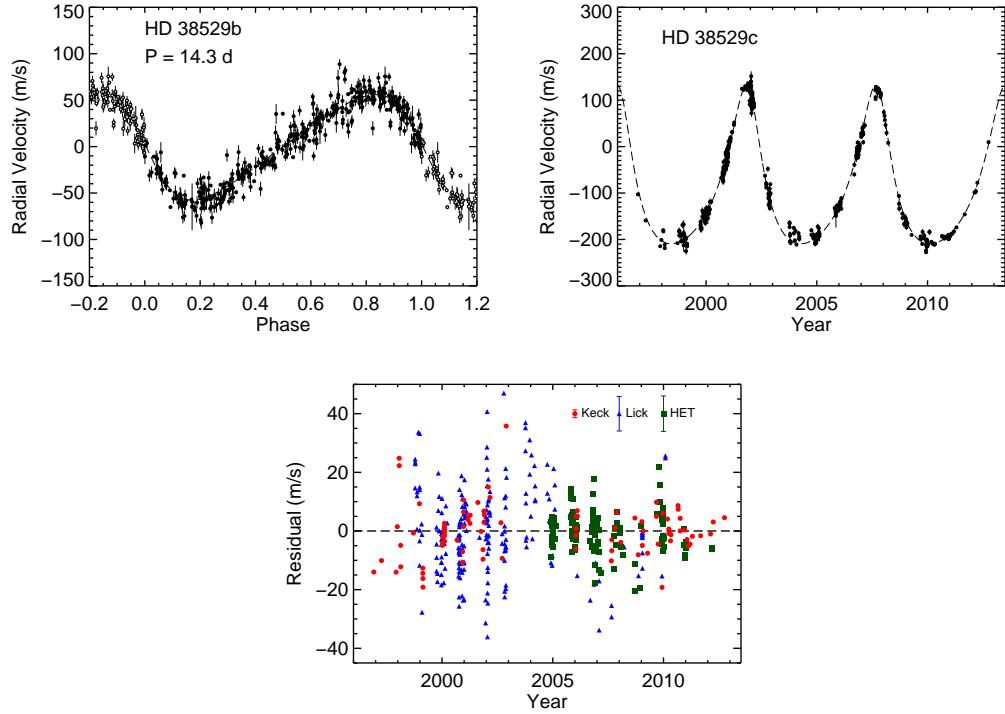


Figure 5.2 Top two panels: radial velocity signal (black dots) induced by HD 38529b and *c*, respectively, and the best-fit orbital solution (dashed line). Error bars shown are internal errors for each observation. The radial velocity signal for each planet was extracted by subtracting off the best-fit orbital velocities of the other planet from the total observed RVs. Bottom panel: residual velocities with respect to the best two-planet orbital solution. The red dots are for Keck data (data sets 3 and 4 in Henry et al. (2013)), the blue triangles are for Lick data (data sets 5 and 6), and the green squares are for the HET data (data sets 1 and 2). The typical size of internal error bars for each telescope (\pm median internal errors) are plotted on the upper right of this panel. This figure is published as Figure 2 in Henry et al. (2013) and was made by me.

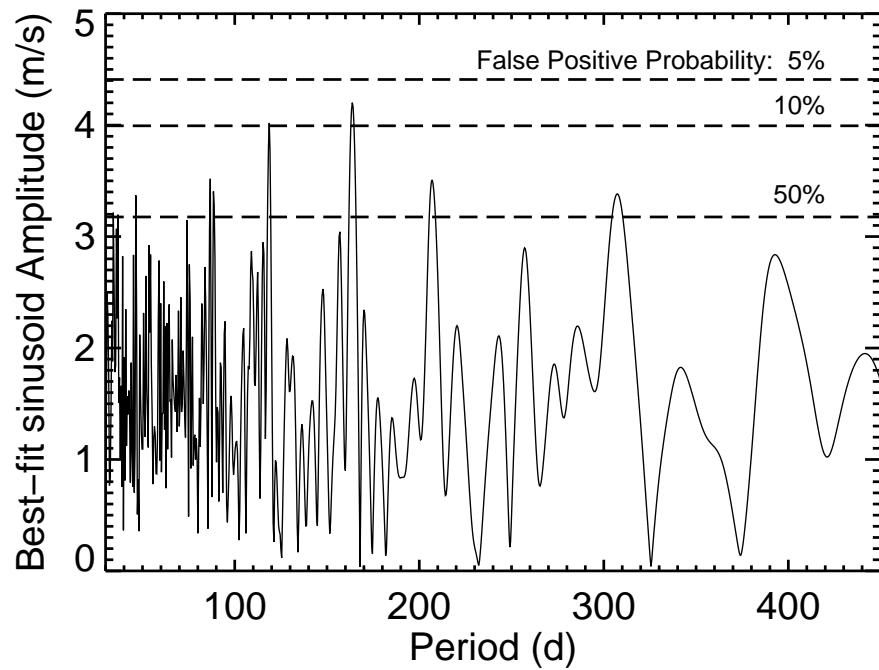


Figure 5.3 Amplitude of best-fit sinusoids to the residuals of the two-planet Keplerian solution (solid line). Any peak in this period window that has amplitude larger than the top dashed line is considered to be significant for having $< 5\%$ false positive probability. Similar meanings for the two lower dashed lines ($< 10\%$ and $< 50\%$). No period within this window has less than 5% false positive probability, and the two peaks with $< 10\%$ false positive probability are at 119 days and 164 days. We see no significant peak around 194 days as reported by Benedict et al. (2010). This figure is published as Figure 3 in Henry et al. (2013) and was made by me.

Chapter 6

The Discovery of HD 37605 c and a Dispositive Null Detection of Transits of HD 37605 b

The content in this chapter was published in ApJ, and the copyright belongs to IOP Publishing; all texts, figures, and tables are used in this thesis with permission. Most of the texts were written by Sharon Xuesong Wang, with the exception of Section 6.1 and Section 6.3.5 (both by Jason T. Wright), Section 6.3.4 (by Matthew J. Payne), and the first three paragraphs in Section 6.4.2 (by Stephen R. Kane and Victoria Antoci).

For figures and tables: Figure 6.3 and 6.4 were made by Mathew J. Payne, Figure 6.5 and 6.6 were made by Gregory W. Henry, and Figure 6.7 was made by Stephen R. Kane. The rest of the figures were made by Sharon Xuesong Wang. All tables were compiled by Sharon Xuesong Wang, although some of the data came from contributing authors: Table 6.1 was based on SME analysis results done by Jeff A. Valenti; Table 6.3 contains orbital parameters estimated using MCMC by Mathew J. Payne; the data in Table 6.4 are provided by Gregory W. Henry; and the data in Table 6.5 are provided by Stephen R. Kane, Victoria Antoci, Diana Dragomir, and Jaymie M. Matthews.

6.1 Introduction

6.1.1 Context

Jupiter analogs orbiting other stars represent the first signposts of true Solar System analogs, and the eccentricity distribution of these planets with $a > 3$ AU will reveal how rare or frequent true Jupiter analogs are. To date, only 9 “Jupiter analogs” have been well-characterized in the peer reviewed literature¹ (defined here as $P > 8$ years, $4 > M \sin i > 0.5 M_{\text{Jup}}$, and $e < 0.3$; Wright et al. 2011, exoplanets.org). As the duration

¹HD 13931 b (Howard et al. 2010), HD 72659 b (Moutou et al. 2011), 55 Cnc d (Marcy et al. 2002), HD 134987 c (Jones et al. 2010), HD 154345 b (Wright et al. 2008, but with possibility of being an activity cycle-induced signal), μ Ara c (Pepe et al. 2007), HD 183263 c (Wright et al. 2009a), HD 187123 c (Wright et al. 2009a), and GJ 832 b (Bailey et al. 2009).

of existing planet searches approach 10–20 years, more and more Jupiter analogs will emerge from their longest-observed targets (Wittenmyer et al. 2012; Boisse et al. 2012).

Of the over 700 exoplanets discovered to date, nearly 200 are known to transit their host star (Wright et al. 2011, [exoplanets.org](#); Schneider et al. 2011, [exoplanet.eu](#)), and many thousands more candidates have been discovered by the *Kepler* telescope. Of all of these planets, only three orbit stars with $V < 8$ ² and all have $P < 4$ days. Long period planets are less likely than close-in planets to transit unless their orbits are highly eccentric and favorably oriented, and indeed only 2 transiting planets with $P > 20$ days have been discovered around stars with $V < 10$, and both have $e > 0.65$ (HD 80606, Laughlin et al. 2009, Fossey et al. 2009; HD 17156, Fischer et al. 2007, Barbieri et al. 2007; both highly eccentric systems were discovered first with radial velocities).

Long period planets not known to transit can have long transit windows due to both the large duration of any edge-on transit and higher phase uncertainties (since such uncertainties scale with the period of the orbit). Long term radial velocity monitoring of stars, for instance for the discovery of low amplitude signals, can produce collateral benefits in the form of orbit refinement for a transit search and the identification of Jupiter analogs (e.g., Wright et al. 2009a). Herein, we describe an example of both.

6.1.2 Initial Discovery and Followup

The inner planet in the system, HD 37605 b , was the first planet discovered with the Hobby-Eberly Telescope (HET) at McDonald Observatory (Cochran et al. 2004). It is a super Jupiter ($M \sin i = 2.41 M_{\text{Jup}}$) on an eccentric orbit $e = 0.67$ with an orbital period in the “period valley” ($P = 55$ days; Wright et al. 2009a).

W.C., M.E., and P.J.M. of the University of Texas at Austin, continued observations in order to get a much better orbit determination and to begin searching for transits. With the first new data in the fall of 2004, it became obvious that another perturber was present in the system, first from a trend in the radial velocity (RV) residuals (i.e., a non-zero dv/dt ; Wittenmyer et al. 2007), and later from curvature in the residuals. By 2009, the residuals to a one-planet fit were giving reasonable constraints on the orbit of a second planet, HD 37605 c , and by early 2011 the orbital parameters of the c component were clear, and the Texas team was preparing the system for publication.

6.1.3 TERMS Data

The Transit Ephemeris Refinement and Monitoring Survey (TERMS; Kane et al. 2009) seeks to refine the ephemerides of the known exoplanets orbiting bright, nearby stars with sufficient precision to efficiently search for the planetary transits of planets with periastron distances greater than a few hundredths of an AU (Kane et al. 2011c; Pilyavsky et al. 2011a; Dragomir et al. 2011). This will provide the radii of planets not experiencing continuous high levels of insolation around nearby, easily studied stars.

In 2010, S.M. and J.T.W. began radial velocity observations of HD 37605 b at HET from Penn State University for TERMS, to refine the orbit of that planet for a fu-

²55 Cnc e (McArthur et al. 2004; Demory et al. 2011), HD 189733 (Bouchy et al. 2005), and HD 209458 (Henry et al. 2000; Charbonneau et al. 2000).

ture transit search. These observations, combined with Keck radial velocities from the California Planet Survey (CPS) consortium from 2006 onward, revealed that there was substantial curvature to the radial velocity residuals to the original Cochran et al. (2004) solution. In October 2010 monitoring was intensified at HET and at Keck Observatory by A.W.H., G.W.M., J.T.W., and H.I., and with these new RV data and the previously published measurements from Wittenmyer et al. (2007) they obtained a preliminary solution for the outer planet. The discrepancy between the original orbital fit and the new fit (assuming one planet) was presented at the January 2011 meeting of the American Astronomical Society (Kane et al. 2011d).

6.1.4 Synthesis and Outline

In early 2011, the Texas and TERMS teams combined efforts and began joint radial velocity analysis, dynamical modeling, spectroscopic analysis, and photometric observations (Kane et al. 2012). The resulting complete two-planet orbital solution allows for a sufficiently precise transit ephemeris for the b component to be calculated for a thorough transit search. We herein report the transit exclusion of HD 37605 b and a stable dynamical solution to the system.

In § 6.2, we describe our spectroscopic observations and analysis, which provided the radial velocities and the stellar properties of HD 37605. § 6.3 details the orbital solution for the HD 37605 system, including a comparison with MCMC Keplerian fits, and our dynamical analysis. We report our photometric observations on HD 37605 and the dispositive null detection³ of non-grazing transits of HD 37605 b in § 6.4. After § 6.5, Summary and Conclusion, we present updates on $M \sin i$ of two previously published systems (HD 114762 and HD 168443) in § 6.6. In the Appendix we describe the algorithm used in the package BOOTTRAN (for calculating orbital parameter error bars; see § 6.3.2).

6.2 Spectroscopic Observations and Analysis

6.2.1 HET and Keck Observations

Observations on HD 37605 at HET started December of 2003. In total, 101 RV observations took place over the course of almost eight years, taking advantage of the queue scheduling capabilities of HET. The queue scheduling of HET allows for small amounts of telescope time to be optimally used throughout the year, and for new observing priorities to be implemented immediately, rather than on next allocated night or after TAC and scheduling process (Shetrone et al. 2007). The observations were taken through the High Resolution Spectrograph (HRS; Tull 1998) situated at the basement of the HET building. This fiber-fed spectrograph has a typical long-term Doppler error of 3 – 5 m/s (Baluev 2009). The observations were taken with the spectrograph configured at a resolving power of $R = 60,000$. For more details, see Cochran et al. (2004).

