

DEMONEX: The DEDicated MONitor of EXotransits

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of
Philosophy in the Graduate School of The Ohio State University

By

Jason D. Eastman

Graduate Program in Astronomy

The Ohio State University

2011

Dissertation Committee:

Professor B. Scott Gaudi, Advisor

Professor Richard Pogge

Professor Krzysztof Stanek

Copyright by
Jason D. Eastman
2011

ABSTRACT

Transiting planets are extraordinarily useful laboratories to study the formation and evolution of planets as a whole. In order to capitalize on this potential, we built DEMONEX, the DEDicated MONitor of EXotransits, a low-cost, 0.5 meter, robotic telescope, mostly from commercially-available parts. The primary goal of DEMONEX is to monitor bright stars hosting transiting planets in order to provide a homogeneous data set for all known transiting systems visible from its location at Winer Observatory in Sonoita, Arizona. The main advantage of DEMONEX is that it can be used every night for transit follow-up. With over 60 known transiting planets visible from Winer Observatory, more than 90% of its time can be spent observing transits. In chapter 1, we describe the hardware, scheduling, observing, data reduction software, and we present some results from the first three years of operation.

One of the most powerful tools that observing multiple transits gives us is the ability to observe Transit Timing Variations (TTVs) – slight variations in the otherwise perfectly periodic times of transits due to, among other things, other planets in the system. However, the times that many people report are ambiguous

at a level than threatens the utility of the technique. In chapter 2, we detail the steps necessary to accurately convert the time as measured on Earth, the JD_{UTC} , to an astrophysically-meaningful time, the BJD_{TDB} , with a precision that far surpasses the practical limit for TTVs.

In order to analyze the vast amounts of data that DEMONEX has collected, in chapter 3 we discuss our analysis software, EXOFAST. It is a fast, robust suite of routines written in IDL designed to fit exoplanetary transits and radial velocity simultaneously and characterize the parameter uncertainties of and covariances of exoplanetary systems with a Markov Chain Monte Carlo method. We take particular care to make it robust, fast, and self-consistent. We outline ways in which the code can easily be extended to include additional effects or generalized for the characterization of any data set – including non-planetary data sets. Of particular interest, we discuss the effects of eccentricity bias and demonstrate the effectiveness of our new strategy to mitigate it, derive a method to analytically fit the linear and quadratic limb darkening coefficients of a planetary transit, and explain how we achieved improved robustness and over a factor of 100 improvement in the run time of the transit model calculation. Our source code, along with easy-to-use online tools for both the timing code and EXOFAST can be found online.

Dedicated to the loving memory of my dad, whose contagious love of astronomy was the seed of this dissertation.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my wife and namesake of NGC 3628, Sarah LoPresti, who has made many difficult sacrifices to make this possible. I am unfathomably lucky to have such a wonderful and understanding partner.

I wish to express my deep gratitude for the love and encouragement from my family throughout my life that has led me to this point. Knowing I have the unconditional support from so many people is humbling – I am confident I never would have made it here without it.

I will remember the many friends I have made at OSU for a lifetime. From late nights at the Library and diagonalizing Fisher matrices to traveling across the country for my wedding, it is good to know that I will always have friends like them and I can only hope that I have given back as much as I have taken. Robert Siverd has been a tremendous benefit to me over the four years we shared an office with his unique insights, vast scripting and computer knowledge, and insatiable curiosity.

I would like to thank everyone in the ISL for giving me the opportunity to contribute to fascinating projects like aluminizing, the 4K, and MODS, and

particularly Thomas O'Brien for his help in diagnosing and fixing DEMONEX's numerous deficiencies.

Mark and Pat Trueblood graciously hosted me during the initial month-long setup of DEMONEX and continue to go above and beyond to ensure that it runs smoothly. David Will kept the computers running, despite network outages, operator errors, disk failures, and even a hacker.

I am thankful for all the faculty at OSU for cultivating such a rich research environment, particularly Richard Pogge, whose expertise on topics from instrumentation to wine and just about everything in between has been an enormous benefit to me; Darren DePoy, who helped get DEMONEX off the ground; Paul Martini, who had the patience to work with me my first year as figured out where my scientific interests lie; Andy Gould, who always keeps Coffee interesting; and Krzysztof Stanek, who I am I proud to call my fellow American.

Finally, I would like extend a special thanks to my advisor, B. Scott Gaudi, whose unwavering dedication and encyclopedic knowledge on all topics astronomy is enviable. He invested a large fraction of his startup money on this thesis, had patience with me as I ignored his advice (only to learn it the hard way), and taught me more than I could ever articulate.

VITA

February 18, 1983	Born – Omaha, Nebraska
2005	B.A. Astronomy and Physics, Boston University
2005 – 2006	Price Instrumentation Fellow, The Ohio State University
2006 – 2007	Research Assistant, The Ohio State University
2007	M.S. Astronomy, The Ohio State University
2007 – 2008	Price Instrumentation Fellow, The Ohio State University
2008 – 2010	Teaching Assistant and Research Assistant, The Ohio State University
2010 – 2011	Presidential Fellow, The Ohio State University

PUBLICATIONS

Research Publications

1. Carraro, G., Janes, K. A., & **Eastman, J. D.** “Photometry of neglected open clusters in the first and fourth Galactic quadrants”, MNRAS, 364, 179, (2005).
2. Yost, S.A., et al. (incl. **Eastman, J.D.**), “Exploring Broadband GRB

- Behavior during γ -Ray Emission”, *ApJ*, 657, 925, (2007).
3. Phillips, M. M., et al. (incl. **Eastman, J.D.**), “The Peculiar SN 2005hk: Do Some Type Ia Supernovae Explode as Deflagration?”, *PASP*, 119, 360, (2007).
 4. **Eastman, J. D.**, Martini, P., Sivakoff, G., Kelson, D. D., Mulchaey, J. S., & Tran, K., “First Measurement of a Rapid Increase in the AGN Fraction in High-Redshift Clusters of Galaxies”, *ApJL*, 664, 9, (2007).
 5. Frieman, J. A., et al. (incl. **Eastman, J.D.**), “The Sloan Digital Sky Survey-II Supernova Survey: Technical Summary”, *AJ*, 135, 338, (2008).
 6. Zheng, C., et al. (incl. **Eastman, J.D.**), “First-Year Spectroscopy for the Sloan Digital Sky Survey-II Supernova Survey”, *AJ*, 135, 1766, (2008).
 7. Grier, C. J., et al. (incl. **Eastman, J. D.**), “The Mass of the Black Hole in the Quasar PG 2130+099”, *ApJ*, 688, 837, (2008).
 8. Carter, J. A., Yee, J. C., **Eastman, J. D.**, Gaudi, B. S., Winn, J. N., “Analytic Approximations for Transit Light-Curve Observables, Uncertainties, and Covariances”, *ApJ*, 689, 499, (2008).
 9. Valtonen, M. J., et al. (incl. **Eastman, J.D.**), “Tidally Induced Outbursts in OJ 287 during 2005-2008”, *ApJ*, 698, 781, (2009).
 10. Yee, J. C., et al. (incl. **Eastman, J.D.**), “Extreme Magnification Microlensing Event OGLE-2008-BLG-279: Strong Limits on Planetary Companions to the Lens Star”, *ApJ*, 703, 2082, (2009).
 11. Denney, K. D., et al. (incl. **Eastman, J.D.**), “A Revised Broad-line Region Radius and Black Hole Mass for the Narrow-line Seyfert 1 NGC 4051”, *ApJ*, 702, 1353, (2009).
 12. Denney, K. D., et al. (incl. **Eastman, J.D.**), “Diverse Kinematic Signatures from Reverberation Mapping of the Broad-Line Region in AGNs”, *ApJL*, 704, L80, (2009).
 13. Villforth, C., et al. (incl. **Eastman, J.D.**), “Variability and stability in blazar jets on time-scales of years: optical polarization monitoring of OJ 287 in 2005-2009”, *MNRAS*, 402, 2087, (2010).
 14. Fleming, S. W., et al. (incl. **Eastman, J. D.**), “Discovery of a Low-

Mass Companion to a Metal-Rich F Star with the MARVELS Pilot Project”, ApJ, 718, 1186, (2010).

15. **Eastman, J. D.**, Siverd, R. J., & Gaudi. B. S., “Achieving better than 1 minute accuracy in the Heliocentric and Barycentric Julian Dates”, PASP, 122, 935, (2010).

16. Gould, A., et al. (incl. **Eastman, J.D.**), “Frequency of Solar-like Systems and of Ice and Gas Giants Beyond the Snow Line from High-magnification Microlensing Events in 2005-2008”, ApJ, 720, 1073, (2010).

17. Denney, K. D., et al. (incl. **Eastman, J.D.**), “Reverberation Mapping Measurements of Black Hole Masses in Six Local Seyfert Galaxies”, ApJ, 721, 715, (2010).

18. Khan, R., et al. (incl. **Eastman, J. D.**), “Pre-discovery and Follow-up Observations of the Nearby SN 2009nr: Implications for Prompt Type Ia SNe”, ApJ, 726, 106, (2010).

19. Lee, B. L., et al. (incl. **Eastman, J.D.**), “MARVELS-1b: A Short-Period, Brown Dwarf Desert Candidate from the SDSS-III MARVELS Planet Search”, ApJ, 728, 32, (2010).