³A dispositive null detection is one that disposes of the question of whether an effect is present, as opposed to one that merely fails to detect a purported or hypothetical effect that may yet lie beneath the detection threshold. The paragon of dispositive null detections is the Michelson-Morley demonstration that the luminiferous ether does not exist (Michelson & Morley 1887).

Observations at Keck were taken starting August 2006. A set of 33 observations spanning over five years were made through the HIRES spectrometer (Vogt et al. 1994) on the Keck I telescope, which has a long-term Doppler error of 0.9 – 1.5 m/s (e.g. Howard et al. 2009). The observations were taken at a resolving power of $R = 55,000$. For more details, see Howard et al. (2009) and Valenti et al. (2009).

Both our HET and Keck spectroscopic observations were taken with an iodine cell placed in the light path to provide wavelength standard and information on the instrument response function⁴ (IRF) for radial velocity extraction (Marcy & Butler 1992; Butler et al. 1996b). In addition, we also have observations taken without iodine cell to produce stellar spectrum templates – on HET and Keck, respectively. The stellar spectrum templates, after being deconvolved with the IRF, are necessary for both radial velocity extraction and stellar property analysis. The typical working wavelength range for this technique is roughly 5000 Å– 6000 Å.

6.2.2 Data Reduction and Doppler Analysis

In this section, we describe our data reduction and Doppler analysis of the HET observations. We reduced the Keck data with the standard CPS pipeline, as described in, for example, Howard et al. (2011) and Johnson et al. (2011a).

We have constructed a complete pipeline for analyzing HET data – from raw data reduction to radial velocity extraction. The raw reduction is done using the REDUCE package by Piskunov & Valenti (2002). This package is designed to optimally extract echelle spectra from 2-D images (Horne 1986). Our pipeline corrects for cosmic rays and scattered light. In order to make the data reduction process completely automatic, we have developed our own algorithm for tracing the echelle orders of HRS and replaced the original semi-automatic algorithm from the REDUCE package.

After the raw data reduction, the stellar spectrum template is deconvolved using IRF derived from an iodine flat on the night of observation. There were two deconvolved stellar spectrum templates (DSST) derived from HET/HRS observations and one from Keck/HIRES. Throughout this work, we use the Keck DSST, which is of better quality thanks to a better known IRF of HIRES and a superior deconvolution algorithm in the CPS pipeline (Howard et al. 2009, 2011).

Then the pipeline proceeds with barycentric correction and radial velocity extraction for each observation. We have adopted the Doppler code from CPS (e.g. Howard et al. 2009, 2011; Johnson et al. 2011a). The code is tailored to be fully functional with HET/HRS-formatted spectra, and it is capable of working with either an HET DSST or a Keck one.

The 101 HET RV observations include 44 observations which produced the published velocities in Cochran et al. (2004) and Wittenmyer et al. (2007), 34 observations also done by the Texas team in follow-up work after 2007, and 23 observations taken as part of TERMS program. We have performed re-reduction on these 44 observations together with all the rest 57 HET observations through our pipeline. This has the advantage

⁴Some authors refer to this as the “point spread function” or the “instrumental profile” of the spectrograph.

of eliminating one free parameter in the Keplerian fit – the offset between two Doppler pipelines.

Two out of the 101 HET observations were excluded due to very low average signal-to-noise ratio per pixel (< 20), and one observation taken at twilight was also rejected as such observation normally results in low accuracy due to the significant contamination by the residual solar spectrum (indeed this velocity has a residual of over 100 m/s against best Keplerian fit, much larger than the ~ 8 m/s RV error).

All the HET and Keck radial velocities used in this work (98 from HET and 33 from Keck) are listed in Table 6.7.

6.2.3 Stellar Analysis

HD 37605 is a K0 V star ($V \sim 8.7$) with high proper motion at a distance of 44.0 ± 2.1 pc (ESA 1997; van Leeuwen 2008). We derived its stellar properties based on analysis on a high-resolution spectrum taken with Keck HIRES (without iodine cell in the light path). Table 6.1 lists the results of our analysis⁵, including the effective temperature T_{eff} , surface gravity $\log g$, iron abundance [Fe/H], projected rotational velocity $v \sin i$, bolometric correction BC, bolometric magnitude M_{bol} , stellar luminosity L_{\star} , stellar radius R_{\star} , stellar mass M_{\star} and age. HD 37605 is found to be a metal rich star ([Fe/H] ~ 0.34) with $M_{\star} \sim 1.0 M_{\odot}$ and $R_{\star} \sim 0.9 R_{\odot}$.

We followed the procedure described in Valenti & Fischer (2005) and also in Valenti et al. (2009) with improvements. Briefly, the observed spectrum is fitted with a synthetic spectrum using Spectroscopy Made Easy (SME; Valenti & Piskunov 1996) to derive T_{eff} , $\log g$, [Fe/H], $v \sin i$, and so on, which are used to derive the bolometric correction BC and L_{\star} consequently. Then an isochrone fit by interpolating tabulated Yonsei-Yale isochrones (Demarque et al. 2004) using derived stellar parameters from SME is performed to calculate M_{\star} and $\log g_{\text{iso}}$ values (along with age and stellar radius). Next, Valenti et al. (2009) introduced an outside loop which re-runs SME with $\log g$ fixed at $\log g_{\text{iso}}$, followed by another isochrone fit deriving a new $\log g_{\text{iso}}$ using the updated SME results. The loop continues until $\log g$ values converge. This additional iterative procedure to enforce self-consistency on $\log g$ is shown to improve the accuracy of other derived stellar parameters (Valenti et al. 2009). The stellar radius and $\log g$ reported here in Table 6.1 are derived from the final isochrone fit, which are consistent with the purely spectroscopic results. The gravity ($\log g = 4.51$) is also consistent with the purely spectroscopic gravity (4.44) based on strong Mg b damping wings, so for HD 37605 the iteration process is optional.

Cochran et al. (2004) reported the values of T_{eff} , $\log g$, and [Fe/H] for HD 37605, and their estimates agree with ours within 1σ uncertainty. Santos et al. (2005) also estimated T_{eff} , $\log g$, [Fe/H], and M_{\star} , all of which agree with our values within 1σ . Our stellar mass and radius estimates are also consistent with the ones derived from the empirical method by Torres et al. (2010).

⁵Note that the errors on the stellar radius R_{\star} and mass M_{\star} listed in Table 6.1 are not intrinsic to the SME code, but are $5\% \times R_{\star}$ and $5\% \times M_{\star}$. This is because the intrinsic errors reported by SME do not include the errors stemming from the adopted stellar models, and a more realistic precision for R_{\star} and M_{\star} would be around $\sim 5\%$. Intrinsic errors reported by SME are $0.015 L_{\odot}$ for R_{\star} and $0.017 M_{\odot}$ for M_{\star} .

Our SME analysis indicates that the rotation of the star ($v \sin i$) is likely < 1 km/s (corresponding to rotation period $\gtrsim 46$ days). We have used various methods to estimate stellar parameters from the spectrum, including the incorporation of color and absolute magnitude information and the Mg b triplet to constrain $\log g$, and various macroturbulent velocity prescriptions. All of these approaches yield results consistent with an undetectable level of rotational broadening, with an upper limit of 1-2 km/s, consistent with the tentative photometric period 57.67 days derived from the APT data (See §6.4.1).

6.3 Orbital Solution

6.3.1 Transit Ephemeris

The traditional parameters for reporting the ephemerides of spectroscopic binaries are P, K, e, ω , and T_p , the last being the time of periastron passage (?). This information is sufficient to predict the phase of a planet at any point in the future in principle, but the uncertainties in those parameters alone are insufficient to compute the uncertainty in orbital phase without detailed knowledge of the covariances among the parameters.

This problem is particularly acute when determining transit or secondary eclipse times for planets with near circular orbits, where σ_{T_p} and σ_ω can be highly covariant. In such cases the circular case is often not excluded by the data, and so the estimation of e includes the case $e = 0$, where ω is undefined. If the best or most likely value of e in this case is small but not zero, then it is associated with some nominal value of ω , but σ_ω will be very large (approaching π). Since T_p represents the epoch at which the true anomaly equals 0, T_p will have a similarly large uncertainty (approaching P), despite the fact that the phase of the system may actually be quite precisely known!

In practice even the ephemerides of planets with well measured eccentricities suffer from lack of knowledge of the covariance in parameters, in particular T_p and P (whose covariance is sensitive to the approximate epoch chosen for T_p). To make matters worse, the nature of “ 1σ ” uncertainties in the literature is inconsistent. Some authors may report uncertainties generated while holding all or some other parameters constant (for instance, by seeing at what excursion from the nominal value χ^2 is reduced by 1), while others using bootstrapping or MCMC techniques may report the variance in a parameter over the full distribution of trials. In any case, covariances are rarely reported, and in some cases authors even report the most likely values on a parameter-by-parameter basis rather than a representative “best fit”, resulting in a set of parameters that is not self-consistent.

The TERMS strategy for refining ephemerides therefore begins with the recalculation of transit time uncertainties directly from the archival radial velocity data. We used bootstrapping (see Appendix) with the time of conjunction, T_c (equivalent to transit center, in the case of transiting planets) computed independently for each trial. For systems whose transit time uncertainty makes definitive observations implausible or impossible due to the accumulation of errors in phase with time, we sought additional RV measurements to “lock down” the phase of the planet.

6.3.2 The 37605 System

There are in total 137 radial velocities used in the Keplerian fit for the HD 37605 system. In addition to the 98 HET velocities and 33 Keck ones (see §6.2.2), we also included six⁶ velocities from Cochran et al. (2004) which were derived from observations taken with the McDonald Observatory 2.1 m Telescope (hereafter the 2.1 m telescope).

We used the RVLIN package by ?) to perform the Keplerian fit. This package is based on the Levenberg–Marquardt algorithm and is made efficient in searching parameter space by exploiting the linear parameters. The uncertainties of the parameters are calculated through bootstrapping (with 1,000 bootstrap replicates) using the BOOTTRAN package, which is described in detail in the Appendix⁷.

The best-fit Keplerian parameters are listed in Table 6.2. The joint Keplerian fit for HD 37605b and HD 37605c has 13 free parameters: the orbital period P , time of periastron passage T_p , velocity semi-amplitude K , eccentricity e , and the argument of periastron referenced to the line of nodes ω for each planet; and for the system, the velocity offset between the center of the mass and barycenter of solar system γ and two velocity offsets between the three telescopes (Δ_{Keck} and Δ_{HET} , with respect to the velocities from the 2.1 m telescope as published in Cochran et al. 2004). We did not include any stellar jitter or radial velocity trend in the fit (i.e., fixed to zero). The radial velocity signals and the best Keplerian fits for the system, HD 37605b only, and HD 37605c only are plotted in the three panels of Fig. 6.1, respectively.

Adopting a stellar mass of $M_\star = 1.000 \pm 0.017 M_\odot$ (as in Table 6.1), we estimated the minimum mass ($M \sin i$) for HD 37605b to be $2.802 \pm 0.011 M_{\text{Jup}}$ and $3.366 \pm 0.072 M_{\text{Jup}}$ for HD 37605c. While HD 37605b is on a close-in orbit at $a = 0.2831 \pm 0.0016$ AU that is highly eccentric ($e = 0.6767 \pm 0.0019$), HD 37605c is found to be on a nearly circular orbit ($e = 0.013 \pm 0.015$) out at $a = 3.814 \pm 0.058$ AU, which qualifies it as one of the “Jupiter analogs”.

In order to see whether the period and mass of the outer planet, HD 37605c, are well constrained, we mapped out the χ^2_ν values for the best Keplerian fit in the P_c - $M_c \sin i$ space (subscript ‘ c ’ denoting parameters for the outer planet, HD 37605c). Each χ^2_ν value on the P_c - $M_c \sin i$ grid was obtained by searching for the best-fit model while fixing the period P_c for the outer planet and requiring constraints on K_c and e_c to maintain $M \sin i$ fixed. As shown in Fig. 6.2, our data are sufficient to have both P_c and $M_c \sin i$ well-constrained. This is also consistent with the tight sampling distributions for P_c and $M_c \sin i$ found in our bootstrapping results.

The rms values against the best Keplerian fit are 7.86 m/s for HET, 2.08 m/s for Keck, and 12.85 m/s for the 2.1 m telescope. In the case of HET and Keck, their rms values are slightly larger than their typical reported RV errors (~ 5 m/s and ~ 1 m/s, respectively). This might be due to stellar jitter or underestimated systematic errors in the velocities. We note that the χ^2_ν is reduced to 1.0 if we introduce a stellar jitter of 3.6 m/s (added in quadrature to all the RV errors).

⁶The velocity from observation on BJD 2,453,101.6647 was rejected as it was from a twilight observation, which had both low precision ($\sigma_{\text{RV}} = 78.12$ m/s) and low accuracy (having a residual against the best Keplerian fit of over 100 m/s).

⁷The BOOTTRAN package is made publicly available online at <http://exoplanets.org/code/> and the Astrophysics Source Code Library.

6.3.3 Comparison with MCMC Results

We compared our best Keplerian fit from RVLIN and uncertainties derived from BOOTTRAN (abbreviated as RVLIN+BOOTTRAN hereafter) with that from a Bayesian framework following Ford (2005) and Ford (2006) (referred to as the MCMC analysis hereafter). Table 6.3 lists the major orbital parameters from both methods for a direct comparison. Fig. 6.3 illustrates this comparison, but with the MCMC results presented in terms of 2-D confidence contours for P , e , K , $M \sin i$, and ω of both planets, as well as for T_c of HD 37605b.

For the Bayesian analysis, we assumed priors that are uniform in log of orbital period, eccentricity, argument of pericenter, mean anomaly at epoch, and the velocity zero-point. For the velocity amplitude (K) and jitter (σ_j), we adopted a prior of the form $p(x) = (x+x_o)^{-1}[\log(1+x/x_o)]^{-1}$, with $K_o = \sigma_{j,o} = 1$ m/s, i.e. high values are penalized. For a detailed discussion of priors, strategies to deal with correlated parameters, the choice of the proposal transition probability distribution function, and other details of the algorithm, we refer the reader to the original papers: Ford (2005, 2006); Ford & Gregory (2007). The likelihood for radial velocity terms assumes that each radial velocity observation (v_i) is independent and normally distributed about the true radial velocity with a variance of $\sigma_i^2 + \sigma_j^2$, where σ_i is the published measurement uncertainty. σ_j is a jitter parameter that accounts for additional scatter due to stellar variability, instrumental errors and/or inaccuracies in the model (i.e., neglecting planet-planet interactions or additional, low amplitude planet signals).

We used an MCMC method based upon Keplerian orbits to calculate a sample from the posterior distribution (Ford 2006). We calculated 5 Markov chains, each with $\sim 2 \times 10^8$ states. We discarded the first half of the chains and calculate Gelman-Rubin test statistics for each model parameter and several ancillary variables. We found no indications of non-convergence amongst the individual chains. We randomly drew 3×10^4 solutions from the second half of the Markov chains, creating a sample set of the converged overall posterior distribution of solutions. We then interrogated this sample on a parameter-by-parameter basis to find the median and $68.27\% (1\sigma)$ values reported in Table 6.3. We refer to this solution set below as the “best-fit” MCMC solutions.

We note that the periods of the two planets found in this system are very widely separated ($P_c/P_b \sim 50$), so we do not expect planet-planet interactions to be strong, hence we have chosen to forgo a numerically intensive N-body DEMCMC fitting procedure (see e.g. Johnson et al. 2011b; Payne & Ford 2011) as the non-Keplerian perturbations should be tiny (detail on the magnitude of the perturbations is provided in §6.3.4). However, to ensure that the Keplerian fits generated are stable, we took the results of the Keplerian MCMC fits and injected those systems into the Mercury n-body package (Chambers 1999) and integrated them forward for $\sim 10^8$ years. This allows us to verify that all of the selected best-fit systems from the Keplerian MCMC analysis are indeed long-term stable. Further details on the dynamical analysis of the system can be found in §6.3.4.

We assumed that all systems are coplanar and edge-on for the sake of this analysis, hence all of the masses used in our n-body analyses are minimum masses.

As shown in Table 6.3 and Fig. 6.3, the parameter estimates from RVLIN+BOOTTRAN and MCMC methods agree with each other very well (all within 1σ error bar). In some

cases, the MCMC analysis reports error bars slightly larger than bootstrapping method ($\sim 20\%$ at most). We note that the relatively large MCMC confidence intervals are not significantly reduced if one conducts an analysis at a fixed jitter level (e.g. $\sigma_J = 3.5\text{m/s}$) unless one goes to an extremely low jitter value (e.g. $\sim 1.5\text{m/s}$). That is, the larger MCMC error bars do not simply result from treating the jitter as a free parameter. For the uncertainties on minimum planet mass $M \sin i$ and semi-major axes a , the MCMC analysis does not incorporate the errors on the stellar mass estimate. Note here, as previously mentioned in § 6.3.1, that the “best-fit” parameters reported by the MCMC analysis here listed in Table 6.3 are not a consistent set, as the best estimates were evaluated on a parameter-by-parameter basis, taking the median from marginalized posterior distribution of each. Assuming no jitter, The best Keplerian fit from RVLIN has a reduced chi-square value $\chi^2_\nu = 2.28$, while the MCMC parameters listed in Table 6.3 give a higher χ^2_ν value of 2.91.

6.3.4 Dynamical Analysis

We used the best-fit Keplerian MCMC parameters as the basis for a set of long-term numerical (n-body) integrations of the HD 37605 system using the Mercury integration package (Chambers 1999). We used these integrations to verify that the best-fit systems: (i) are long-term stable; (ii) do not exhibit significant variations in their orbital elements on the timescale of the observations (justifying the assumption that the planet-planet interactions are negligible); (iii) do not exhibit any other unusual features. We emphasize again that the planets in this system are well separated and we do not expect any instability to occur: for the masses and eccentricities in question, a planet at $a_b \sim 0.28$ AU will have companion orbits which are Hill stable for $a \gtrsim 0.83$ AU (Gladman 1993), so while Hill stability does not preclude outward scatter of the outer planet, the fact that $a_c \sim 3.8 \gg 0.83$ AU suggests that the system will be far from any such instability.

We integrated the systems for $> 10^8$ years ($\sim 10^7 \times$ the orbital period of the outer planet and $> 10^2 \times$ the secular period of the system), and plot in Fig. 6.4 the evolution of the orbital elements a , e , & ω . On the left-hand side of the plot we provide short-term detail, illustrating that over the ~ 10 year time period of our observations, the change in orbital elements will be very small. On the right-hand side we provide a much longer-term view, plotting 10^7 out of $> 10^8$ years of system evolution, demonstrating that (i) the secular variation in some of the elements (particularly the eccentricity of the outer planet; see e_c in red) over a time span of $\sim 4 \times 10^5$ years can be significant: in this case we see $0.03 < e_c < 0.11$, but (ii) the system appears completely stable, as one would expect for planets with a period ratio $P_c/P_b \sim 50$. Finally, at the bottom of the figure we display the range of parameter space covered by the $e_i \cos \omega_i$, $e_i \sin \omega_i$ parameters ($i = b$ in blue for inner planet and $i = c$ in red for outer planet), demonstrating that the orbital alignments circulate, i.e. they do not show any signs of resonant confinement, which confirms our expectation of minimal planet-planet interaction as mentioned before.

As noted above, our analysis assumed coplanar planets. As such the planetary masses used in these dynamical simulations are minimum masses. We note that for inclined systems, the larger planetary masses will cause increased planet-planet perturbations.

To demonstrate this is still likely to be unimportant, we performed a 10^8 year simulation of a system in which $1/\sin i = 10$, pushing the planetary masses to $\sim 30 M_{\text{Jup}}$. Even in such a pathological system the eccentricity oscillations are only increased by a factor of ~ 2 and the system remains completely stable for the duration of the simulation.

We also performed a separate Transit Timing Variation (TTV) analysis, using the best-fit MCMC systems as the basis for a set of highly detailed short-term integrations. From these we extracted the times of transit and found a TTV signal ~ 100 s, or ~ 0.001 day, which is much smaller than the error bar on T_c (~ 0.07 day). Therefore we did not take into account the effect of TTV when performing our transit analysis in the next section.

6.3.5 Activity Cycles and Jupiter Analogs

The coincidence of the Solar activity cycle period of 11 years and Jupiter's orbital period near 12 years illustrates how activity cycles could, if they induced apparent line shifts in disk-integrated stellar spectra, confound attempts to detect Jupiter analogs around Sun-like stars. Indeed, Dravins (1985) predicted apparent radial velocity variations of up to 30 m/s in solar lines due to the Solar cycle, and Deming et al. (1987) reported a tentative detection of such a signal in NIR CO lines of 30 m/s in just 2 years, and noted that such an effect would severely hamper searches for Jupiter analogs. That concern was further amplified when Campbell et al. (1991) reported a positive correlation between radial velocity and chromospheric activity in the active star κ^1 Cet, with variations of order 50–100 m/s.

Wright et al. (2008) found that the star HD 154345 has an apparent Jupiter analog (HD 154345 *b*), but that this star also shows activity variations in phase with the radial velocity variations. They noted that many Sun-like stars, including the precise radial velocity standard star HD 185144 (σ Dra) show similar activity variations and that rarely, if ever, are these signals well-correlated with signals similar in strength to that seen in HD 154345 (~ 15 m/s), and concluded that the similarity was therefore likely just an inevitable coincidence. Put succinctly, activity cycles in Sun-like stars are common (Baliunas et al. 1995), but few Jupiter analogs have been discovered, meaning that the early concern that activity cycles would mimic giant planets is not a severe problem.

Nonetheless, there is growing evidence that activity cycles can, in some stars, induce radial velocity variations, and the example of HD 154345 still warrants care and concern. Most significantly, Dumusque et al. (2011) found a positive correlation between chromospheric activity and precise radial velocity in the average measurements of a sample of HARPS stars, and provided a formula for predicting the correlation strength as a function of the metallicity and effective temperature of the star. Their formulae predict a value of 2 m/s for the most suspicious case in the literature, HD 154345 (compared to an actual semiamplitude of ~ 15 m/s), but are rather uncertain. It is possible that in a few, rare cases, the formula might significantly underestimate the amplitude of the effect.

The top panel of Fig. 6.5 plots the T12 APT observations from all five observing seasons (data provided in Table 6.4; see details on APT photometry in § 6.4.1). The dashed line marks the mean relative magnitude ($\Delta(b + y)/2$) of the first season. The

seasonal mean brightness of the star increases gradually from year to year by a total of ~ 0.002 mag, which may be due to a weak long-term magnetic cycle. However, no evidence is found in support of such a cycle in the Mount Wilson chromospheric Ca II H & K indices (Isaacson & Fischer 2010), although the S values vary by approximately 0.1 over the span of a few years. The formulae of Lovis et al. (2011) predict a corresponding RV variation of less than 2 m/s due to activity, far too small to confound our planet detection with $K = 49$ m/s.

Since we do not have activity measurements for this target over the span of the outer planet's orbit in HD 37605, we cannot definitively rule out activity cycles as the origin of the effect, but the strength of the outer planetary signal and the lack of such signals in other stars known to cycle strongly dispels concerns that the longer signal is not planetary in origin.

6.4 The Dispositive Null Detection of Transits of HD 37605 b

We have performed a transit search for the inner planet of the system, HD 37605 b . This planet has a transit probability of 1.595% and a predicted transit duration of 0.352 day, as derived from the stellar parameters listed in Table 6.1 and the orbital parameters given in Table 6.2. From the minimum planet mass ($M \sin i = 2.802 \pm 0.011 M_{\text{Jup}}$; see Table 6.2) and the models of Bodenheimer et al. (2003), we estimate its radius to be $R_p = 1.1 R_{\text{Jup}}$. Combined with the stellar radius of HD 37605 listed in Table 6.1, $R_* = 0.901 \pm 0.015 R_{\odot}$, we estimate the transit depth to be 1.877% (for an edge-on transit, $i = 90^\circ$). We used both ground-based (APT; §6.4.1) and space-based (MOST; §6.4.2) facilities in our search.

6.4.1 APT Observations and Analysis

The T12 0.8-m Automatic Photoelectric Telescope (APT), located at Fairborn Observatory in southern Arizona, acquired 696 photometric observations of HD 37605 between 2008 January 16 and 2012 April 7. Henry (1999) provides detailed descriptions of observing and data reduction procedures with the APTs at Fairborn. The measurements reported here are differential magnitudes in $\Delta(b + y)/2$, the mean of the differential magnitudes acquired simultaneously in the Strömgren b and y bands with two separate EMI 9124QB bi-alkali photomultiplier tubes. The differential magnitudes are computed from the mean of three comparison stars: HD 39374 ($V = 6.90$, $B-V = 0.996$, K0 III), HD 38145 ($V = 7.89$, $B-V = 0.326$, F0 V), and HD 38779 ($V = 7.08$, $B-V = 0.413$, F4 IV). This improves the precision of each individual measurement and helps to compensate for any real microvariability in the comp stars. Intercomparison of the differential magnitudes of these three comp stars demonstrates that all three are constant to 0.002 mag or better from night to night, consistent with typical single-measurement precision of the APT (0.0015–0.002 mag; Henry 1999).

Fig. 6.5 illustrates the APT photometric data and our transit search. As mentioned in § 6.3.5, the top panel shows all of our APT photometry covering five observing seasons, which exhibits a small increasing trend in the stellar brightness. To search for the transit signal of HD 37605 b , the photometric data were normalized so that all five seasons had

the same mean (referred to as the “normalized photometry” hereafter). The data were then phased at the orbital period of HD 37605b, 55.01307 days, and the predicted time of mid-transit, T_c , defined as Phase 0. The normalized and phased data are plotted in the middle panel of Fig. 6.5. The solid line is the predicted transit light curve, with the predicted transit duration (0.352 day or 0.0064 phase unit) and transit depth (1.877% or ~ 0.020 mag) as estimated above. The scatter of the phased data from their mean is 0.00197 mag, consistent with APT’s single-measurement precision, and thus demonstrates that the combination of our photometric precision and the stability of HD 37605 is easily sufficient to detect the transits of HD 37605b in our phased data set covering five years. A least-squares sine fit of the phased data gives a very small semi-amplitude of 0.00031 ± 0.00011 mag (consistent with zero) and so provides strong evidence that the observed radial-velocity variations are not produced by rotational modulation of surface activity on the star.