FIELDS OF STUDY

Major Field: Astronomy

TABLE OF CONTENTS

Abstract	ii
Dedication	iv
Acknowledgments	v
Vita	vii
List of Tables	xiv
List of Figures	xv
Chapter 1 Introduction	1
Chapter 2 DEMONEX: The DEDicated MONitor of EXotransits	5
2.1 Introduction	5
2.2 Science Goals	6
2.3 The System	9
2.4 Automation	12
2.5 Results	14
Chapter 3 Achieving Better Than 1 Minute Accuracy in the Heliocentric and Barycentric Julian Dates	20
3.1 Introduction	20
3.2 Theoretical Timing Precision	23
3.2.1 Reference Frames: The Rømer Delay	24

3.2.2	Time Standards: The clock correction	28
3.2.3	Shapiro Delay	35
3.2.4	Einstein Delay	35
3.3	Practical Considerations	36
3.3.1	Coordinates	37
3.3.2	Computer Clock	37
3.3.3	Flux-weighted mean time of exposure	38
3.3.4	Plane-Wave Approximation	39
3.3.5	Geocenter	41
3.3.6	BJD'_{UTC} versus BJD_{UTC}	41
3.3.7	Computer Precision	42
3.3.8	Negligible ($< 1\mu\text{s}$) effects	42
3.4	Calculating the BJD_{TDB}	43
3.4.1	Our Code	46
3.5	Conclusion	50
3.6	Current Implementations	52
3.6.1	Software	52
3.6.2	Space telescopes	55
Chapter 4	EXOFAST	66
4.1	Introduction	66
4.2	Overview of the Problem	68
4.2.1	Finding the best fit	68
4.2.2	Finding the uncertainties: MCMC	72
4.3	Radial Velocity Model	76

4.3.1	RV Parameterization	79
4.3.2	Eccentricity Bias	83
4.3.3	Radial Velocity Best Fit	90
4.4	Transit Model	92
4.4.1	Limb Darkening	93
4.4.2	Planetary Path	95
4.4.3	Inputs	98
4.4.4	Other biases	99
4.4.5	Transit Best Fit	100
4.5	Radial Velocity and Transit	101
4.5.1	Inputs	102
4.5.2	Stellar Mass and Radius	102
4.5.3	Planetary Mass	105
4.5.4	Roemer Delay	106
4.5.5	Combined Best Fit	108
4.6	Outputs	108
4.6.1	Plots	109
4.6.2	Parameters and Uncertainties	111
4.6.3	Models	112
4.7	Validation	112
4.7.1	Known Systems	112
4.7.2	New Systems	113
4.7.3	Simulated data	113
4.8	Additional Utilities	113

4.8.1	READEXO	113
4.8.2	Sys-Rem	114
4.9	Online Tools	115
4.9.1	Ephemerides	115
4.9.2	Transit fitting	116
4.9.3	RV fitting	116
4.9.4	Quadratic limb darkening	117
4.10	Summary and Suggestions for Future Extensions	117
Appendix A Analytic fit of the linear and quadratic limb darkening		134
Appendix B EXOFAST_OCCULTQUAD		137
B.0.1	Speed Enhancements	137
B.0.2	Bug fixes and conceptual enhancements	139
Appendix C Random Number Generator		142
Appendix D Performance		145
Bibliography		147

LIST OF TABLES

4.1	An example output of EXOFAST_LATEXTAB	132
4.1	An example output of EXOFAST_LATEXTAB	133

LIST OF FIGURES

2.1	Fraction of nights with transits	16
2.2	DEMONEX in Arizona	17
2.3	Three primary transits of TrES-3b	18
2.4	Allan plot of DEMONEX noise	19
3.1	Difference between BJD_{TDB} and the uncorrected JD_{TT}	60
3.2	Difference between the BJD_{TDB} and the HJD_{TDB}	61
3.3	Difference between the uniform BJD_{TDB} and the BJD_{UTC}	62
3.4	Difference between BJD_{TDB} with a small offset in Right Ascension	63
3.5	Error in using the geocentric BJD_{TDB}	64
3.6	Difference between BJD_{UTC} and BJD'_{UTC}	65
4.1	Error induced from zero-point in T_C	119
4.2	Error covariance induced from zero-point in T_C	120
4.3	Prior probability of e and ω	121
4.4	Prior probability of e and ω (zoomed)	122
4.5	Difference between the measured and actual eccentricity	123
4.6	Parameter distribution of e	124
4.7	Execution time of the MCMC fit	125
4.8	Prior probability of stellar mass	126
4.9	Prior probability of stellar radii	127

4.10 RV and transit parameter distributions	128
4.11 RV and transit parameter covariances	129
4.12 Best-fit transit model	130
4.13 Best fit RV model	131

CHAPTER 1

INTRODUCTION

A mere 16 years after the discovery of the first “Hot Jupiter” around a main sequence star (Mayor & Queloz 1995), over 500 exoplanets have been confirmed, and over twice that many candidates discovered. With that, we are beginning to develop a coherent theory of planet formation and evolution that encompasses the astonishing diversity of exoplanetary systems, including the mysterious class of Hot Jupiters heralded by the first discovery.

The transit method has been at the center of this revolution, not only because it has expanded the region of parameter space to which we are sensitive, but more importantly, it has provided a seemingly endless wealth of information about each planet (see Winn (2010) for a comprehensive review). For example, we can measure the planet radius and orbital inclination, and with the minimum mass from Radial Velocity (RV) studies, we find the true mass and average density, constraining the planet’s structure (Guillot 2005; Sato et al. 2005; Charbonneau et al. 2006; Fortney et al. 2006). Photometric observations during both primary transits and secondary eclipses enable the study of their atmospheres (Charbonneau et al. 2002;

Vidal-Madjar et al. 2003) and thermal emission (Deming et al. 2005; Charbonneau et al. 2006; Deming et al. 2006). Further, the projected angle between the spin axis of the star and the orbit of the planet can be measured via spectroscopic observations during transit to provide diagnostic information of the physical processes at work in the migration of Hot Jupiters (Gaudi & Winn 2007). Indeed, the combination of radial velocity and transit data provides the most thorough insight into a planetary system of any demonstrated planet detection and characterization method to date.

Currently, over 90% of nights have a least one transit to observe. Most observations of transiting planets are made on different, shared telescopes with different cameras by different groups and analyzed in different ways. Not only does this method make comparisons from system to system less reliable, but there is simply not enough telescope time, money, or manpower available to observe each of the currently known transits in this manner, much less the hundreds more that are expected to be discovered in the coming years.

We have built and automated a 0.5 meter telescope, dedicated to observing phenomena related to transiting planets, called DEMONEX, the DEDicated MONitor of EXotransits. In chapter 2, we outline the science goals of DEMONEX, describe the system, its operation, and present some of our first results.

Timing plays a particularly important role in the study of exoplanets. Indeed, many of the ways in which exoplanets are discovered involve the detection of transient

or time-variable phenomena, including the radial velocity, transit, microlensing, and astrometry techniques. Furthermore, in some cases much can be learned about planetary systems from the precise timing of these phenomena. As examples, the measurement of terrestrial parallax in microlensing allows one to infer the mass of the primary lens and so the planetary companion (e.g., Gould et al. 2009), and one can constrain the eccentricity of transiting planets by comparing the times of primary transits and secondary eclipses (e.g., Deming et al. 2005; Charbonneau et al. 2005; Knutson et al. 2007). Possibly the most promising application of timing in exoplanets, however, comes from transit timing variations (TTVs). With an exquisitely periodic phenomena like transiting planets, we will be able to measure many effects using the departures from strict periodicity, such as the gravitational perturbations from additional planets (Miralda-Escudé 2002; Holman & Murray 2005; Agol et al. 2005), trojans (Ford & Gaudi 2006; Ford & Holman 2007), and moons (Kipping 2009), stellar quadrupoles (Miralda-Escudé 2002), tidal deformations, general relativistic precession (Jordán & Bakos 2008; Pál & Kocsis 2008; Heyl & Gladman 2007), orbital decay (e.g., Sasselov 2003), and proper motion (Rafikov 2009).

In chapter 3, we summarize the effects one must consider in order to achieve timing accuracy of $1 \mu\text{s}$ – well beyond the accuracy that will likely be required by the exoplanet community for the foreseeable future.

Because of the wealth of information that can be derived from these planets, it is important to carefully consider the best ways to extract such information from the data sets we acquire in a consistent, homogeneous manner that makes results easy to compare. Ford (2005) concluded that Markov Chain Monte Carlo (MCMC) is the best way to extract robust inferences about the model parameters from the data, and recently this has become a standard tool for exoplanet research(e.g. Gregory 2005; Ford 2006; Winn et al. 2010).

In chapter 4, we provide a brief summary of the general problem of fitting data sets, how MCMC works, and why it is preferred over alternatives. Next, we detail how our analysis software, EXOFAST, works to fit both transit and radial velocity data simultaneously. All of the source code, in addition to online interfaces to the most useful codes are available online¹.

¹<http://astrutils.astronomy.ohio-state.edu/exofast/>

CHAPTER 2

DEMONEX: THE DEDICATED MONITOR OF EXOTRANSITS

2.1. INTRODUCTION

Transiting planets are enormously useful for measuring many detailed properties of extrasolar planets and planetary systems (see reviews by Charbonneau et al. (2007) and Winn (2009)). The most obvious properties that can be measured are the planet mass and radius, and so average density. These parameters alone allow for interesting constraints on the internal composition and structure of giant planets (Guillot 2005; Sato et al. 2005; Charbonneau et al. 2006; Fortney et al. 2006). In addition to these basic parameters, transiting planets enable the study of the atmospheres (Charbonneau et al. 2002; Vidal-Madjar et al. 2003), thermal emission (Deming et al. 2005; Charbonneau et al. 2006; Deming et al. 2006), rotation rate (Seager & Hui 2002), spin-orbit alignment (Gaudi & Winn 2007), and the existence of rings and moons (Brown et al. 2001; Barnes & Fortney 2004) of close-in giant planets. Transiting planets are also useful for probing the existence of other bodies in the system, via precise timing of the transits (Holman & Murray 2005; Agol et al.

2005; Steffen & Agol 2005; Ford & Gaudi 2006; Ford & Holman 2007). In particular, with sufficiently precise timing observations, it is possible to constrain the presence of planets with masses as small as the Earth (Agol & Steffen 2007).

Over the last ten years ground and space-based surveys for transiting planets have detected over 500 planets¹, with the majority ($\sim 2/3$) orbiting bright ($V \lesssim 13$) stars. This number will increase substantially as ongoing ground-based RV, ground-based transit, and space-based transit surveys continue to yield planets at unprecedented rates. However, as we are flooded with additional planets, it becomes difficult to follow up all of the known transits with traditionally scheduled, shared telescopes. Figure 2.1 shows the fraction of nights with at least one full primary transit or secondary eclipse visible from Arizona for three years beginning in May of 2008.

2.2. SCIENCE GOALS

Because there are so many events each night, we often have to prioritize between multiple events. This section is organized based on our priorities, with our highest priority first.

Observing all possible primary transits and averaging multiple transits together will refine the planetary properties and provide better constraints on planetary

¹ See <http://exoplanet.eu/> for an updated list and references.

formation and evolution models. The precision with which we can measure the planetary parameters is ultimately limited by stellar models, though very few planets have been characterized this precisely. In addition, with the times measured from each individual transit, we can look for Transit Timing Variations (TTVs), which would be likely to be caused by additional planets in the system perturbing the transiting planet's orbit (Agol et al. 2005). Lastly, because all planets are observed and reduced in the same manner, the sample will be homogeneous which will allow for robust planet-to-planet comparisons.

The Kilodegree Extremely Little Telescope (Pepper et al. 2007) is a wide field transiting planet survey that is about to produce viable transiting planet candidates. DEMONEX will provide invaluable follow-up on these candidates to rule out false positives and save time on larger telescopes.

The signal from each secondary eclipse is typically too weak to detect in a single event, but by stacking many eclipses of the same target, we will be able to measure or place scientifically interesting limits on the albedos of, or thermal emission from, the planets. This tells us about the atmospheric composition and energy budget of the planet. If we detect the secondary emission, we will also be able to constrain the eccentricity from the relative timing of the primary and secondary transits.