The bottom panel of Fig. 6.5 plots the phased data around the predicted time of mid-transit, T_c , at an expanded scale on the abscissa. The horizontal bar below the transit window represents the $\pm 1\sigma$ uncertainty on T_c (0.138 day or 0.0025 phase unit for T_c ’s near BJD 2,455,901.361; see § 6.3.2). The light curve appears to be highly clustered, or binned, due to the near integral orbital period ($P \sim 55.01$ days) and consequent incomplete sampling from a single observing site. Unfortunately, none of the data clusters chance to fall within the predicted transit window, so we are unable to rule out transits of HD 37605b with the APT observations.

Periodogram analysis of the five individual observing seasons revealed no significant periodicity between 1 and 100 days. This suggests that the star is inactive and the observed $K \sim 200$ m/s RV signal (for HD 37605b) is unlikely to be the result of stellar activity.

Analysis of the complete, normalized data set, however, suggests a weak periodicity of 57.67 ± 0.30 days with a peak-to-peak amplitude of just 0.0012 ± 0.0002 mag (see Fig. 6.6). We tentatively identify this as the stellar rotation period. This period is consistent with the projected rotational velocity of $v \sin i < 1$ km/s derived from our stellar analysis described in §6.2.3. It is also consistent with the analysis of Isaacson & Fischer (2010), who derived a Mount Wilson chromospheric Ca II H & K index of $S = 0.165$, corresponding to $\log R'_{HK} = -5.03$. Together, these results imply a rotation period $\gtrsim 46$ days and an age of ~ 7 Gyr (see Table 6.1). Similarly, Ibukiyama & Arimoto (2002) find an age of > 10 Gyr using isochrones along with the Hipparcos parallax and space motion, supporting HD 37605’s low activity and long rotation period.

6.4.2 MOST Observations and Analysis

As noted earlier, the near-integer period of HD 37605b makes it difficult to observe from a single longitude. The brightness of the target and the relatively long predicted transit duration creates additional challenges for ground-based observations. We thus observed HD 37605 during 2011 December 5–6 (around the predicted T_c at BJD 2,455,901.361 as listed in Table 6.2) with the MOST (Microvariability and Oscillations of Stars) satellite launched in 2003 (Walker et al. 2003; Matthews et al. 2004) in the Direct Imaging mode. This observing technique is similar to ground-based CCD photometry, allowing to apply

traditional aperture and PSF procedures for data extraction (see e.g. Rowe et al. 2006, for details). Outlying data points caused by, e.g., cosmic rays were removed.

MOST is orbiting with a period of ~ 101 minutes (14.19 cycles per day, cd^{-1}), which leads to a periodic artifact induced by the scattered light from the earthshine. This signal and its harmonics are further modulated with a frequency of 1 cd^{-1} originating from the changing albedo of the earth. To correct for this phenomenon, we constructed a cubic fit between the mean background and the stellar flux, which was then subtracted from the data. The reduced and calibrated MOST photometric data are listed in Table 6.5.

The MOST photometry is shown in Figure 6.7 for the transit window observations. The vertical dashed lines indicate the beginning and end of the 1σ transit window defined by adding σ_{T_c} (0.069 day) on both sides of the predicted transit duration of 0.352 days. The solid line shows the predicted transit model for the previously described planetary parameters. The rms scatter of the photometry is 0.17%, and within the predicted transit window there are 58 MOST observations. Therefore, the standard error on the mean relative photometry (which is measured to be 0.00%) is $0.17\%/\sqrt{58} = 0.022\%$. This means that, for the predicted transit window and a predicted depth of 1.877%, we can conclude a null detection of HD 37605b’s transit with extremely high confidence (149σ).

Note that the above significance is for an edge-on transit with an impact parameter of $b = 0.0$. A planetary trajectory across the stellar disk with a higher impact parameter will produce a shorter transit duration. However, the gap between each cluster of MOST measurements is 0.06 days which is 17% of the edge-on transit duration. In order for the duration to be fit within the data gaps, the impact parameter would need to be $b > 0.996$. To estimate a more conservative lower limit for b , we now assume the most unfortunate case where the transit center falls exactly in the middle of one of the measurement gaps, and also consider the effect of limb darkening by using the non-linear limb darkening model by Mandel & Agol (2002) with their fitted coefficients for HD 209458. Even under this scenario, we can still conclude the null detection for any transit with $b < 0.951$ at $\gtrsim 5\sigma$ (taking into account that there are at least ~ 20 observations will fall within the transit window in this case, though only catching the shallower parts of the transit light curve).

All of the above is based on the assumption that the planet has the predicted radius of $1.1 R_{\text{Jup}}$. If in reality the planet is so small that even a $b = 0$ transit would fall below our detection threshold, it would mean that the planet has a radius of $< 0.36 R_{\text{Jup}}$ (a density of $> 74.50 \text{ g/cm}^3$), which seems unlikely. It is also very unlikely that our MOST photometry has missed the transit window completely due to an ill-predicted T_c . In the sampling distribution of T_c from BOOTTRAN (with 1000 replicates; see § 6.3.2 and Appendix), there is no T_c that would put the transit window completely off the MOST coverage. In the marginalized posterior distribution of T_c calculated via MCMC (see § 6.3.3 and Fig. 6.3), there is only 1 such T_c out of 3×10^4 (0.003%).

6.5 Summary and Conclusion

In this paper, we report the discovery of HD 37605c and the dispositive null detection of non-grazing transits of HD 37605b, the first planet discovered by HET. HD 37605c is

the outer planet of the system with a period of ~ 7.5 years on a nearly circular orbit ($e = 0.013$) at $a = 3.814$ AU. It is a “Jupiter analog” with $M \sin i = 3.366 M_{\text{Jup}}$, which adds one more sample to the currently still small inventory of such planets (only 10 including HD 37605c; see §6.1). The discovery and characterization of “Jupiter analogs” will help understanding the formation of gas giants as well as the frequency of true solar system analogs. This discovery is a testimony to the power of continued observation of planet-bearing stars.

Using our RV data with nearly 8-year long baseline, we refined the orbital parameters and transit ephemerides of HD 37605b. The uncertainty on the predicted mid-transit time was constrained down to 0.069 day (at and near $T_c = 2,455,901.361$ in BJD), which is small compared to the transit duration (0.352 day). In fact, just the inclusion of the two most recent points in our RV data have reduced the uncertainty on T_c by over 10%. We have performed transit search with APT and the MOST satellite. Because of the near-integer period of HD 37605b and the longitude of Fairborn Observatory, the APT photometry was unable to cover the transit window. However, its excellent photometric precision over five observing seasons enabled us to rule out the possibility of the RV signal being induced by stellar activity. The MOST photometric data, on the other hand, were able to rule out an edge-on transit with a predicted depth of 1.877% at a $\gg 10\sigma$ level, with a 5σ lower limit on the impact parameter of $b \leq 0.951$. This transit exclusion is a further demonstration of the TERMS strategy, where follow-up RV observations help to reduce the uncertainty on transit timing and enable transit searches.

Our best-fit orbital parameters and errors from RVLIN+BOOTTRAN were found to be consistent with those derived from a Bayesian analysis using MCMC. Based on the best-fit MCMC systems, we performed dynamic and TTV analysis on the HD 37605 system. Dynamic analysis shows no sign of orbital resonance and very minimal planet-planet interaction. We derived a TTV of ~ 100 s, which is much smaller than σ_{T_c} .

We have also performed a stellar analysis on HD 37605, which shows that it is a metal rich star ($[\text{Fe}/\text{H}] = 0.336 \pm 0.030$) with a stellar mass of $M_\star = 1.000 \pm 0.017 M_\odot$ with a radius of $R_\star = 0.901 \pm 0.015$. The small variation seen in our photometric data (amplitude < 0.003 mag over the course of four years) suggests that HD 37605 is consistent as being an old, inactive star that is probably slowly rotating. We tentatively propose that the rotation period of the star is 57.67 ± 0.30 days, based on a weak periodic signal seen in our APT photometry.

6.6 Note on Previously Published Orbital Fits

In early 2012, we repaired a minor bug in the BOOTTRAN package, mostly involving the calculation and error bar estimation of $M \sin i$. As a result, the $M \sin i$ values and their errors for two previously published systems (three planets) need to be updated. They are: HD 114762b (Kane et al. 2011b), HD 168443b, and HD 168443c (Pilyavsky et al. 2011b). Table 6.6 lists the updated $M \sin i$ and error bars.

One additional system, HD 63454 (Kane et al. 2011f), was also analyzed using BOOTTRAN. However, the mass of HD 63454b is small enough compared to its host mass and thus was not affected by this change.

Table 6.1. STELLAR
PARAMETERS

Parameter	Value
Spectral type ^a	K0 V
Distance (pc) ^a	44.0 ± 2.1
V	8.661 ± 0.013
T_{eff} (K)	5448 ± 44
$\log g$	4.511 ± 0.024
[Fe/H]	0.336 ± 0.030
BC	-0.144
M_{bol}	5.301
L_{\star} (L_{\odot})	0.590 ± 0.058
R_{\star} (R_{\odot})	0.901 ± 0.045^c
M_{\star} (M_{\odot})	1.000 ± 0.050^c
$v \sin i$	< 1 km/s
Age ^b	~ 7 Gyr

^aESA (1997); van Leeuwen (2008).

^bIsaacson & Fischer (2010), see § 6.4.1.

^c5% relative errors, not the SME intrinsic errors. See footnote 5 for details.

Table 6.2. KEPLERIAN FIT PARAMETERS

Parameter	HD 37605 b	HD 37605 c
P (days)	55.01307 ± 0.00064	2720 ± 57
T_p (BJD) ^a	2453378.241 ± 0.020	2454838 ± 581
T_c (BJD) ^b	2455901.361 ± 0.069	...
K (m/s)	202.99 ± 0.72	48.90 ± 0.86
e	0.6767 ± 0.0019	0.013 ± 0.015
ω (deg)	220.86 ± 0.28	221 ± 78
$M \sin i$ (M_{Jup})	2.802 ± 0.011	3.366 ± 0.072
a (AU)	0.2831 ± 0.0016	3.814 ± 0.058
γ (m/s)		-50.7 ± 4.6
Δ_{Keck} (m/s) ^c		55.1 ± 4.7
Δ_{HET} (m/s) ^c		36.7 ± 4.7
χ_{ν}^2		2.28 (<i>d.o.f.</i> = 124)
rms (m/s)		7.61
Jitter (m/s) ^d		3.6

^aTime of Periastron passage.

^bTime of conjunction (mid-transit, if the system transits).

^cOffset with respect to the velocities from the 2.1 m telescope.

^dIf a jitter of 3.6 m/s is added in quadrature to all RV errors, χ_{ν}^2 becomes 1.0.

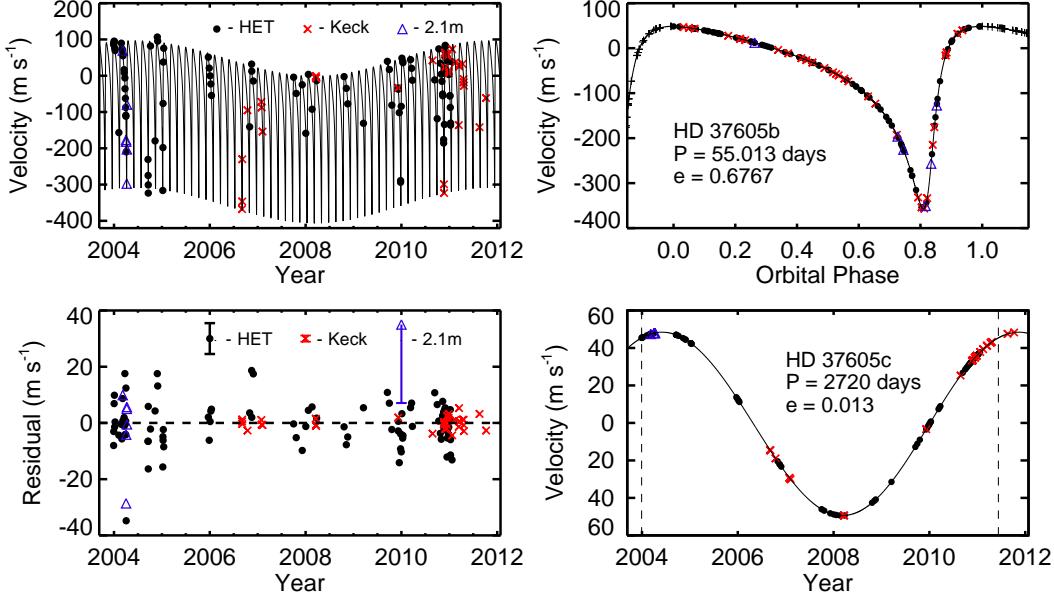


Figure 6.1 Radial velocity and Keplerian model plots for the HD 37605 system. In all panels, HET observations are labeled with black filled circles, Keck observations are labeled with red crosses, and the velocities from the 2.1 m telescope (Cochran et al. 2004) are labeled with blue triangles. Best Keplerian fits are plotted in black solid lines. **Top left:** The best-fit 2-planet Keplerian model (solid line) and the observed radial velocities from 3 telescopes. The HET and Keck velocities have been adjusted to take into account the velocity offsets (i.e., subtracting Δ_{HET} and Δ_{Keck} from the velocities, respectively; see Table 6.2 and § 6.3.2). **Bottom left:** Residual velocities after subtracting the best-fit 2-planet Keplerian model. The legend gives the typical size of the error bars using the \pm median RV error for each telescope (for 2.1 m telescope only the lower half is shown). **Top right:** RV signal induced by HD 37605b alone, phased up to demonstrate our coverage. **Bottom right:** RV signal induced by HD 37605c alone. The two vertical dashed lines denote the date of our first observation, and the date when HD 37605c closes one orbit, respectively.