Although 80% of all planets have been discovered with Radial Velocity (RV) surveys, these systems generally provide significantly less information than transiting

systems. Among the 199 RV planets with periods longer than 10 days, three are expected to transit. Unfortunately, there has never been a systematic search of all RV planets for transits. If we found those expected to transit, they would be a valuable addition to the sparsely explored region of parameter space, owing to the different selection effects between transit and RV surveys.

DEMONEX is also generally useful for photometry of any bright object. With its array of filters, it is well suited for microlensing, supernovae, and many other ancillary sciences.

Many systems for which the Rossiter-McLaughlin effect has been measured shows the planet's orbit is aligned with the star's rotation axis (Bundy & Marcy 2000; Queloz et al. 2000; Winn et al. 2005). This, along with the coplanarity of the planets in the Solar System, suggests that many exoplanetary systems are likely coplanar, and thus the probability of finding additional transiting planets in these systems with known transiting planets is greatly increased from the a priori geometric probability. By monitoring the host stars of known transiting planets, we may find additional, smaller or long period planets. A system with two transiting planets would allow us to directly measure the stellar mass and remove the dependence on stellar models. These sorts of high-risk, very high-reward observations are well suited to small, dedicated telescopes like DEMONEX.

2.3. THE SYSTEM

DEMONEX is an f/8 Meade 0.5 m RCX400 on a MAX Mount (figure 2.2). The installed dew shield prevents dew, reduces scattered light, and protects it from dust. It also increases the seeing, which is acceptable because all of our targets are bright, and even desired as it reduces systematics induced by flat fielding errors and allows us to take longer exposures before reaching saturation, which increases our efficiency. There is a large gap at the end of the telescope that allows the corrector lens to move in order to focus the telescope. However, due to the large amount of dust that entered our system through the gap and the lack of drivers to control the focus robotically, we sealed it off and use a commercial focuser attached to the camera.

The science CCD is a Fairchild CCD3041 packaged by Finger Lakes Instrumentation (FLI). It is a 2K x 2K chip, with 15 μm pixels, for a field of view of 25.7'x25.7', and 0.75"/pixel. The large field of view is required to have an adequate number of bright comparison stars for differential photometry. Its gain is 1.6 e-/ADU, with a read noise of 12 e- and readout time of 6.75 s. Its exceptionally high z-band throughput ($\sim 50\%$) is well suited to transit observations because redder bandpasses are less affected by limb darkening.

The focuser and filter wheel are also from FLI. The CFW-5-7 filter wheel holds seven 50 mm square filters. Most observations are done in the z' band to reduce

the effects of limb darkening, but g' , r' , i' , V , R , I , and clear are also available for auxiliary science and experimentation.

A piggybacked Orion 80 mm ED and SBIG ST402-ME camera are used for autoguiding. An adequate guide star can usually be found with integration times well below one second. Differential flexure between the guider and science camera is compensated for by monitoring the position of the target in the science image and shifting the guide star appropriately.

We have an IP Power 9258, an IP-based relay that can individually power cycle the telescope, science camera, or guider via Visual Basic scripts in order to automatically recover from a variety of failures.

An Axis 215 PTZ network camera continuously monitors DEMONEX, and has been invaluable in remotely diagnosing and fixing problems. Its high sensitivity makes it an ideal choice for nighttime monitoring. We automatically generate time-stamped, high speed movies every night for inspection the following morning. When combined with our verbose engineering log, it is easy to pinpoint failures when they occur. Its ability to pan, tilt, and zoom also makes it an effective tool for more detailed diagnosis and to walk on-site support staff through more difficult procedures.

All of our components are plugged into a Powerware 9125, an uninterruptable power supply (UPS) with over an hour of runtime. In case of a power failure, this is

more than enough time to park the telescope and cleanly shut down the computers until power is restored. More important, the UPS isolates the output power from all input anomalies. The MAX mount is highly susceptible to electronic noise, and the power at our remote location is very noisy. Without this power conditioning, the mount behaves irrationally, interpreting power spikes as unintended commands. The UPS is also able to remotely power cycle all components simultaneously to recover from serious failures.

Everything is controlled via computer in the control room 35 feet away, as discussed in the next section.

Several telescopes, including DEMONEX, are housed in a large observatory with a roll-off roof at the Winer Observatory in Sonoita Arizona. All normal operation is autonomous, as described in section 2.4. Most problems can be diagnosed and fixed remotely from our location in Columbus, Ohio.

The observatory director manages a weather station and closes the roll-off roof if there are threatening conditions. The roof status, as well as weather data and other diagnostic tools are posted on a webpage, which is queried and recorded in a log before each exposure. If the roof is closed, we wait until it opens.

2.4. AUTOMATION

DEMONEX is completely automated using a Windows XP box equipped with TheSky6, CCDSoft, astrometry.net, and VBScripts for observations, and a Linux box with IDL, SExtractor, and astrometry.net for reductions. Each night, DEMONEX begins taking biases and darks then waits for the Sun to set to -3 degrees, when it is low enough for twilight flats. It points to the optimally flat part of the sky (Chromey & Hasselbacher 1996) and takes flats in each band that will be used during the night, beginning with the reddest band. It scales the exposure time to keep the background at 20,000 counts. We impose a minimum 10 second exposure time to avoid shutter effects, and we stop when the background from a 60 second exposure is less than three times the bias level. Since we typically observe in z band, we only take z band flats. In this case, we get around 17 flats each twilight.

After flats are finished, we wait until the Sun is below -12 degrees and point to our first target. The MAX mount has no absolute motor encoders, and its home position is not reliable enough to ensure our target always falls inside our field of view. Sometimes, the pointing can be wrong by 5 degrees or more. We use a local copy of astrometry.net (Lang et al. 2010), which provides a coordinate solution to any astronomical image without any knowledge of its pointing, to determine the telescope's position and center the object.

We observe each target visible according to the priorities outlined in section 2.2. It continues until the sun rises, and then takes another set of flats in the morning for quality control, this time starting with the bluest band.

It then goes to its home position, turns off, compresses and backs up the data, emails a log for the night, and logs into the linux box to begin reductions. If a fatal error is encountered, it attempts to park the telescope and sends a text message and email, attaching a verbose engineering log to aide in diagnosis.

The reduction pipeline does the usual bias and dark calibrations. Since we use twilight flats, we must employ some quality control to avoid using flats contaminated by clouds. A separate, mode-scaled, median-combined flat field is created from the flats in the evening and morning. We divide one by the other and record the variance as a figure of merit. We compare that figure of merit with the same quantity for previous nights, except we add an empirically determined penalty based on how much time has passed between when the flats and science observations were taken, and use the flat with the lowest figure of merit. While this is a conservative approach that will eliminate good flats taken either in the evening or morning if the other is bad (indeed, only 10% of the flats taken are actually used), it is simple and effective.

For each object, a reference image is selected by determining which has the most stars with the smallest Full Width at Half Maximum (FWHM) with SExtractor(Bertin & Arnouts 1996). A coordinate solution is performed on this

reference image with `astrometry.net`, and the offsets and rotations of all other images are determined using a modified version of ISIS's `interp` script (Alard & Lupton 1998; Alard 2000), fed with the SExtractor source positions. Aperture photometry is performed with the IDL astronomy library routine `APER`, using the reference image positions transformed with the calculated offsets, with many aperture sizes. For each aperture size, the target star is divided by the ensemble of stars that minimizes its out-of-transit RMS. The aperture with the lowest out-of-transit RMS is used, The resulting lightcurves are emailed for visual inspection an hour or two after the final flat for the night was taken. A webpage is also generated with thumbnails and statistics for each image for inspection.

2.5. RESULTS

An example of three primary transits of TrES-3b, is shown in figure 2.3. The co-added and binned version is below. For each individual transit, we are able to get transit timing precision of about 30 seconds, which is good enough to detect an Earth-mass planet in 2:1 resonance around a hot Jupiter (Agol et al. 2005).

By co-adding many transits of the same planet, we are able to reduce the uncertainty in the planetary parameters to the point at which it is dominated by stellar models. Many follow-up efforts suffer from significant red-noise that prevents the accurate measurement of the planetary parameters (Alonso et al. 2009).

However, figure 2.4 shows the RMS of the residual to the best fit model of the three transits as a function of the number of points binned. The solid black line shows the theoretically ideal, pure white noise. A signal dominated by red noise would flatten out, but our data is consistent with pure white noise to at least 0.03%.

We have contributed data for several other projects, including data for the microlensing event OGLE-2008-BLG-279 (Yee et al. 2009), supernovae, and variable stars.

We chose four high-metallicity stars to focus our out-of-transit monitoring campaign, one for each season which we observe when nothing else is available.

During the past two years, DEMONEX has generated nearly 8 terabytes of data while observing 325 primary transits, 236 secondary transits, 14 RV planets, 15 ancillary science targets, and the four targets for additional transits 108 times. As it continues to gather data, we are extending the software pipeline to analyze the light curves and fit the planetary parameters for each system.

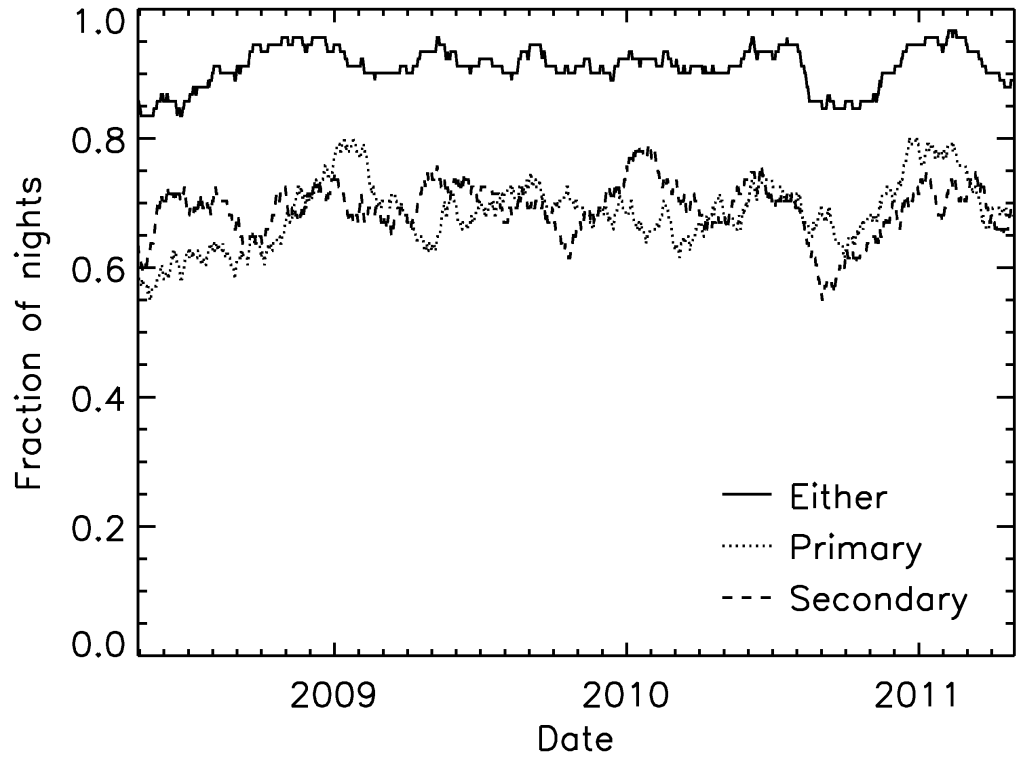


Fig. 2.1.— The fraction of nights with a primary transit (dotted), secondary transit (dashed), or either (solid) over a 3 year period for all transiting planets visible from Arizona, starting in May 2008.