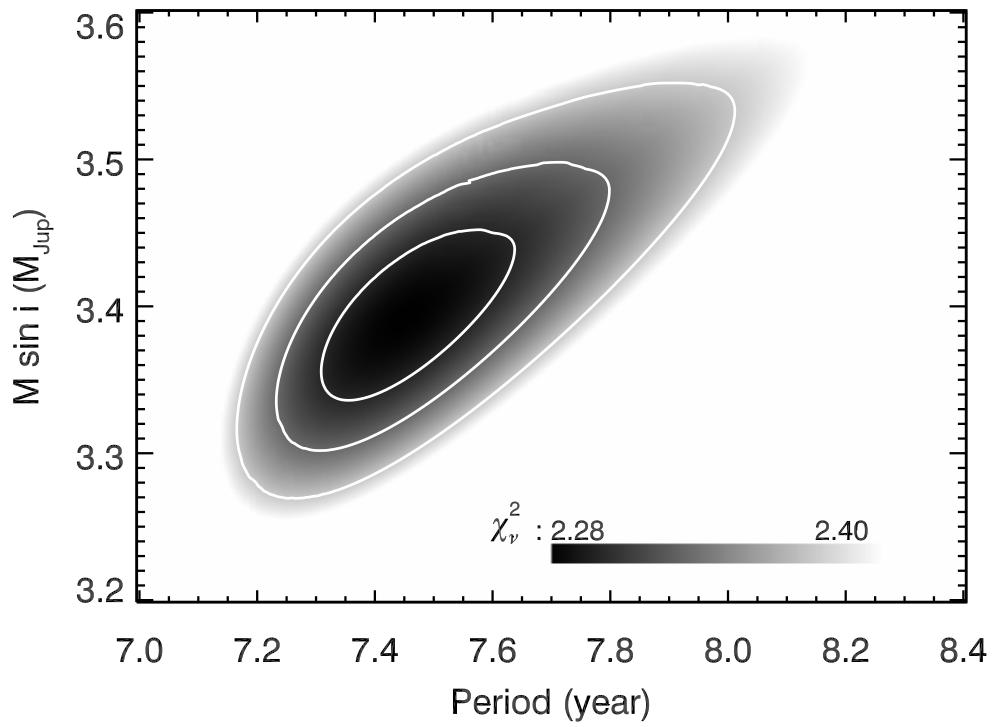


Figure 6.2 χ^2_ν map for the best Keplerian fits with fixed values of period P and minimum planet mass $M \sin i$ for HD 37605c. This is showing that both P and $M \sin i$ are well-constrained for this planet. The levels of the contours mark the 1σ (68.27%), 2σ (95.45%) and 3σ (99.73%) confidence intervals for the 2-D χ^2 distribution.

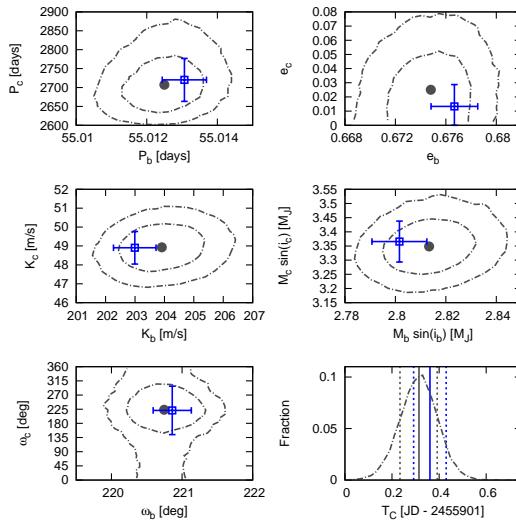


Figure 6.3 Comparison between the Bayesian (MCMC) analysis and RVLIN+BOOTTRAN results. **Top four and bottom left:** Contours of the posterior distributions of selected orbital parameters (P , e , K , $M \sin i$, and ω) based on the MCMC analysis (dashed dotted line). The x -axes are orbital parameters of the inner planet, b , and the y -axes are those of the outer planet, c . The inner contours mark the 68.27% (' 1σ ') 2-D confidence regions and the outer ones are 95.45% (' 2σ ') ones. Also plotted are the best Keplerian fit from RVLIN (blue squares) and $\pm 1\sigma$ error bars estimated via bootstrapping (blue bars). **Bottom right:** Marginalized posterior distribution of time of conjunction (mid-transit) T_c of HD 37605 b in dashed dotted line. The solid grey vertical line is the median of the distribution, and the dashed grey vertical lines mark 1σ confidence interval. The solid blue vertical line is the best estimate of T_c from RVLIN+BOOTTRAN, with $\pm 1\sigma$ error bars plotted in blue dashed vertical lines. See § 6.3.3 for details.

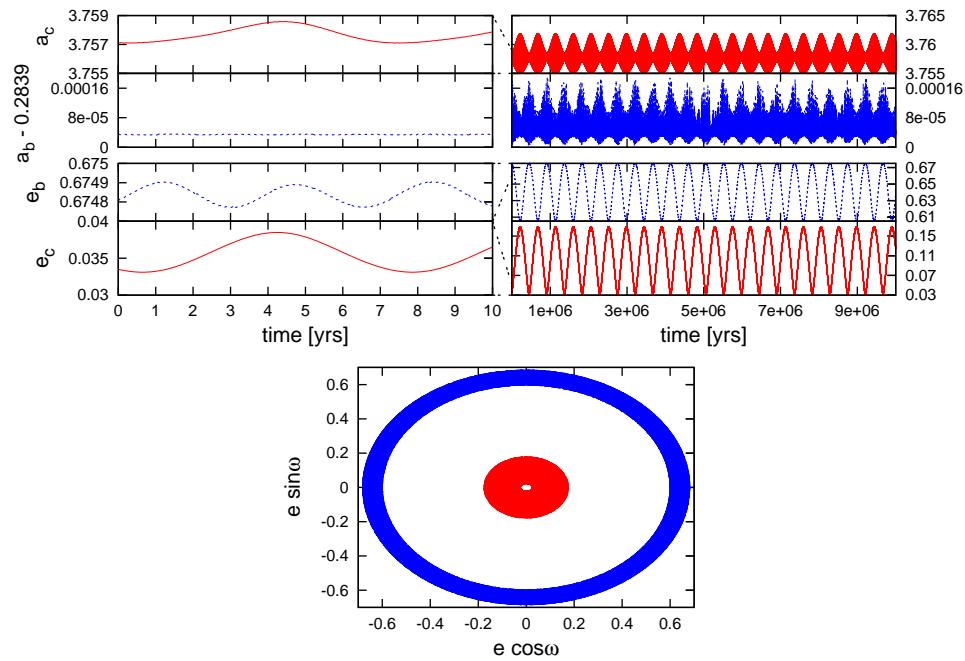


Figure 6.4 Dynamic evolution of the best-fit MCMC system. On the left we plot the short-term evolution over 10 years, on the right we plot the evolution over 10^7 years ($< 1/10$ of our dynamic simulation time scale). The top plots describe the evolution of the semi-major axes and eccentricities of the inner planet (a_b & e_b , blue lines) and the outer planet (a_c & e_c , red lines), while the bottom plot describes the parameter space covered by the $e \cos \omega, e \sin \omega$ quantities over 10^8 years (blue for inner planet and red for outer planet). We find that over the short-term (e.g., our RV observation window of ~ 10 years), the parameter variations are negligible, but in the long term significant eccentricity oscillations can take place (particularly noticeable in the eccentricity of the outer planet). See § 6.3.4 for details.

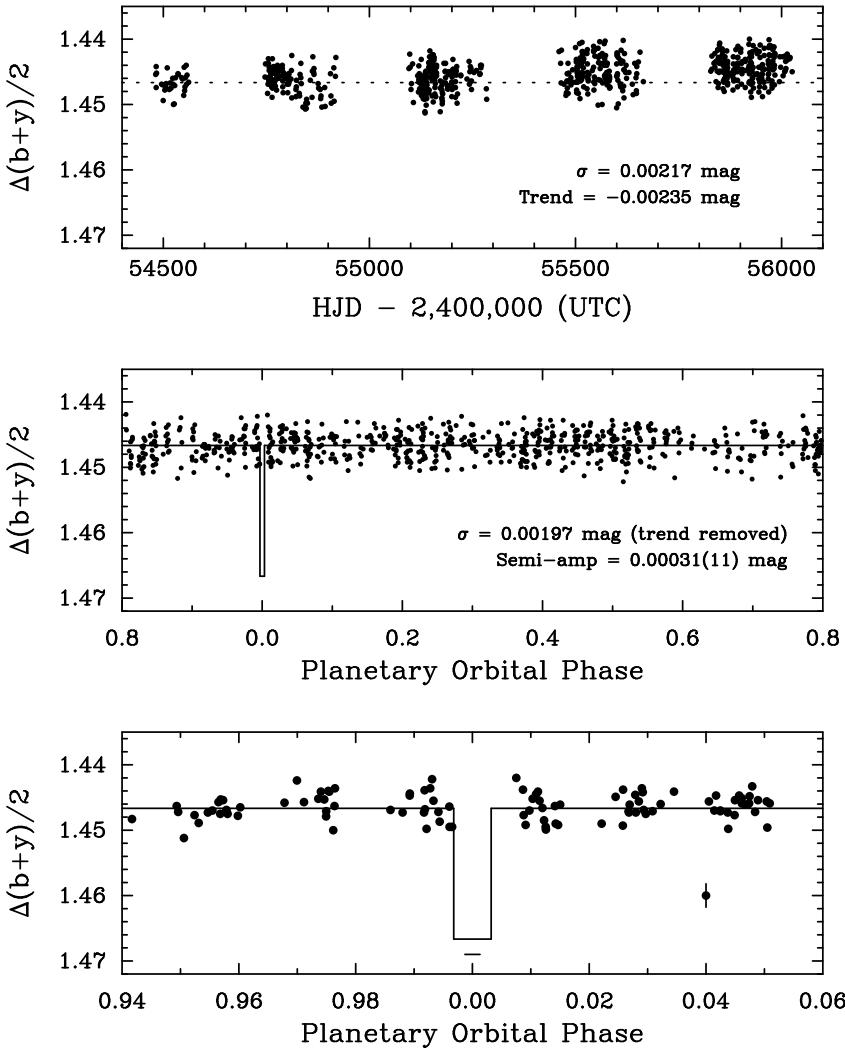


Figure 6.5 Photometric observations of HD 37605 acquired over five years with the T12 0.8m APT. The top panel shows the entire five-year data set; the dotted line represents the mean brightness of the first observing season. A long-term brightening trend is evident with a total range in the seasonal means of 0.002 mag. The middle panel shows the photometric data normalized so that each season has the same mean as the first and then phased to the orbital period of HD 37605b (55.01307 day). The solid line is the predicted transit light curve, with Phase 0.0 being the predicted time of mid-transit, T_c . A least-squares sine fit of the phased data produces the very small semi-amplitude of 0.00031 ± 0.00011 mag, providing strong evidence that the observed radial-velocity variations are not produced by rotational modulation of surface activity on the star. The bottom panel plots the observations near T_c at an expanded scale on the abscissa. The horizontal bar below the transit window represents the $\pm 1\sigma$ uncertainty in T_c . Unfortunately, none of the APT observations fall within the predicted transit window, so we are unable to rule out transits with the APT observations. See § 6.4.1 for more.