Fig. 2.2.— DEMONEX during sunset awaiting observations at Winer Observatory in Sonoita, Arizona.

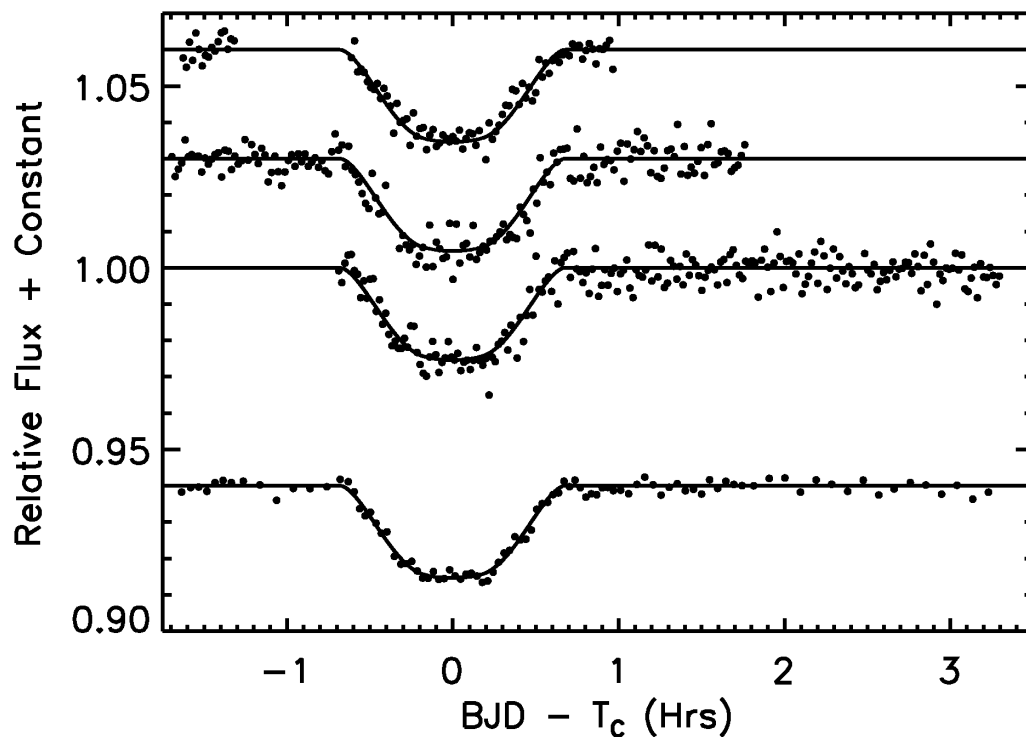


Fig. 2.3.— The light curve of three primary transits of TrES-3b taken with DEMONEX operating robotically. The solid line is the best fit model. The bottom light curve shows the three above averaged together. Each are offset by a constant for clarity.

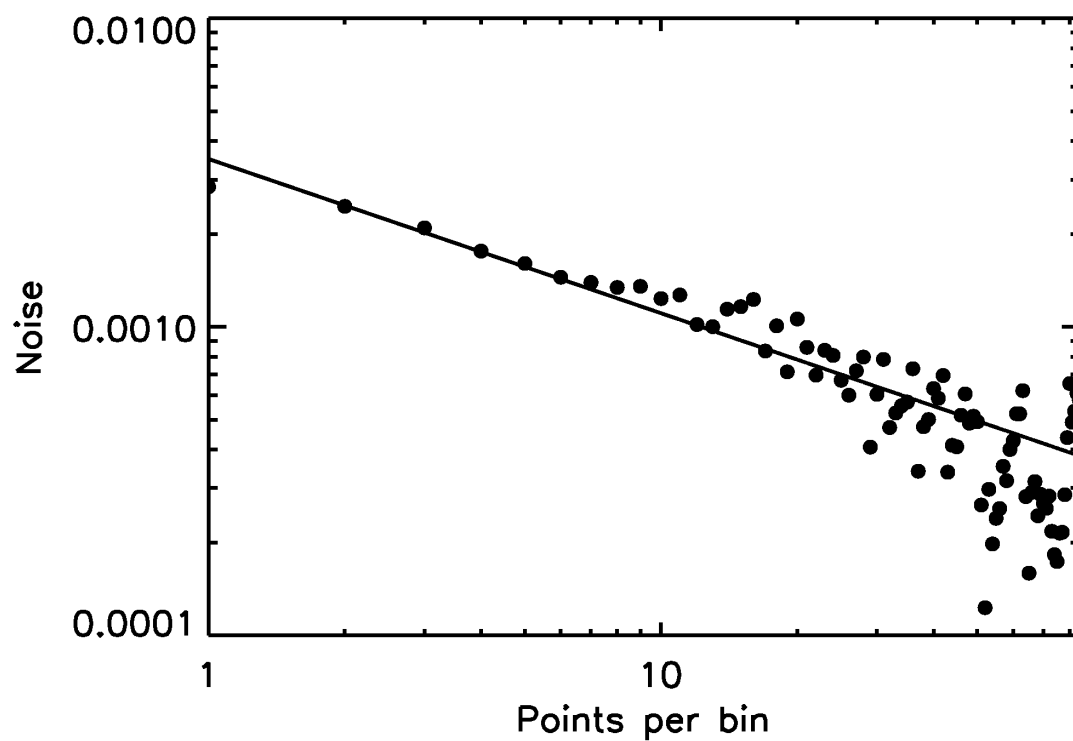


Fig. 2.4.— The noise in the data as a function of points per bin. The solid black line shows the theoretically predicted line for the case of the ideal, pure white noise.

CHAPTER 3

ACHIEVING BETTER THAN 1 MINUTE ACCURACY IN THE HELIOCENTRIC AND BARYCENTRIC JULIAN DATES

3.1. INTRODUCTION

Precise timing of astrophysical events is one of the fundamental tools of astronomy, and is an important component of essentially every area of study. There are two basic sources of uncertainty in timing: the astrophysical data characterizing the event, and the time stamp with which the event is referenced. Unfortunately, since the accuracy of the time stamp is something that is often taken for granted, the improvements in data are sometimes not accompanied (or are not uniformly accompanied) by the requisite improvements in accuracy of the time stamp used. This situation can lead to confusion, or even spurious inferences.

Because of their great potential, TTVs have become the focus of many groups. The typical data-limited transit timing precisions of most observations are around 1 minute, with the best transit time precision yet achieved of a few seconds (Pont et al. 2007). However, as discussed above, accuracies of transit times are limited

not only by the data themselves, but also by the time stamp used. In order to make these difficult measurements useful, it is critical that a time stamp be used that is considerably more accurate than the uncertainty due to the data themselves. Furthermore, since thorough characterization of TTVs will require the use of all available data spanning many years from several groups, this time stamp must be stable in the long term, and all groups must clearly convey how it was calculated.

Unfortunately, we have discovered that, in the exoplanet community, the Julian Date (JD) and its geocentric (GJD), heliocentric (HJD), and barycentric (BJD) counterparts are currently being quoted in several different, and often unspecified, time standards. In addition, the site arrival time and its time standard is not quoted. This general lack of homogeneity and specificity leaves quoted time stamps ambiguous at the 1 minute level. More alarmingly, the most commonly-used time standard, the Coordinated Universal Time (UTC), is discontinuous and drifts with the addition of each leap second roughly each year.

The pulsar community has solved the problem of precise timing well beyond the level that is currently necessary for exoplanet studies, and we can benefit from the techniques they have developed over the past 40 years. In particular, their current state of the art program (`TEMPO2`) models pulsar arrival times to 1 ns precision (Hobbs et al. 2006; Edwards et al. 2006). This program is highly specialized and generally cannot be applied outside of pulsar timing observations, but many of the effects they consider are relevant to optical observers in the exoplanet community.

Section 3.2 provides the background required to understand each of the effects that could change the arrival time of a photon. They are listed in order of decreasing magnitude, so latter subsections can be ignored for low-precision measurements.

Section 3.3 discusses the practical limitations to achieving high-precision timing. We begin with the effects which may cause errors that are comparable to or exceed the BJD correction. These should be read and understood by everyone. We continue with remaining effects, in order of decreasing magnitude, which can be ignored for low-precision (> 30 ms) measurements. We conclude § 3 by listing additional effects, the errors due to which are negligible ($< 1\mu\text{s}$).

We begin § 3.4 by detailing the procedure one must follow in order to calculate the BJD_{TDB} , which is designed to be a useful reference for those already familiar with the concepts of precision timing. In the latter part of this section, we describe our particular IDL and web-based implementation of this procedure.

Lastly, in the Appendix, we discuss some of our specific findings about the time stamps currently in use and how these are calculated throughout the exoplanet community.

While we focus on the effects of timing on the optical/infrared exoplanet community, timing precision of order 1 minute is necessary for many other areas, such as the study of rapidly rotating white dwarfs (Euchner et al. 2006). This article should be equally applicable in such cases.

3.2. THEORETICAL TIMING PRECISION

The biggest source of confusion comes from the fact that time standards and reference frames are independent from one another, even though there are many overlapping concepts between the two. We will use the following terminology: “reference frame” will refer to the geometric location from which one could measure time – different reference frames differ by the light-travel time between them; “time standard” will refer to the way a particular clock ticks and its arbitrary zero point, as defined by international standards; and “time stamp” is the combination of the two, and determines the timing accuracy of the event.

The BJD_{TDB} , the time stamp we advocate, can be calculated using the equation:

$$BJD_{TDB} = JD_{UTC} + \Delta_{R\odot} + \Delta_C + \Delta_{S\odot} + \Delta_{E\odot}, \quad (3.1)$$

where JD_{UTC} is the Julian Date in Coordinated Universal Time; $\Delta_{R\odot}$ is the Rømer Delay, discussed in § 4.5.4; Δ_C is the clock correction discussed in § 3.2.2; $\Delta_{S\odot}$ is the Shapiro delay discussed in § 3.2.3; and $\Delta_{E\odot}$ is the Einstein delay, discussed in § 3.2.4.

The order of these terms is such that they are of decreasing magnitude, so one need only keep the terms up to the precision required. The timing precision required by current exoplanet studies (~ 1 s) requires only the terms up to and including Δ_C .

Because future Solar System ephemerides may enable more precise calculations of the arrival time at the Barycenter, or in order to allow others to check that the original conversion was done accurately enough for their purpose, the site arrival time (e.g., the JD_{UTC}) should always be quoted in addition to the BJD_{TDB} .

3.2.1. REFERENCE FRAMES: THE RØMER DELAY

Due to the finite speed of light, as the Earth travels in its orbit, light from an astrophysical object may arrive early or be delayed by as much as 8.3 minutes from the intrinsic time of the extraterrestrial event. This is called the Rømer delay, Δ_R , in honor of Ole Rømer's demonstration that the speed of light is finite. Since most observers cannot observe during daylight, a bias is introduced and in practice the delay (as distinct from the early arrival time) is only as much as 7 minutes, for a peak-to-peak variation of 15 minutes. Figure 3.1 shows an example of this effect for a maximally affected object on the ecliptic. In order to show the observational bias, our example assumes the object is at 0^{h} right ascension and 0° declination. This curve shifts in phase with ecliptic longitude and in amplitude with ecliptic latitude.

We also place our observer at the Earth's equator, but note that the asymmetry will be larger at different latitudes.

The solution to this problem is to calculate the time when a photon would have arrived at an inertial reference frame. This time delay is the dot product of the unit vector from the observer to the object, \hat{n} , and the vector from the origin of the new reference frame to the observer, \vec{r}

$$\Delta_{R\odot} = \frac{\vec{r} \cdot \hat{n}}{c}, \tag{3.2}$$

where c is the speed of light and \hat{n} can be written in terms of its right ascension (α) and declination (δ),

$$\hat{n} = \begin{pmatrix} \cos(\delta) \cos(\alpha) \\ \cos(\delta) \sin(\alpha) \\ \sin(\delta) \end{pmatrix}. \tag{3.3}$$

This equation is general as long as α , δ , and \vec{r} are in the same coordinate system (e.g., Earth mean equator J2000) and the object located at (α, δ) is infinitely far away. Other forms of this equation in the literature assume that we have the angular coordinates of the new origin or that the Earth and the new origin are in the same

plane (e.g., Binnendijk 1960; Henden & Kaitchuck 1982; Hirshfeld & Sinnott 1997), but we explain in § 3.4 why this form is most practical for calculating the delay.

The HJD, which uses the Sun as the origin of the new reference frame, is only accurate to 8 s because of the acceleration of the Sun due primarily to Jupiter and Saturn (Fig. 3.2). It was popular when people first began considering this effect because it is relatively simple to calculate from tables without a computer (Landolt & Blondeau 1972), and remains popular because self-contained algorithms exist to approximate it without any external tables (e.g., Duffett-Smith 1989). However, because the HJD is not useful when accuracies of better than 8 s are needed, most of the algorithms in use today use approximations that are only precise at the 1 s level, and it becomes impossible to back out the original JD from the HJD unless we know the exact algorithm used.

Because of these problems, the HJD was formally deprecated by International Astronomical Union (IAU) Resolution A4 in 1991, in favor of the BJD, a time referenced to the Solar System Barycenter (SSB).

The analogous correction to the Rømer delay in our Solar System can also be significant in the target system. We refer to this as Δ_R . For example, for transiting planets with $a \approx 0.06$ AU, Δ_R can be as large as 30 s. In general, the position of the planet during primary transit has become the unspoken standard reference frame for

transiting planets, while the host star’s photosphere is the unspoken standard for Radial Velocity (RV) planets.

In theory, the timing would be much more stable in the target’s barycentric reference frame, but the accuracy with which we can convert to this frame depends on the measurements of the system. Since different observers may use different values as measurements improve, quoting the JD in the frame of the target’s barycenter may obfuscate the long term reliability of timing. Therefore, we argue it is better to quote Julian Date in the SSB reference frame, and correct for Δ_R only when comparing observations at different phases in the planet’s orbit.

This Δ_R correction is not necessary for TTVs of the primary transit, since the planet is always in the same phase. Nevertheless, we should explicitly state the object’s reference frame to avoid any potential ambiguity, particularly when comparing any combination of primary transits, secondary transits, RVs, and another primary transit of a different planet in the same system, when it may not be obvious which origin is being used.

For RV measurements, which are taken at many different phases, the effect is much smaller and can generally be ignored because the star’s orbit around the barycenter is small. For a typical hot Jupiter, (i.e., a Jupiter mass planet in a 3 day orbit around a solar mass star), the maximum time difference in the RV signal (for an edge-on orbit) is 20 ms, which would change the measured RV by $\sim 50 \mu\text{m}$

s^{-1} . While planets farther out will cause a larger timing offset, the difference in the measured RV is even smaller.

3.2.2. TIME STANDARDS: THE CLOCK CORRECTION

To be clear, the JD can be specified in many time standards (Seidelmann 1992), and while the IAU has made no explicit statement regarding the allowed time standards of the GJD, HJD, or BJD, their meaning in any given time standard is unambiguous. Unfortunately, they have been specified in many standards, usually implicitly.

However, the particular time standard used affects how useful the time stamp is as an absolute reference. We must be careful not to directly compare BJDs or HJDs in different time standards, as each has different offsets, periodic terms, and/or rates, which can introduce systematic errors of over 1 minute. For this reason, it is critical that any stated BJD or HJD also specify the time standard used when one-minute accuracies are important, and the uncertainty of a time that is quoted without a standard should be assumed to be at least 1 minute.

First, it may be useful to summarize the relevant standards of time:

Universal Time, UT1 – Defined by the mean solar day, and so drifts forward and backward with the speeding and slowing of the Earth’s rotation. Generally,

it slows due to the tidal braking of the Moon, though changes in the Earth's moment of inertia and complex tidal interactions make its exact behavior unpredictable. It is rarely used directly in astronomy as a time reference, but we mention it for context.

International Atomic Time, TAI – Based on an average of atomic clocks all corrected to run at the rate at the geoid at 0 K, with 1 s equal to “9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom,” as defined by Resolution 1 of the thirteenth meeting of the Conférence Générale des Poids et Mesures (CGPM) in 1967. This definition is based on the duration of the Ephemeris Time second, which was previously defined as 1/31,556,925.9747 of the tropical year for 1900 January 0 at 12 hours Ephemeris Time by Resolution 9 of the eleventh CGPM in 1960. TAI is the fundamental basis for many other time standards, and is the default time standard of the Sloan Digital Sky Survey.

Coordinated Universal Time, UTC – Runs at the same rate as TAI, except that it is not allowed to differ from UT1 by more than 0.9 s. Every 6 months, at the end of 31 December and 30 June, the International Earth Rotation and Reference Systems Service (IERS)¹ may elect to add (or subtract) a leap second to UTC in order to keep it within 0.9 s of UT1. UTC is therefore

¹At <http://www.iers.org>.

discontinuous and drifts relative to TAI with the addition of each leap second, which occur roughly once per year. As of January 2009, the current number of leap seconds, N , is 34. The full table of leap seconds is available online and is typically updated several months in advance of when an additional leap second is to be added². UTC is the current international standard for broadcasting time. As a result, when a modern, network-connected computer's clock is synchronized to a Network Time Protocol (NTP) server, it will be in UTC. Thus, this is the system of time most familiar to astronomers and non-astronomers alike (modulo time zones and daylight savings time).

Universal Time, UT – An imprecise term, and could mean UT1, UTC, or any of several other variations. In general, such imprecise language should be avoided, as the potential ambiguity is up to 1 s. In the context of a time stamp, it is likely UTC, but some people may intentionally use UT to imply 1 s accuracy. While explicitness is preferred (i.e., UTC \pm 1 s), any time stamp quoted in UT should be assumed to be uncertain at the 1 s level unless the time standard has been independently verified.

Terrestrial Time, TT(TAI) – A simple offset from TAI of 32.184 s released in real time from atomic clocks and never altered. This offset is to maintain continuity between it and its predecessor, the Ephemeris Time (ET).

²At <ftp://maia.usno.navy.mil/ser7/tai-utc.dat>.

Terrestrial Time, TT(BIPM) – A more precise version of TT(TAI). The International Bureau of Weights and Measures (BIPM) reanalyzes TT(TAI) and computes a more precise scale to be used for the most demanding timing applications. The current difference between TT(TAI) and TT(BIPM) is $\sim 30\mu\text{s}$, and must be interpolated from a table maintained by the BIPM and published online with a 1 month delay³.

Terrestrial Time, TT – Sometimes called Terrestrial Dynamical Time (TDT), can refer to either TT(TAI) or TT(BIPM). From this point on, we will not make the distinction, but when accuracies of better than $30\mu\text{s}$ are required, TT(BIPM) must be used.

Barycentric Dynamical Time, TDB – Corrects TT for the Einstein delay to the geocenter, $\Delta_{E\odot}$, which is the delay due to time dilation and gravitational redshift from the motions of the Sun and other bodies in the Solar System. The conversion from TT to TDB cannot be written analytically, but is usually expressed as a high-order series approximation (Irwin & Fukushima 1999). The difference is a predominantly a periodic correction with a peak-to-peak amplitude of 3.4 ms and a period of 1 yr. TDB was slightly modified by IAU Resolution B3 in 2006, converging on the same definition as the JPL Ephemeris Time, T_{eph} , also called Coordinate Time (CT) in the JPL ephemerides of Solar System objects.

³At <ftp://tai.bipm.org/TFG/TT%28BIPM%29>.

Barycentric Coordinate Time, TCB – Physically and mathematically

equivalent to the TDB as defined in 2006 (Standish 1998), and differs only by an offset and rate of about 0.5 s yr^{-1} due primarily to time dilation in the Sun's gravitational potential. TDB and TCB were roughly equal at 1977 Jan 1.0 TAI, and now differ by about 16 s.

Caution must be always be exercised, however, as these definitions are subject to change at any point, though usually with a few year's notice.

Assuming the time is measured according to current definition of UTC, then the clock correction, from UTC to TDB, can be written as the sum of the corrections from UTC to TAI, TAI to TT⁴, and TT to TDB:

$$\Delta_C = N + 32.184s + (TDB - TT). \quad (3.4)$$

Of course, if one wishes to express the BJD in another time standard (or start with something other than UTC), the clock correction would change accordingly. However, not every time standard is well-suited to precise, astrophysical time stamps, and the use of any time standard other than TDB should be viewed simply as an adequate approximation to TDB.