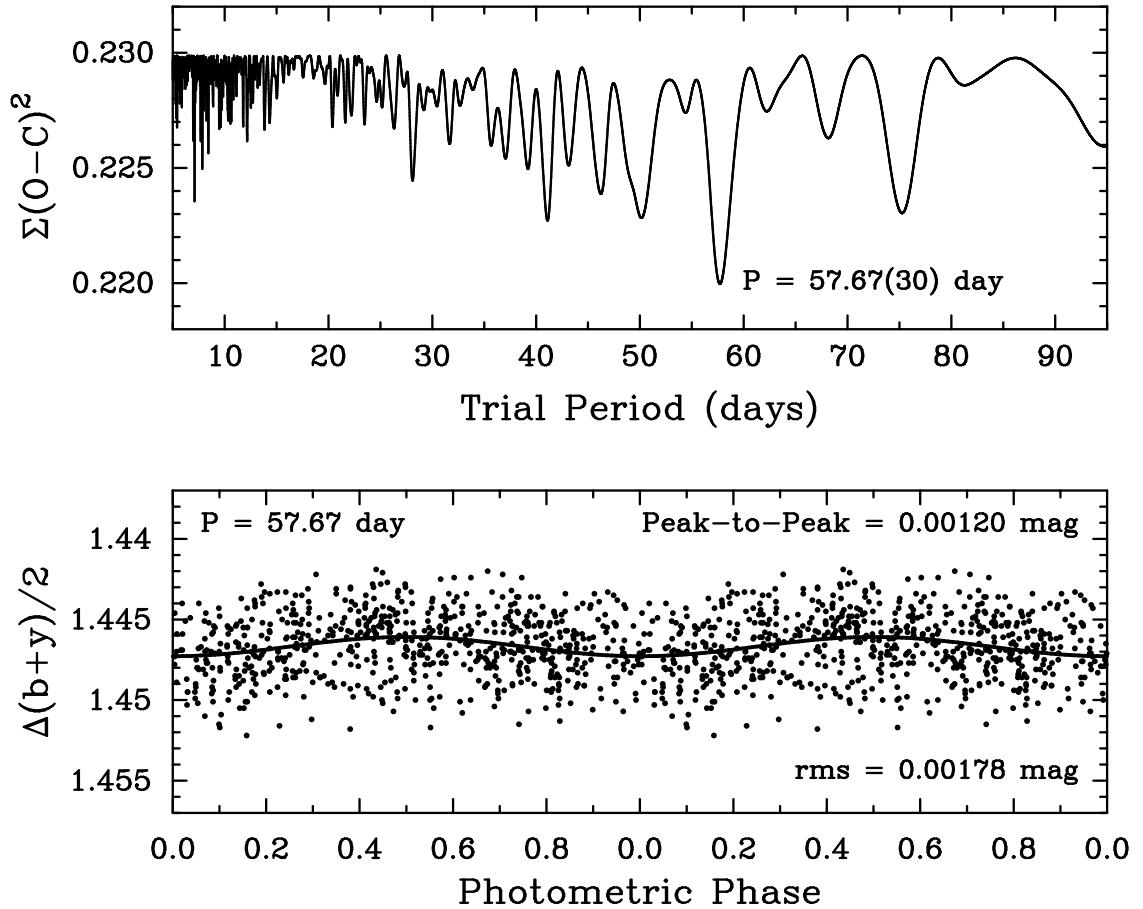


Figure 6.6 Brightness variability in HD 37605 possibly induced by stellar rotation at $P = 57.67 \pm 0.30$ days. Top panel is the periodogram of the complete, normalized data set. Bottom panel shows the normalized photometry folded with this possible rotation period. The peak-to-peak amplitude is 0.00120 ± 0.00021 mag. See § 6.4.1 for more.

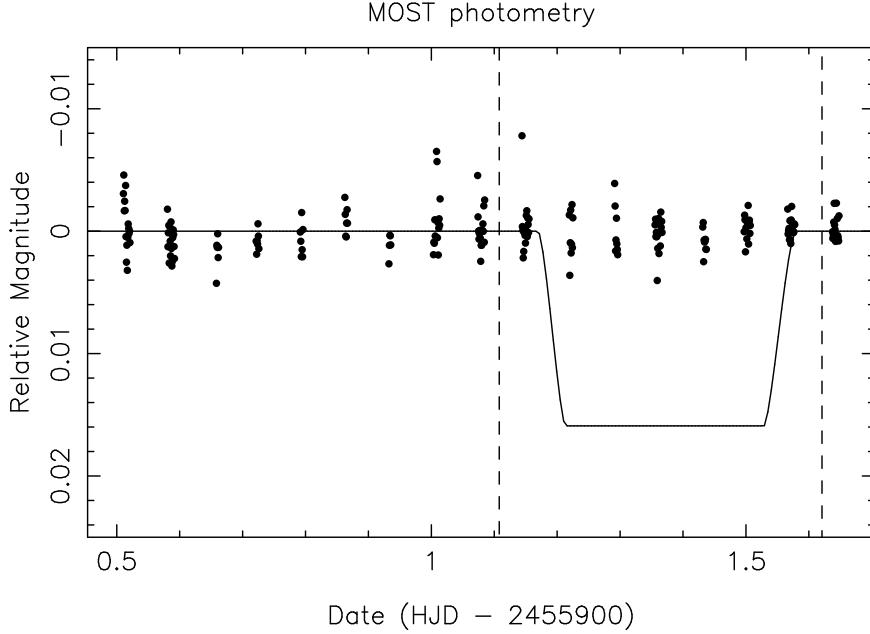


Figure 6.7 Photometric observations of HD 37605 by the MOST satellite, which rule out the edge-on transit of HD 37605 b at a $\gg 10\sigma$ level. The solid line is the predicted transit light curve, and the dashed vertical lines are the 1σ transit window boundaries defined by adding σ_{T_c} (0.069 day) on both sides of the predicted transit window (0.352-day wide). See § 6.4.2 for more details.

Table 6.3. COMPARISON WITH MCMC RESULTS

Parameter	HD 37605 b		HD 37605 c	
	RVLIN+BOOTTRAN	MCMC ^a	RVLIN+BOOTTRAN	MCMC ^a
P (days)	55.01307 ± 0.00064	$55.01250^{+0.00073}_{-0.00075}$	2720 ± 57	2707^{+57}_{-42}
T_p (BJD)	2453378.243 ± 0.020	$2453378.243^{+0.025}_{-0.024}$	2454838 ± 581	2454838^{+354}_{-435}
T_c (BJD)	2455901.361 ± 0.069	$2455901.314^{+0.077}_{-0.081}$
K (m/s)	202.99 ± 0.72	$203.91^{+0.92}_{-0.88}$	48.90 ± 0.86	$48.93^{+0.82}_{-0.82}$
e	0.6767 ± 0.0019	$0.6748^{+0.0022}_{-0.0023}$	0.013 ± 0.015	$0.025^{+0.022}_{-0.017}$
ω (deg)	220.86 ± 0.28	$220.75^{+0.33}_{-0.32}$	221 ± 78	223^{+50}_{-52}
M (deg) ^b	62.31 ± 0.15	$62.27^{+0.18}_{-0.18}$	117 ± 78	118^{+56}_{-51}
$M \sin i$ (M_{Jup})	2.802 ± 0.011	$2.814^{+0.012}_{-0.012}$	3.366 ± 0.072	$3.348^{+0.065}_{-0.062}$
a (AU)	0.2831 ± 0.0016	$0.2833364^{+0.0000027}_{-0.0000027}$	3.814 ± 0.058	$3.809^{+0.053}_{-0.040}$
Jitter (m/s) ^c	3.6	$2.70^{+0.53}_{-0.46}$		

^aMedian values of the marginalized posterior distributions and the 68.27% ('1 σ ') confidence intervals.

^bMean anomaly of the first observation (BJD 2,453,002.671503).

^cLike RVLIN, BOOTTRAN assumes no jitter or fixes jitter to a certain value, while MCMC treats it as a free parameter. See § 6.3.3.

Table 6.4. PHOTOMETRIC
OBSERVATIONS OF HD 37605
FROM THE T12 0.8m APT

Heliocentric Julian Date (HJD – 2,400,000)	$\Delta(b + y)/2$ (mag)
54,481.7133	1.4454
54,482.6693	1.4474
54,482.7561	1.4442
54,483.6638	1.4452
54,495.7764	1.4469
54,498.7472	1.4470

Note. — This table is presented in its entirety in the electronic edition of the Astrophysical Journal. A portion is shown here for guidance regarding its form and content.

Table 6.5. PHOTOMETRIC
OBSERVATIONS OF HD 37605 ON
MOST

Heliocentric Julian Date (HJD – 2,451,545)	Relative Magnitude (mag)
4355.5105	-0.0032
4355.5112	-0.0047
4355.5119	-0.0018
4355.5126	-0.0026
4355.5133	-0.0018
4355.5140	-0.0039

Note. — This table is presented in its entirety in the electronic edition of the Astrophysical Journal. A portion is shown here for guidance regarding its form and content.

Table 6.6. Updated $M \sin i$ and Errors for HD 114762 b and HD 168443 b, c

Planet	$M \sin i \pm \text{std. error } (M_{\text{Jup}})$
HD 114762 b^a	11.086 ± 0.067
HD 114762 b^b	11.069 ± 0.063
HD 168443 b	7.696 ± 0.015
HD 168443 c	17.378 ± 0.044

^aFor best orbital fit with RV trend (dv/dt).

^bFor best orbital fit without RV trend (dv/dt).

Table 6.7. HET and KECK RADIAL VELOCITIES FOR HD 37605

BJD−2440000	Velocity (m/s)	Uncertainty (m/s)	Telescope
13002.671503	122.4	8.8	HET
13003.685247	126.9	5.6	HET
13006.662040	132.1	5.2	HET
13982.116400	-312.6	1.1	Keck
13983.110185	-291.6	1.1	Keck
13984.104595	-175.0	1.1	Keck

Note. — This table is presented in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content.

Chapter 7

Summary and Future Directions

In this chapter, I summarize the main findings of this thesis briefly and outline some future directions that are not direct extensions on the works documented in previous chapters.

My work focuses on improving the RV precision of HET/HRS and Keck/HIRES, with the goal to improve the RV precision of HET/HRS to beyond 3 m/s and to the level of Keck/HIRES, around 1-2 m/s. The code I adopted for RV extraction is the widely used CPS Doppler code written by Marcy, Butler, Johnson, with contributions from Valenti and many others (Chapter 2). I examined the causes behind HET/HRS's lower RV precision in comparison to Keck/HIRES and came to the conclusion that the gas in the HRS iodine cell was not at its set-temperature of 70°C. However, the iodine atlas that was used for modeling HET/HRS data for years was taken at an iodine gas temperature of around 70°C. This discrepancy lead to the poor goodness of fits for HET/HRS spectra and higher RV RMS. The other cause is inaccurate IP modeling, where the current IP functions often have a hard time to converge onto a good solution due to the inaccurate iodine model, a lack of good initial guesses, and/or intrinsic complexity with the IP function itself. I believe that a more robust temperature control system and a (series of) validated FTS scan(s) for the HET/HRS iodine cell, plus a good IP function, will improve the RV precision of HET/HRS, including its archival data (Chapter 3).

My primary goal for improving Keck/HIRES precision is to pin down the causes for the RV systematic errors that manifest as correlations or trends between RV and BC. I conclude that the leading cause is the errors in Keck/HIRES's deconvolved stellar spectral template (DSST), which adds about 1 m/s to the error budget and is directly driving the RV-BC correlation. There are also secondary causes being algorithmic errors and telluric contamination, entering the error budget at 0.2-0.5 m/s level. I believe that a better algorithm for generating DSST, proper treatment for telluric contamination, and adoption for a more robust algorithm will bring the RV precision of Keck/HIRES to the next level (Chapter 4).

I have also done some works in characterizing planetary orbits using RV data. Most notably, I have published the BOOTTRAN package, which computes error bars for Keplerian fits via bootstrapping (Chapter 5). We have discovered a new planet, HD 37605c, using HET/HRS data analyzed by the adopted CPS Doppler code and also Keck/HIRES RVs (Chapter 6).

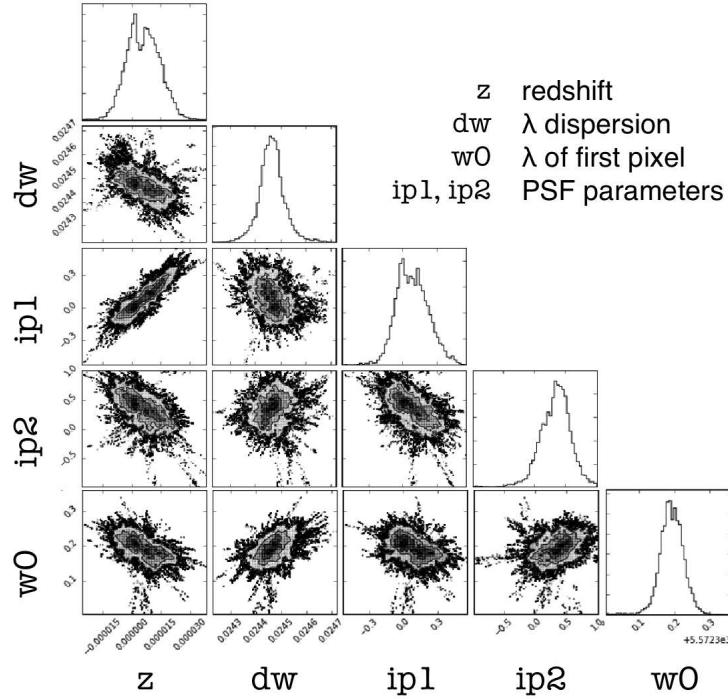


Figure 7.1 Initial result from the new Doppler pipeline: 1-D and 2-D marginalized posterior distribution for 5 parameters (out of 17, including two for Gaussian processes) in a fit for a 2Å spectral chunk.

In the concluding sections of Chapter 3 and 4, we have summarized future works which are natural next steps to improve the RV precision of HET/HRS and Keck/HIRES. Beyond these next steps, there are also other independent paths I plan to take in continuing my journey in exoplanet discoveries using precise Doppler spectroscopy.