⁴Ignoring the TT(BIPM) correction.

Most readily available programs that calculate the time stamp assume that the user has already applied the UTC-to-TT part of this clock correction, which is often not true. We feel this assumption has contributed the widespread confusion regarding time stamps.

As this last point is our primary motivation for writing this article, we elaborate here on the effects of time standards on the reliability of time stamps. For the sake of simplicity, we only discuss the effects of time standards on the BJD. Each of these effects also applies to the HJD, though the improvement in accuracy of the time stamp is negligible compared to the accuracy of the HJD reference frame for all but the UTC time standard.

The least preferred, though most commonly used, time standard for the BJD is UTC (BJD_{UTC}), and is equivalent to ignoring Δ_C altogether. Because UTC is discontinuous and drifts with the addition of each leap second, comparing two BJD_{UTC} time stamps could result in spurious differences if any leap seconds have been introduced between observations. Therefore, 1 s timing accuracies cannot be achieved using the BJD_{UTC} over a span that straddles the addition of one or more leap seconds (roughly 1 yr). Figure 3.3 shows the difference between the BJD_{UTC} time stamp with the uniform BJD_{TDB} time stamp (described below) from 1961 Jan 1, when UTC was defined (though its definition has evolved over the years), to 2010 December 31, the furthest future date for which the value of UTC can be accurately predicted at this writing.

BJD in TT (BJD_{TT}), which is equivalent to ignoring the (TDB-TT) term in equation (3.4), corrects for the discontinuity and drift introduced by leap seconds and is appropriate for timing accurate to 3.4 ms.

BJD in TDB (BJD_{TDB}) is usually the best time stamp to use in practice, as it further corrects the BJD_{TT} for all known effects on the motions, and therefore rates, of our atomic clocks. While BJD_{TDB} is not perfect, any more accurate time stamp is unique for each target.

BJD in TCB (BJD_{TCB}) corrects for the gravitational potential, primarily from the Sun, which causes the clock to run slower than it otherwise would. However, if one is concerned about effects of this magnitude, the analogous correction of putting it in the gravitational potential of each observed object is also required. Since these two rates are small (10^{-8}), and opposite in sign, we believe it is best to ignore BJD_{TCB} except perhaps as an intermediate step in calculating the target-specific frame. Technically, the use of TCB was recommended by IAU Resolution B1.3 in 2000. However, because of the greater practicality of using TDB (see § 3.4), and the drifting difference between TCB and TDB and TT, we believe its use will only lead to confusion without any foreseeable benefit to the exoplanet community.

3.2.3. SHAPIRO DELAY

The Shapiro delay, Δ_S (Shapiro 1964), is a general relativistic effect in which light passing near a massive object is delayed. For an object at an angle θ from the center of the Sun, the Shapiro delay is

$$\Delta_{S\odot} = \frac{2GM_{\odot}}{c^3} \log(1 - \cos \theta). \quad (3.5)$$

This can be as large as 0.1 ms for observations at the limb of the Sun, but for objects more than 30° from the Sun, the correction is less than $20 \mu\text{s}$.

There is also the analogous correction, Δ_S , for the target system. Similar to Δ_R , Δ_S depends on the measurements of the target, which may be refined over time. Therefore, the time should generally be quoted without Δ_S , but include it when comparing times where this could be significant.

3.2.4. EINSTEIN DELAY

As we discussed in § 3.2.2, relativity dictates that the motion of the observer influences the rate at which the observed clock ticks. The use of TDB corrects for

an observer moving with the geocenter, but in reality we observe from the surface of the Earth or from a satellite, for which there is an additional term to $\Delta_{E\odot}$:

$$\Delta_{E\odot} = \frac{\vec{r}_o \cdot \vec{v}_{\oplus}}{c^2}. \quad (3.6)$$

Here, \vec{r}_o is the location of the observer with respect to the geocenter, and \vec{v}_{\oplus} is the velocity of the geocenter. Again, there is an analogous correction, Δ_E , for the target system, which should be ignored when quoting the time but included when comparing times if necessary.

3.3. PRACTICAL CONSIDERATIONS

Of course, the accuracy of the output time stamp is only as good as our assumptions and the inputs needed to specify the time standard and reference frame. Here we discuss their effects on the accuracy of the time stamps. The first four subsections, through § 3.3.4, have no reasonable upper bound. In each case, the accuracy of the inputs must be evaluated depending on the accuracy of time stamp required. The later three subsections through § 3.3.7 are organized in decreasing magnitude, all of which can be ignored for accuracies no better than 21.3 ms. Finally, § 3.3.8 discusses effects of less than 1 μ s that we have ignored.

3.3.1. COORDINATES

An error of 1' in the position of the target amounts to a timing error of as much as 0.28 s in the BJD (Fig. 3.4). Such error would be common if the coordinates of the field center are used instead of the specific object's coordinates. In particular, if doing a survey, one may wish to assign the same BJD to all objects in a given frame. However, with a 10° offset, which is possible with some wide-field transit surveys, the error can be as large as 200 s for objects at the edge of the field. An error of 0.25'' will yield 1 ms timing offsets, and 0.25 mas accuracy is necessary for 1 μ s timing.

3.3.2. COMPUTER CLOCK

The accuracy of a typical computer clock depends on its intrinsic stability, the computer workload, its operating system, and the reliability of the network connection. Older computers with a parallel port CCD interface may produce unreliable timing because the clock may slow or stop completely during CCD readout. Without any special effort, a modern Windows machine with a network connection is accurate to ~ 2 seconds⁵, and with third-party software like Dimension 4⁶, we have found it to be stable to 0.1 s. An NTP-synchronized Linux machine is typically accurate to ~ 50 ms.

⁵At <http://support.microsoft.com/kb/939322>.

⁶At <http://www.thinkman.com/dimension4>.

Of course, the stability of the clock only sets a lower limit on the absolute accuracy of the time recorded in the FITS image header. NTP synchronization attempts to measure and compensate for network latency, but the accuracy of time stamps also depends on the particular software package taking the image and the hardware it uses, which is difficult to calibrate. Unless independently verified, the time recorded in image headers should not be trusted to better than 0.25 s. However, various solutions exist for higher precision timing, such as GPS-triggered shutters.

In particular, it is worth emphasizing the 1 s error in the *Hubble Space Telescope* (HST) clock and potential 6.5 s error in the Kepler clock, described in more detail in the Appendix, both of which have already achieved data-limited transit timing precisions of that order (e.g., Pont et al. 2007; Kipping & Bakos 2010).

3.3.3. FLUX-WEIGHTED MEAN TIME OF EXPOSURE

When calculating the time of exposure of an image, we typically use the time at midexposure. However, the precise time of exposure that should be used is the flux-weighted mean time of exposure. The magnitude of this error depends on the intrinsic stability of source, the stability of the atmosphere, and the exposure time. In the diabolical case of a cloud completely covering the object during one half of the exposure, the error could be as large as half the exposure time. For a typical Hot Jupiter that dims by 1% over the course of a 15 minute ingress, the error introduced

into the time stamp during ingress or egress by using the mid exposure time is 0.25% of the exposure time – 150 ms for a 1 minute exposure. Near the peak of some high-magnification microlensing events (i.e., Gould et al. 2009), the flux may double in as little as 6 minutes. During such instances, using the mid exposure time will result in a time stamp error of as much as 1/4 of the exposure time – 15 s for a 1 minute exposure.

3.3.4. PLANE-WAVE APPROXIMATION

Equation (3.2) assumes the object is infinitely far away, and therefore the incoming wavefronts are plane waves. In reality, the wavefronts are spherical, which introduces a distance-dependent, systematic error. The maximum error introduced by the plane-wave approximation is 1000 s for the Moon⁷, 100 s for the Main Asteroid Belt, 5 s for the Kuiper Belt, 1 ms at the distance of Proxima Centauri, and 150 ns at the distance to the Galactic Center.

The fully precise equation, assuming spherical wavefronts, is

$$\Delta_{R\odot} = \frac{|\vec{r} + \hat{n}d| - d}{c}, \quad (3.7)$$

⁷The maximum error for the Moon is not because of the departure from the plane wave, but because the plane-wave formalism can place the Moon on the wrong side of the SSB.

where d is the distance from the observer to the target. In such instances, the distances must be derived from precise ephemerides for both the target and observer.

Although this form is generally applicable, it is not generally practical because at large distances, double precision floating point arithmetic cannot reliably recover the small difference between $|\vec{r} + \hat{n}d|$ (the distance from the barycenter to the target) and d .

To solve this problem, the pulsar community (e.g., Lorimer 2008)⁸ uses the two-term Taylor expansion of equation (3.7) about $1/d = 0$:

$$\Delta_{R\odot} \approx \frac{\vec{r} \cdot \hat{n}}{c} + \frac{\vec{r} \cdot \vec{r} - (\vec{r} \cdot \hat{n})^2}{2cd}. \quad (3.8)$$

One may recognize the first term as plane wave approximation (eq. [3.2]). In practice, the accuracy of this two-term approximation exceeds the accuracy of the “exact” calculation using double precision at a distance of 10,000 AU (0.05 pc). At the distance of Proxima Centauri, the accuracy of this approximation is at worst 1 ns.

However, this Taylor expansion is divergent when $d < |\vec{r}|$. It should never be used for objects inside 1 AU and may still be inadequate for other objects inside the Solar System. It has a maximum error of 1 day for the Moon, 20 s for the Main

⁸Note: the sign of the second term in Lorimer 2008, eq. 9, should be - rather than +.

Asteroid Belt, and 40 ms for the Kuiper Belt. In these cases, using the exact formula (eq. [3.7]) may be easiest.

Therefore, we recommend that for precise calculations of any Solar System body, equation (3.7) should be used. For better than 1 ms timing of any object outside the Solar System, equation (3.8) should be used.

3.3.5. GEOCENTER

Most readily available time stamp calculators use the position of the geocenter, rather than the location of the observer on the surface of the Earth. Neglecting the light-travel time from the surface of the Earth to the center introduces a 21.3 ms amplitude variation with a period of 1 sidereal day. In practice, most observers can only observe their targets at night, creating a systematic bias of between 8 ms and 21.3 ms (Fig. 3.5).

3.3.6. BJD'_{UTC} VERSUS BJD_{UTC}

Usually, when people calculate the BJD_{UTC} , they neglect Δ_C , and input JD_{UTC} for algorithms designed to take JD_{TT} . This effectively uses the positions of the Earth, Sun, and planets offset in time by $32.184 + N$ s to calculate the correction. When the BJD_{UTC} is calculated in this manner, we denote it as BJD'_{UTC} . The correct way to calculate the BJD_{UTC} would be to first calculate the BJD_{TDB} , then

subtract the Δ_C correction⁹. Figure 3.6 shows, for a maximally affected object on the ecliptic, an example of the difference between the BJD'_{UTC} and the fully correct BJD_{UTC} . Fortunately, this amounts to at most a 13.4 ms difference (though growing with the UTC-TT difference), which is below the precision of most clocks and the geocentric correction that is usually ignored. Therefore, to an accuracy of ~ 50 ms, one can safely say $\text{BJD}_{\text{TDB}} \approx \text{BJD}_{\text{TT}} \approx \text{BJD}'_{\text{UTC}} + \Delta_C$, making it easy to convert currently published values of BJD'_{UTC} to the superior BJD_{TDB} .

3.3.7. COMPUTER PRECISION

Representing JD as a double precision floating-point number limits the accuracy to about 1 ms, and any operation done on the full JDs will be even less accurate. Many programs require the use of a Reduced or Modified Julian Date, and/or can return the JD to BJD offset in seconds, but care must be taken at every step of the way never to store the full JD as a double precision number if 1 ms precision is required.

3.3.8. NEGLIGIBLE ($< 1\mu\text{s}$) EFFECTS

The Shapiro delay occurs for other bodies as well, but observations at the limb of Jupiter only delay light by 200 ns.

⁹However, the BJD_{UTC} is a poor approximation to BJD_{TDB} and should not be used.

Typical modern, commercial GPS units use the World Geodetic System (WGS84), which is referenced to the International Terrestrial Reference System (ITRS) with an error of about 15 m, which amounts to a 50 ns error in the time stamp.

The index of refraction of the atmosphere is not exactly 1 and changes with its composition, temperature, and pressure, which changes the speed of light. However, the largest reasonable deviation due to this effect is only tens of ns.

The pulsar community must specify a frequency-dependent dispersion measure delay. At radio wavelengths (21 cm), the delay can be as much as 1 s, but the dispersion delay contributes at less than 1 μ s shortward of $\sim 300 \mu$ m.

3.4. CALCULATING THE BJD_{TDB}

The most practical way to precisely calculate the BJD_{TDB} time stamp is using JPL's DE405 ephemeris¹⁰. It contains the position of thousands of bodies in the Solar System, including the Sun, planets, spacecraft, moons, asteroids, and comets. It is oriented to the International Celestial Reference Frame (ICRF), which is consistent with the FK5 system at J2000.0 within the 50 mas error of FK5, and has its origin at the SSB with its axes fixed with respect to extragalactic objects (Arias et al. 1995). Therefore, it is recommended to use the 3-space Cartesian coordinates

¹⁰At <ftp://ssd.jpl.nasa.gov/pub/eph/export/DE405/de405iom.ps>.

retrieved from the JPL DE405 ephemeris directly with the J2000 object coordinates in equation (3.2).

The following is an outline of the steps required to properly calculate the BJD_{TDB} .

1. Calculate the midexposure time in JD_{UTC} . Most FITS image headers give DATE-OBS in UTC at the beginning of exposure. If high precision is required (< 1 s), read the caveats about clock precision in § 3.3.2 carefully; depending on the sky conditions or variability of the object, one may need to account for the flux-weighted mean time of exposure.
2. Convert the midexposure time to JD_{TDB} by applying Δ_C (eq. [3.4]). For times accurate to 3.4 ms, one can use the simpler JD_{TT} and calculate BJD'_{TT} (using the positions of planets delayed by the TT-TDB offset). The difference between BJD'_{TT} and BJD_{TT} is no more than ~ 200 ns, which is well below the precision of the BJD_{TT} . If better than $30 \mu\text{s}$ precision is required, the TT(BIPM) - TT(TAI) offset must be applied.
3. Retrieve the JPL ephemeris of the observing station for the times spanning the observing window. JPL's HORIZONS system¹¹ is designed for this. To return inputs for use with equation (3.2) and J2000 target coordinates: select “Vector Table”, the SSB as the coordinate origin, and “Earth mean equator

¹¹At <http://ssd.jpl.nasa.gov/?horizons>.

and equinox of reference epoch” for the reference plane. This will return the Cartesian coordinates of the observing station with respect to the SSB in the J2000 Earth mean equator reference frame at steps as small as 1 minute in CT, which is the same as TDB.

4. If the observing station is on Earth and better than 20 ms timing is required, another ephemeris must be generated (from HORIZONS) for the observer’s position with respect to the non-rotating geocenter, and added to the geocentric positions. We note that the precise conversion from latitude, longitude, and elevation to the Cartesian coordinates with respect to the non-rotating geocenter is not trivial, and requires tables of measured precession and nutation of the Earth.
5. Interpolate the positions of the observing station to each midexposure time JD_{TDB} .
6. Input the interpolated X, Y, and Z positions of the observing station, and the target’s J2000 Earth mean equator coordinates into equation (3.2). Depending on the distance to the target and the precision required, equation (3.7) (and the target’s ephemeris) or equation (3.8) may be required. One must be careful to use sufficiently accurate target coordinates.

7. If greater than 0.1 ms precision is required, apply the Shapiro correction (equation [3.5]). An ephemeris of the Sun is required, which can be generated from HORIZONS.

8. If greater than 1 μ s precision is required, apply the additional Einstein correction for the observing station's position with respect to the geocenter. The geocentric velocity is required, which can also be given by HORIZONS.

3.4.1. OUR CODE

Our IDL code implementing this procedure is available online¹². It requires the JD_{UTC} at midexposure and target coordinates (α, δ) in J2000 as inputs. We outline its procedure here. More explicit details, as well as the calling procedure and dependencies, are commented inside the code.

We compute Δ_C using Craig Markwardt's `TAI-UTC` program¹³ to read the leap second table, and his `TDB2TDT` program to compute the TT-TDB correction, which uses a 791-term Fairhead and Bretagnon analytical approximation to the full numerical integration, with an error of 23 ns (Fairhead & Bretagnon 1990).

Our code will automatically update its leap second table the first time it runs after every January 1 or July 1, but this requires a periodic Internet connection

¹²At <http://astrutils.astronomy.ohio-state.edu/time>.

¹³At <http://cow.physics.wisc.edu/%7Ecraigm/idl/ephem.html>.

and the use of the `wget` program. It will terminate on failure to update, but this protection can be bypassed for those that elect to (or have to) update their table by hand.

By default, we ignore the $\sim 30\mu\text{s}$ TT(BIPM) - TT(TAI) correction, which would require a constant Internet connection, would not apply to data acquired in the previous month, and is likely negligible for most applications. However, our code can optionally correct for it if an up-to-date file is supplied.

To read and interpolate the ephemeris from JPL, we use Craig Markwardt's routines `JPLEPHREAD` and `JPLEPHINTERP` for the Earth, Sun, and other planets. If the observing station is space-borne, the smaller ephemeris used with those programs does not include satellites, so we use an Expect script to automate a telnet session to the HORIZONS system and automatically retrieve the ephemeris, which we quadratically interpolate to the desired times using IDL's `INTERPOL`. The accuracy of this interpolation depends on how quickly the object's position is changing and the step size of the ephemeris. HORIZONS can only return $\sim 90,000$ data points per query, so the smallest step size (1 minute) limits the calculation to a range of 60 days. For the geocenter, a 100 minute step size is sufficient for 60 ns accuracy, but, for example, a 2 minute step size is required for 1 μs accuracy for the *HST* (though it is still limited by its clock accuracy). We have found that a 10 minute step size is adequate for 1 ms timing for most objects and allows a range of nearly 2 yr.

If the observer is on the Earth, and the coordinates (latitude, longitude, and elevation) are given, we correct for the additional delay. If no observer-specific information is given, we assume the observer is at the geocenter, and the result will be biased by ~ 10 ms (Fig. 3.5).

If the target’s ephemeris can be returned by HORIZONS and its unique name is given, we use our Expect script to generate its ephemeris too, and calculate the exact $\Delta_{R_{\odot}}$ (eq. [3.7]). If not, and instead the distance is given, we use the two-term approximation to the spherical wave solution (eq. [3.8]). Otherwise, we use the plane wave approximation (eq. [3.2]).

Lastly, we include the Shapiro correction and the additional Einstein correction due to the position of the observer with respect to the geocenter, either from the surface of the Earth (if given the coordinates), or the spacecraft.

In the geocentric case, our code agrees with BARYCEN¹⁴ to 200 ns (peak to peak) and the authors of BARYCEN report that their code agrees with FXBARY¹⁵ to 1 μ s. The ephemeris we generate for a location on the surface of the Earth agrees with HORIZONS to 20 nano-seconds, and the geocentric BJDs we calculate from HORIZONS ephemeris agree with the BJDs we calculate using Craig Markwardt’s routines within 10 ns.

¹⁴At <http://astro.uni-tuebingen.de/software/idl/aitlib/astro/barycen.pro>.

¹⁵At <http://heasarc.nasa.gov/lheasoft/ftools/fhelp/fxbary.txt>.

The near-exact agreements between these methods is not surprising, and do not necessarily indicate that they are accurate to better than $1 \mu\text{s}$. Our code was inspired by BARYCEN and both rely on Craig Markwardt’s routines (the difference comes from the fact that we index the JPL ephemeris with JD_{TDB} instead of JD_{TT}), and all methods use JPL’s DE405 ephemeris.

The primary advantage of our code is that it includes the JD_{UTC} to JD_{TT} correction (but can optionally ignore it). The choice of starting with JD_{UTC} is a departure from what is typically done with such time stamp calculators, but we feel this is a far more robust starting point. The current confusion has shown that many assume JD_{UTC} as the starting point, which is likely due to a lack of explicitness in the programs and/or unfamiliarity with various time standards. Our hope is that people are unlikely to make the opposite mistake (assume the input should be JD_{TT} instead of JD_{UTC}) since our code is very explicit and calculating the JD_{UTC} is almost always a trivial calculation from the DATE-OBS FITS header keyword.

Additionally, our code can easily correct for the observer’s position on the Earth or from a spacecraft, and can include the spherical wave correction.

In order to schedule observations, even ± 10 minute precision is generally good enough, and one can approximate $\text{BJD}_{\text{TDB}} \approx \text{JD}_{\text{UTC}}$; for more demanding observing schedules, we provide software to iteratively calculate the reverse correction.

Along with the IDL source code, we provide a web-based interface to our codes¹⁶, though not every feature is enabled. Specifically, it is limited to 1 ms precision, can only do one target at a time, only does the plane wave approximation, and is limited to 10,000 JDs at a time. Those with applications for which these features are too limited should download our source code and run it locally.