First of all, I am constructing a new RV code in *Python*. The CPS Doppler code is a great legacy code which works at 1-2 m/s level, but it has many drawbacks: It is based on a simple home-constructed Levenberg-Marquardt least χ^2 fitter (LM fitter) which has high requirement on initial guesses for parameters and is terribly inefficient and inadequate in exploring the χ^2 space and finding the true minimum. It also has many legacy house keeping parts and complicated structures that makes it hard to upgrade, adopt for other instruments, and add new modules and functions. The new code carries on the valuable successful parts of the CPS code over, and more importantly, built to be highly modular and thus easy to adopt for other instruments or to plug in modern numerical and statistical tools. It employs advanced statistical and numerical tools such as a MCMC algorithm for Bayesian statistics and Gaussian processes, which will model the correlated noise originated from a complex blaze function, normalization issues, CCD effects, telluric line modeling residuals, imperfect deconvolved stellar spectrum, and so on (see, e.g., Czekala et al. 2014 and Daniel Foreman-Mackey’s *george* package). The preliminary result from this code (fitting for a spectral chunk) is shown in Figure 7.1.

I also plan to implement features for estimating RVs induced by stellar activity in

the new package, such as bisector fitting, analysis on line depth change, and color-dependent RV estimate. This means disentangling the stellar activity and planetary RV signals directly from the spectral data, instead of estimating and subtracting the effects of stellar activity “after the fact”. I also plan to work with astrophysicists specialized in line formation and atmospheric activities to work out the best activity indicators or modeling strategies and build a more astrophysically informed Doppler code.

On a different note, what should be done once the RVs are extracted from the spectra? I described the “black magic” of vanking in Chapter 2, where the readers can see that this intuitively constructed statistical weighting procedure is composed of seemingly arbitrary steps and formulas of outlier rejection and weight evaluation. The error bars are also computed in the vanking process, depending mostly on the weights, but they often appear to be underestimated (not only due to the existence of stellar RV jitter). To improve vanking, I am collaborating with Ben Nelson to formulate a more rigorous and statistically justified “vanking” method. The concept of vanking also needs to be reinvented once we come out of the least- χ^2 frame and move on to the land of posteriors.

In terms of applications to real data, there will be numerous next-generation RV instruments which would benefit from these modern data analysis tools. I will work on the data taken by the MINIature Exoplanet RV Array (MINERVA), which is an array of four 0.7-m telescopes feeding into a temperature and pressure stabilized spectrometer, dedicated for RV surveys (under commissioning as of May 2016). MINERVA is iodine-calibrated, with the possibility to go iodine free. I will take advantage of the unique high-cadence capability of MINERVA to tackle the problem of stellar activity induced RV signal.

I also plan to continue working on HET/HRS data, especially for the upgraded HET/HRS, which will have an image slicer so that all the light in the fiber can be fed through the slit to boost the throughput by a factor of 2-4 (in combination with the upgrade of the telescope). The new HET/HRS will have a new spectral format, which is similar to that of MINERVA and the upgraded Magellan/PFS (Crane et al. 2010). The new format contains several traces for each spectral order, each representing a slice of the image (for HRS and PFS; in the case of MINERVA, each trace comes from one of the four telescopes). It would be interesting to explore how the RVs extracted from each trace should be combined or whether it is better to “smash” the traces together when reducing the image to extract the 1-D spectrum.

MINERVA, the new HET/HRS, and the new Magellan/PFS are all iodine-calibrated instruments. I am also hoping to work on laser-comb calibrated instruments such as WIYN-NEID, where, for example, treatment of telluric contamination becomes an interesting problem, because the conventional way of extracting RVs is to perform cross correlation, where forward modeling of telluric lines does not seem to fit into the picture (yet; and similarly for HET/HPF Mahadevan et al. 2012).

The promise of high RV precision from these next-generation instruments provides exciting opportunities into new discovery spaces in exoplanets. I am excited and proud to carry on the legacy of HET/HRS and Keck/HIRES and be a part of the era when we will breach the 10 cm/s barrier and detect rocky planets in the Habitable Zone, likely the first Earth 2.0.

Bibliography

- Alonso, R., Brown, T. M., Torres, G., et al. 2004, ApJ, 613, L153
- Anglada-Escudé, G., López-Morales, M., & Chambers, J. E. 2010, ApJ, 709, 168
- Anglada-Escudé, G., & Tuomi, M. 2015, Science, 347, 1080
- Artigau, É., Astudillo-Defru, N., Delfosse, X., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9149, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 5
- Bailes, M., Lyne, A. G., & Shemar, S. L. 1991, Nature, 352, 311
- Bailey, J., Butler, R. P., Tinney, C. G., et al. 2009, ApJ, 690, 743
- Bailey, III, J. I., White, R. J., Blake, C. H., et al. 2012, ApJ, 749, 16
- Bakos, G. Á., Noyes, R. W., Kovács, G., et al. 2007, ApJ, 656, 552
- Baliunas, S. L., Donahue, R. A., Soon, W. H., et al. 1995, ApJ, 438, 269
- Baluev, R. V. 2009, MNRAS, 393, 969
- Baranne, A., Queloz, D., Mayor, M., et al. 1996, A&AS, 119, 373
- Barbieri, M., Alonso, R., Laughlin, G., et al. 2007, A&A, 476, L13
- Bean, J. L., Seifahrt, A., Hartman, H., et al. 2010, ApJ, 713, 410
- Bender, C. F., Mahadevan, S., Deshpande, R., et al. 2012, ApJ, 751, L31
- Benedict, G. F., McArthur, B. E., Bean, J. L., et al. 2010, AJ, 139, 1844
- Bertaux, J. L., Lallement, R., Ferron, S., Boonne, C., & Bodichon, R. 2014, A&A, 564, A46
- Beuzit, J.-L., Feldt, M., Dohlen, K., et al. 2008, in Proc. SPIE, Vol. 7014, Ground-based and Airborne Instrumentation for Astronomy II, 701418
- Blake, C. H., Charbonneau, D., & White, R. J. 2010, ApJ, 723, 684
- Bodenheimer, P., Laughlin, G., & Lin, D. N. C. 2003, ApJ, 592, 555

- Boisse, I., Pepe, F., Perrier, C., et al. 2012, ArXiv e-prints, arXiv:1205.5835
- Bond, I. A., Udalski, A., Jaroszyński, M., et al. 2004, ApJ, 606, L155
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977
- Bouchy, F., Udry, S., Mayor, M., et al. 2005, A&A, 444, L15
- Butler, R. P., & Marcy, G. W. 1996, ApJ, 464, L153
- Butler, R. P., Marcy, G. W., Fischer, D. A., et al. 1999, ApJ, 526, 916
- Butler, R. P., Marcy, G. W., Williams, E., et al. 1996a, PASP, 108, 500
- . 1996b, PASP, 108, 500
- Campbell, B., Walker, G. A. H., & Yang, S. 1988, ApJ, 331, 902
- Campbell, B., Yang, S., Irwin, A. W., & Walker, G. A. H. 1991, in Lecture Notes in Physics, Berlin Springer Verlag, Vol. 390, Bioastronomy: The Search for Extraterrestrial Life — The Exploration Broadens, ed. J. Heidmann & M. J. Klein, 19–20
- Chambers, J. E. 1999, MNRAS, 304, 793
- Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, ApJ, 529, L45
- Christian, D. J., Pollacco, D. L., Skillen, I., et al. 2006, MNRAS, 372, 1117
- Clough, S. A., Iacono, M. J., & Moncet, J.-L. 1992, J. Geophys. Res., 97, 15
- Cochran, W. D., Endl, M., Wittenmyer, R. A., & Bean, J. L. 2007, ApJ, 665, 1407
- Cochran, W. D., Endl, M., McArthur, B., et al. 2004, ApJ, 611, L133
- Cosentino, R., Lovis, C., Pepe, F., et al. 2012, in Proc. SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84461V
- Crane, J. D., Shectman, S. A., Butler, R. P., et al. 2010, in Proc. SPIE, Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III, 773553
- Cunha, D., Santos, N. C., Figueira, P., et al. 2014, A&A, 568, A35
- Czekala, I., Andrews, S. M., Mandel, K. S., Hogg, D. W., & Green, G. M. 2014, ArXiv e-prints, arXiv:1412.5177
- Davison, A., & Hinkley, D. 1997, Bootstrap Methods and Their Application, Cambridge Series on Statistical and Probabilistic Mathematics (Cambridge University Press)
- de Wit, J., & Seager, S. 2013, Science, 342, 1473
- Demarque, P., Woo, J.-H., Kim, Y.-C., & Yi, S. K. 2004, ApJS, 155, 667
- Deming, D., Espenak, F., Jennings, D. E., Brault, J. W., & Wagner, J. 1987, ApJ, 316, 771

- Demory, B.-O., Gillon, M., Deming, D., et al. 2011, A&A, 533, A114
- Dragomir, D., Kane, S. R., Pilyavsky, G., et al. 2011, AJ, 142, 115
- Dragomir, D., Kane, S. R., Henry, G. W., et al. 2012, ApJ, 754, 37
- Dravins, D. 1985, in Stellar Radial Velocities, ed. A. G. D. Philip & D. W. Latham, 311–320
- Dumusque, X., Boisse, I., & Santos, N. C. 2014, ApJ, 796, 132
- Dumusque, X., Lovis, C., Ségransan, D., et al. 2011, A&A, 535, A55
- Dumusque, X., Pepe, F., Lovis, C., et al. 2012, Nature, 491, 207
- Endl, M., Caldwell, D. A., Barclay, T., et al. 2014, ApJ, 795, 151
- ESA. 1997, VizieR Online Data Catalog, 1239, 0
- Feng, Y. K., Wright, J. T., Nelson, B., et al. 2015, ApJ, 800, 22
- Figueira, P., Pepe, F., Melo, C. H. F., et al. 2010, A&A, 511, A55
- Fischer, D. A., Marcy, G. W., & Spronck, J. F. P. 2014, ApJS, 210, 5
- Fischer, D. A., Vogt, S. S., Marcy, G. W., et al. 2007, ApJ, 669, 1336
- Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, PASP, 128, 066001
- Ford, E. B. 2005, AJ, 129, 1706
- . 2006, ApJ, 642, 505
- Ford, E. B., & Gregory, P. C. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 371, Statistical Challenges in Modern Astronomy IV, ed. G. J. Babu & E. D. Feigelson, 189
- Fossey, S. J., Waldmann, I. P., & Kipping, D. M. 2009, MNRAS, 396, L16
- Freedman, D. A. 1981, The Annals of Statistics, 9, pp. 1218
- Fressin, F., Torres, G., Charbonneau, D., et al. 2013, ApJ, 766, 81
- Gladman, B. 1993, Icarus, 106, 247
- Griffin, R., & Griffin, R. 1973, MNRAS, 162, 255
- Grundahl, F., Christensen-Dalsgaard, J., Gråe Jørgensen, U., et al. 2011, Journal of Physics Conference Series, 271, 012083
- Halverson, S., Roy, A., Mahadevan, S., et al. 2015, ApJ, 806, 61
- Hatzes, A. P. 2013, Astronomische Nachrichten, 334, 616

- Hatzes, A. P., & Cochran, W. D. 1993, ApJ, 413, 339
- Hatzes, A. P., Cochran, W. D., Endl, M., et al. 2003, ApJ, 599, 1383
- . 2006, A&A, 457, 335
- Haywood, R. D., Collier Cameron, A., Queloz, D., et al. 2014, MNRAS, 443, 2517
- Henry, G. W. 1999, PASP, 111, 845
- Henry, G. W., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2000, ApJ, 529, L41
- Henry, G. W., Kane, S. R., Wang, S. X., et al. 2013, ApJ, 768, 155
- Hoaglin, D. C., Mosteller, F., & Tukey, J. W. 1983, Understanding Robust and Exploratory Data Analysis, Wiley Classics Library Editions (Wiley)
- Horne, K. 1986, PASP, 98, 609
- Howard, A. W., Johnson, J. A., Marcy, G. W., et al. 2009, ApJ, 696, 75
- . 2010, ApJ, 721, 1467
- . 2011, ApJ, 726, 73
- Howard, A. W., Sanchis-Ojeda, R., Marcy, G. W., et al. 2013, Nature, 503, 381
- Ibukiyama, A., & Arimoto, N. 2002, A&A, 394, 927
- Isaacson, H., & Fischer, D. 2010, ApJ, 725, 875
- Jacob, W. S. 1855, MNRAS, 15, 228
- Johnson, J. A., Marcy, G. W., Fischer, D. A., et al. 2006, ApJ, 647, 600
- Johnson, J. A., Clanton, C., Howard, A. W., et al. 2011a, ApJS, 197, 26
- . 2011b, ApJS, 197, 26
- Johnson, M. C., Endl, M., Cochran, W. D., et al. 2016, ArXiv e-prints, arXiv:1602.05200
- Jones, H. R. A., Butler, R. P., Tinney, C. G., et al. 2010, MNRAS, 403, 1703
- Jontof-Hutter, D., Ford, E. B., Rowe, J. F., et al. 2016, ApJ, 820, 39
- Juhola, A. 1975, Carbon, 13, 437
- Kalas, P., Graham, J. R., Chiang, E., et al. 2008, Science, 322, 1345
- Kane, S. R., Ephemeris Refinement, T., & Survey(TERMS), M. 2012, in American Astronomical Society Meeting Abstracts, Vol. 219, American Astronomical Society Meeting Abstracts, #228.07
- Kane, S. R., Henry, G. W., Dragomir, D., et al. 2011a, ApJ, 735, L41

- . 2011b, ApJ, 735, L41
- Kane, S. R., Mahadevan, S., von Braun, K., Laughlin, G., & Ciardi, D. R. 2009, PASP, 121, 1386
- Kane, S. R., Ciardi, D., Fischer, D., et al. 2011c, Detection and Dynamics of Transiting Exoplanets, St. Michel l'Observatoire, France, Edited by F. Bouchy; R. Díaz; C. Moutou; EPJ Web of Conferences, Volume 11, id.06005, 11, 6005
- Kane, S. R., Ciardi, D., Dragomir, D., et al. 2011d, in American Astronomical Society Meeting Abstracts, #128.03
- Kane, S. R., Dragomir, D., Ciardi, D. R., et al. 2011e, ApJ, 737, 58
- . 2011f, ApJ, 737, 58
- Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, Icarus, 101, 108
- Knöckel, H., Bodermann, B., & Tiemann, E. 2004, European Physical Journal D, 28, 199
- Knutson, H. A., Fulton, B. J., Montet, B. T., et al. 2014, ApJ, 785, 126
- Konacki, M., Torres, G., Jha, S., & Sasselov, D. D. 2003, Nature, 421, 507
- Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131
- Koren, S. C., Blake, C. H., Dahn, C. C., & Harris, H. C. 2016, AJ, 151, 57
- Kürster, M., Trifonov, T., Reffert, S., Kostogryz, N. M., & Rodler, F. 2015, A&A, 577, A103
- Latham, D. W., Stefanik, R. P., Mazeh, T., Mayor, M., & Burki, G. 1989, Nature, 339, 38
- Laughlin, G., Deming, D., Langton, J., et al. 2009, Nature, 457, 562
- Li, C.-H., Benedick, A. J., Fendel, P., et al. 2008, Nature, 452, 610
- Lovis, C., Ségransan, D., Mayor, M., et al. 2011, A&A, 528, A112
- Ma, B., Ge, J., & Varosi, F. 2016, in American Astronomical Society Meeting Abstracts, Vol. 227, American Astronomical Society Meeting Abstracts, 137.18
- Macintosh, B., Graham, J. R., Ingraham, P., et al. 2014, Proceedings of the National Academy of Science, 111, 12661
- Mahadevan, S., Ramsey, L., Bender, C., et al. 2012, in Proc. SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84461S
- Mandel, K., & Agol, E. 2002, ApJ, 580, L171
- Marcy, G. W., & Butler, R. P. 1992, PASP, 104, 270

- Marcy, G. W., Butler, R. P., Fischer, D. A., et al. 2002, *ApJ*, 581, 1375
- Marcy, G. W., Isaacson, H., Howard, A. W., et al. 2014, *ApJS*, 210, 20
- Marois, C., Macintosh, B., Barman, T., et al. 2008, *Science*, 322, 1348
- Matthews, J. M., Kuschnig, R., Guenther, D. B., et al. 2004, *Nature*, 430, 51
- Mayor, M., & Queloz, D. 1995, *Nature*, 378, 355
- Mayor, M., Pepe, F., Queloz, D., et al. 2003, *The Messenger*, 114, 20
- Mayor, M., Bonfils, X., Forveille, T., et al. 2009, *A&A*, 507, 487
- Mazeh, T., Nachmani, G., Holczer, T., et al. 2013, *ApJS*, 208, 16
- McArthur, B. E., Endl, M., Cochran, W. D., et al. 2004, *ApJ*, 614, L81
- McCullough, P. R., Stys, J. E., Valenti, J. A., et al. 2006, *ApJ*, 648, 1228
- Meschiari, S., Wolf, A. S., Rivera, E., et al. 2009, *PASP*, 121, 1016
- Michelson, A. A., & Morley, E. 1887, *American Journal of Science*, 34, 333
- Morton, T. D., Bryson, S. T., Coughlin, J. L., et al. 2016, *ApJ*, 822, 86
- Motalebi, F., Udry, S., Gillon, M., et al. 2015, *A&A*, 584, A72
- Moutou, C., Mayor, M., Lo Curto, G., et al. 2011, *A&A*, 527, A63
- Nelson, B., Ford, E. B., & Payne, M. J. 2014a, *ApJS*, 210, 11
- Nelson, B. E., Ford, E. B., Wright, J. T., et al. 2014b, *MNRAS*, 441, 442
- Niedzielski, A., Deka-Szymankiewicz, B., Adamczyk, M., et al. 2016, *A&A*, 585, A73
- Payne, M. J., & Ford, E. B. 2011, *ApJ*, 729, 98
- Pepe, F., Mayor, M., Galland, F., et al. 2002a, *A&A*, 388, 632
- Pepe, F., Mayor, M., Rupprecht, G., et al. 2002b, *The Messenger*, 110, 9
- Pepe, F., Correia, A. C. M., Mayor, M., et al. 2007, *A&A*, 462, 769
- Pepe, F., Cameron, A. C., Latham, D. W., et al. 2013, *Nature*, 503, 377
- Perruchot, S., Kohler, D., Bouchy, F., et al. 2008, in Proc. SPIE, Vol. 7014, Ground-based and Airborne Instrumentation for Astronomy II, 70140J
- Pilyavsky, G., Mahadevan, S., Kane, S. R., et al. 2011a, *ApJ*, 743, 162
- . 2011b, *ApJ*, 743, 162
- Piskunov, N. E., & Valenti, J. A. 2002, *A&A*, 385, 1095

- Plavchan, P., Latham, D., Gaudi, S., et al. 2015, ArXiv e-prints, arXiv:1503.01770
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2002, Numerical recipes in C++ : the art of scientific computing
- Quirrenbach, A., Amado, P. J., Caballero, J. A., et al. 2014, in Proc. SPIE, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 91471F
- Rajpaul, V., Aigrain, S., & Roberts, S. 2016, MNRAS, 456, L6
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, in Proc. SPIE, Vol. 9143, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave, 914320
- Robertson, P., Mahadevan, S., Endl, M., & Roy, A. 2014, Science, 345, 440
- . 2015, Science, 347, 1080
- Rogers, L. A. 2015, ApJ, 801, 41
- Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, J. Quant. Spec. Radiat. Transf., 130, 4
- Rowe, J. F., Matthews, J. M., Seager, S., et al. 2006, ApJ, 646, 1241
- Rowe, J. F., Bryson, S. T., Marcy, G. W., et al. 2014, ApJ, 784, 45
- Santos, N. C., Israelian, G., Mayor, M., et al. 2005, A&A, 437, 1127
- Schneider, J., Dedieu, C., Le Sidaner, P., Savalle, R., & Zolotukhin, I. 2011, A&A, 532, A79
- Shetrone, M., Cornell, M. E., Fowler, J. R., et al. 2007, PASP, 119, 556
- Sithajan, S., Ge, J., & Wang, J. 2016, in American Astronomical Society Meeting Abstracts, Vol. 227, American Astronomical Society Meeting Abstracts, 137.19
- Struve, O. 1952, The Observatory, 72, 199
- Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., et al. 2015, ApJ, 809, 77
- Swift, J. J., Bottom, M., Johnson, J. A., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 027002
- Topley, B. 1926, Nature, 118, 802
- Torres, G., Andersen, J., & Giménez, A. 2010, A&A Rev., 18, 67
- Tull, R. G. 1998, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 3355, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. S. D'Odorico, 387–398
- Valenti, J. A., Butler, R. P., & Marcy, G. W. 1995, 107, 966

- Valenti, J. A., & Fischer, D. A. 2005, ApJS, 159, 141
- Valenti, J. A., & Piskunov, N. 1996, A&AS, 118, 595
- Valenti, J. A., Fischer, D., Marcy, G. W., et al. 2009, ApJ, 702, 989
- van de Kamp, P. 1969, AJ, 74, 757
- van Leeuwen, F. 2008, VizieR Online Data Catalog, 1311, 0
- Vanderburg, A., Johnson, J. A., & Muirhead, P. 2013, in American Astronomical Society Meeting Abstracts, Vol. 221, American Astronomical Society Meeting Abstracts #221, 149.08
- Vanderburg, A., Plavchan, P., Asher Johnson, J., et al. 2016a, MNRAS, arXiv:1604.03143
- Vanderburg, A., Latham, D. W., Buchhave, L. A., et al. 2016b, ApJS, 222, 14
- Vargaftik, N. 1993, Handbook of Thermal Conductivity of Liquids and Gases (Taylor & Francis)
- Vogt, S. S., Butler, R. P., Rivera, E. J., et al. 2010, ApJ, 723, 954
- Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 2198, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. D. L. Crawford & E. R. Craine, 362
- Vogt, S. S., Radovan, M., Kibrick, R., et al. 2014, PASP, 126, 359
- Vogt, S. S., Burt, J., Meschiari, S., et al. 2015, ApJ, 814, 12
- Walker, G., Matthews, J., Kuschnig, R., et al. 2003, PASP, 115, 1023
- Wang, Sharon, X., Wright, J. T., Cochran, W., et al. 2012, ApJ, 761, 46
- Waser, J., & Wieland, K. 1947, Nature, 160, 643
- Weiss, L. M., & Marcy, G. W. 2014, ApJ, 783, L6
- Wittenmyer, R. A., Endl, M., Cochran, W. D., & Levison, H. F. 2007, AJ, 134, 1276
- Wittenmyer, R. A., Endl, M., Cochran, W. D., Levison, H. F., & Henry, G. W. 2009, ApJS, 182, 97
- Wittenmyer, R. A., Horner, J., Tuomi, M., et al. 2012, ArXiv e-prints, arXiv:1205.2765
- Wittenmyer, R. A., Wang, S., Horner, J., et al. 2013, ApJS, 208, 2
- Wolfgang, A., & Lopez, E. 2015, ApJ, 806, 183
- Wolfgang, A., Rogers, L. A., & Ford, E. B. 2015, ArXiv e-prints, arXiv:1504.07557

- Wolszczan, A., & Frail, D. A. 1992, *Nature*, 355, 145
- Wright, J. T. 2006, PhD thesis, University of California, Berkeley, California, USA
- Wright, J. T., & Eastman, J. D. 2014, *PASP*, 126, 838
- Wright, J. T., & Gaudi, B. S. 2013, *Exoplanet Detection Methods*, ed. T. D. Oswalt, L. M. French, & P. Kalas, 489
- Wright, J. T., & Howard, A. W. 2009, *ApJS*, 182, 205
- Wright, J. T., Marcy, G. W., Butler, R. P., et al. 2008, *ApJ*, 683, L63
- Wright, J. T., Upadhyay, S., Marcy, G. W., et al. 2009a, *ApJ*, 693, 1084
- Wright, J. T., Fischer, D. A., Ford, E. B., et al. 2009b, *ApJ*, 699, L97
- Wright, J. T., Fakhouri, O., Marcy, G. W., et al. 2011, *PASP*, 123, 412
- Wu, Y., & Lithwick, Y. 2013, *ApJ*, 772, 74
- Yee, J. C., Albrow, M., Barry, R. K., et al. 2014, ArXiv e-prints, arXiv:1409.2759
- Yi, X., Vahala, K., Li, J., et al. 2016, *Nature Communications*, 7, 10436

Vita

Sharon Xuesong Wang

525 Davey Lab
University Park, PA 16802
(814) 321-7236
email: wang.xuesong.sharon@gmail.com
webpage: bit.ly/sharonxuesongwang

EDUCATION

<i>PhD</i> , Astronomy & Astrophysics, Penn State University	Aug 2016
<i>PhD Minor</i> , Computational Sciences	
Thesis Advisor: Dr. Jason T. Wright	
<i>Bachelor of Science</i> , Physics, Tsinghua University, Beijing, China	Jun 2008
Thesis: <i>Characterizing the Luminosity-Variability Correlation in Gamma-Ray Bursts</i>	
Advisor: Dr. Shuang-Nan Zhang	

PROFESSIONAL EMPLOYMENT

<i>Research Assistant / Fellow</i>	Jan 2009 – present
Department of Astronomy & Astrophysics, Penn State University	
<i>Teaching Assistant</i>	Aug 2008 – Dec 2009
Department of Astronomy & Astrophysics, Penn State University	

AWARDS

<i>Carnegie DTM Fellowship in Astronomy and Planetary Science</i>	since Aug 2016
<i>NASA Earth and Space Science Fellow</i>	since Sep 2014
Proposal title: Finding the Lowest Mass Exoplanets with Improved Radial Velocimetry	
<i>Downsborough Graduate Fellowship</i>	May 2013
<i>Stephen B. Brumbach Fellowship in Astrophysics</i>	May 2010
<i>Zaccheus Daniel Fellowship</i>	2009-2013
<i>Teaching Assistant of the Year Award</i>	Jun 2009
Department of Astronomy & Astrophysics, Penn State University	