3.5. CONCLUSION

Timing of transient events is a powerful tool for characterizing many astronomical phenomena. In the field of exoplanets in particular, the search for variations in the times of primary transits and secondary eclipses, or transit timing variations, is one of the most promising new techniques for studying planetary systems.

The accuracy with which transit times, and indeed any transient phenomenon, can be measured is limited not only by the data themselves, but by the time stamp to which the transit time is referenced. As the quality of transit timing data crosses the threshold of 1 minute precision, the precise time standard and reference frame in which event times are quoted becomes important. Achieving uniform and accurate time stamps with accuracies of better than 1 minute that can reliably be compared to one another requires extraordinary care in both our techniques and our terminology.

¹⁶At <http://astrutils.astronomy.ohio-state.edu/time/utc2bjd.html>.

We have found that the time standards adopted by various groups that measure transit times can differ by as much as a minute, and are typically left unspecified. As these ambiguities can be significant compared to the timing precisions that are quoted, they may therefore lead to spurious detections of transit timing variations or biased eccentricity measurements.

Here we have summarized the effects one must consider in order to achieve timing precision of $1 \mu\text{s}$. We argue that the BJD_{TDB} is nearly the ideal time stamp, being as reliable as any time stamp can be without being unique to each target system. On the other hand, BJD_{UTC} and the HJD in any form should be avoided whenever possible. Most importantly, we emphasize that the time standard should always be explicitly stated. Any time stamp that is quoted without a time standard should be assumed to be uncertain to at least 1 minute. Unless the time standards used in programs or algorithms have been independently confirmed, one should avoid using ones that do not precisely specify the input and output time standard.

In addition, the arrival time at the observing site along with its time standard (e.g., JD_{UTC}) should also be specified. This will remove any ambiguity in the time stamp, allow others to apply improved corrections should more precise ephemerides become available in the future, and allow others to check that original conversion was done accurately enough for their purpose.

Finally, we have written an IDL program for general use that facilitates the use of BJD_{TDB} to an accuracy of $1 \mu\text{s}$, provided that the inputs are sufficiently precise, and we provide a web-based interface to its most useful features.

3.6. CURRENT IMPLEMENTATIONS

We looked in detail at several readily available tools for the BJD/HJD calculation, and have been in contact with many people in the exoplanet community and the help desks for several major space telescopes. We summarize our findings here to demonstrate how easily errors of up to 1 minute can be introduced and to stress the importance of specifying the time precisely. We caution the reader not to trust our general findings for specific cases, but always to confirm what has been done in each case where 1 minute timing accuracy is required but the time standard has not been specified explicitly.

3.6.1. SOFTWARE

Ground-based observers have typically used one of the following methods to calculate the BJD. However, most FITS image headers give the DATE-OBS and TIME-OBS keywords in UTC. We have found that most people, when starting with JD_{UTC} end up quoting HJD'_{UTC} or BJD'_{UTC} .

HORIZONS

JPL’s HORIZONS ephemeris calculator, which is used by many to calculate the BJDs from space telescopes, and can be used to calculate ground-based BJDs, returns the time in $JD_{CT} = T_{\text{eph}} = JD_{\text{TDB}}$ when the ephemeris type is “Vector Table”. Any conversion that uses a HORIZONS ephemeris in JD_{TDB} but indexes it with JD_{UTC} , as had been done by several people we spoke with, will calculate BJD'_{UTC} , which can be offset from the true BJD_{UTC} by up to 13.4 ms (as shown in Fig. 3.6), and offset from the uniform BJD_{TDB} by more than 1 minute (as shown in Fig. 3.3).

IRAF

IRAF’s `setjd` calculates the HJD, but calls for UT, which is likely to be interpreted as JD_{UTC} . In this case, it would calculate the drifting quantity HJD'_{UTC} . If TT were used instead, it would calculate the HJD_{TT} , accurate to ~ 8 s.

IDL

The IDL routines `HELIO_JD` (for HJD)¹⁷, from the IDL Astronomy Library curated by Wayne Landsman, and `BARYCEN` (for BJD), from the Institut für Astronomie und Astrophysik IDL Library, maintained by Jörn Wilms, both call for the GJD, which, we remind the reader, can be specified in any time standard.

¹⁷At http://idlastro.gsfc.nasa.gov/ftp/pro/astro/helio_jd.pro.

Often, this is interpreted as JD_{UTC} , in which case they would calculate HJD'_{UTC} or BJD'_{UTC} , respectively. If TT were used, they would calculate the HJD_{TT} , accurate to ~ 8 s, or BJD_{TDB} , accurate to the geocentric correction (21.3 ms).

AXBARY

NASA's High Energy Astrophysics Science Archive Research Center (HEASARC) created the tools `FXBARY` and later the improved version `FAXBARY`, both of which call `AXBARY` to calculate the BJD. Their documentation is precise and correct, but quite long and may be difficult for the uninitiated to follow. Therefore, it would not be surprising for users of `AXBARY` to input either UTC or TT, in which case they could generate either the BJD'_{UTC} to the accuracy of the leap seconds or the BJD_{TDB} to the accuracy of the geocentric correction (21.3 ms).

ONLINE TOOLS

Currently, common Google results turn up various applets, spreadsheets, programs, or algorithms to calculate the HJD that explicitly call for JD_{UTC} or JD_{UT} as an input. Unless explicitly mentioned otherwise, it is usually safe to assume the time standard used as input will be the time standard used throughout their calculation. Thus, these algorithms and applets will very likely calculate HJD'_{UTC} . However, they are perfectly capable of calculating HJD_{TT} if given JD_{TT} as an input.

3.6.2. SPACE TELESCOPES

EPOXI

EPOXI has the mid-exposure time BJD_{TDB} in the header for the intended pointing under the FITS header keyword `KPKSSBJT`. This can be used directly, as long as one is careful about the intended target, so it would be very surprising if a BJD from EPOXI was not BJD_{TDB} .

We recalculated the BJD_{TDB} using the HORIZONS ephemeris as described in §3.4 and an example FITS header given to us by the EPOXI help desk. In their example FITS header, they pointed at the Moon, which is not infinitely far away, so we must use equation 3.7 for $\Delta_{R_{\odot}}$. With this method, we agree with the `KPKSSBJT` header value to ~ 1 ms, the limit of the precision of the keyword.

We also redid the BJD_{TDB} calculation of HAT-P-4b as described in the *Report on the Calibration of EPOXI spacecraft timing and reduction to Barycentric Julian Date* of August 2009 by Hewagama et al.. We find agreement with the quoted `KPKSSBJT` FITS header keyword to 47 ms. While this is much better than the 0.41 s difference calculated by Hewagama et al. (a difference they attribute to “cumulative rounding limits” in their method), we believe our method to be far more precise. However, we could not obtain access to the original headers and were unable to determine the source of the discrepancy. Given the very good agreement with the

calculation of the Moon above, our best guess is that the target coordinates used by the EPOXI pipeline differed from the published values for HAT-P-4b. The 47 ms difference could be explained by a $\sim 27''$ discrepancy in R.A., a $\sim 45''$ discrepancy in declination, or some combination thereof.

CHANDRA X-RAY OBSERVATORY

Chandra stores their DATE-OBS keyword in TT. Their more precise TSTART and TSTOP keywords are expressed in seconds after 1998 Jan 1, 00:00:00 TT. This departs from what is typically done, which may lead to confusion, but it makes the conversion to a uniform time stamp much more straightforward and less likely to drift by the leap seconds. They provide extensive directions online to calculate the BJD_{TDB} using AXBARY, so it is likely that anyone using Chandra who quotes a BJD is using BJD_{TDB} .

HUBBLE SPACE TELESCOPE

The FITS headers of *HST* state that their DATE-OBS and TIME-OBS keywords are UT. We contacted the *HST* help desk for clarification, since UT is ambiguous. The *HST* help desk response stated that their clock reports UTC accurate to ~ 10 ms, but "due to variabilities and quantization in the particular science instruments' operations, the actual time light begins falling on the detector is not known to better than about ~ 1 second, $\pm 50\%$ (rough estimate)." It is

possible for the HST engineering team to calibrate these variations, but they have limited resources and have no plans to do so. It is thought that this error is some combination of random and systematic errors, but the precise breakdown is unknown.

This potential ~ 1 s systematic error may have important implications for the reliability of the transit times quoted with HST observations, most importantly, the 3 s error of the transit time of HD189733b (Pont et al. 2007).

HST does not calculate the HJD or BJD at any point, leaving the calculation up to each individual observer. Our experience with ground based observers suggests that most people will end up quoting an HJD'_{UTC} or BJD'_{UTC} .

KEPLER

The Kepler Data Release Notes 2 describe how to calculate the BJD from UTC, but do not include the correction to TDB. They mention the HORIZONS ephemeris, but neglect to mention that its output time is in CT, not UTC; thus it appears they calculate BJD'_{UTC} , though we were unable to confirm this. In addition, the time stamp uncertainty may be much larger than typical, so it is worth quoting from the Kepler Data Release Notes 5 (released 2010 June 4): The advice of the DAWG [Data Analysis Working Group] is not to consider as scientifically significant relative timing variations less than the read time (0.5 s) or absolute timing accuracy

better than one frame time (6.5 s) until such time as the stability and accuracy of time stamps can be documented to near the theoretical limit.

SPITZER SPACE TELESCOPE

The *Spitzer* pipeline calculates the HJD_{UTC} for the intended pointing (presumably the target) at the end of the exposure, subtracts the full exposure time, and records the result in the header as HJD. Depending on the exposure time, this will produce roughly a 10 ms effect similar to that shown in Figure 3.6, and depending on how close the intended pointing was to the object of interest, may produce a ~ 0.1 s effect similar to Figure 3.4. However, this effect is negligible compared to both the ~ 8 s accuracy of the HJD (Fig. 3.2) and the number of leap seconds that may have elapsed between observations (Fig. 3.3).

One typically quotes the HJD at the midexposure time. Since *Spitzer* quotes the HJD_{UTC} at the beginning of the exposure, using the unmodified *Spitzer* HJDs would produce a systematic offset of half the exposure time, though experienced observers correct for this.

Also contributing to this confusion, the FITS header keyword `UTCS_OBS` is incorrectly documented. While the documentation states that it is seconds after J2000 ET, it is actually seconds after January 1st, 2000 12:00 UTC + $N - 32$.

Therefore, trusting the documentation as is will unwittingly lead to a difference of $N + 32.184$ s.

Most of the people we have asked opted to calculate their own BJD using the HORIZONS ephemeris. However, they have typically quoted the BJD'_{UTC} .

We would like to thank Craig Markwardt for his fundamental routines that make ours possible, the help desks at the various space telescopes and at IERS for answering our questions, the anonymous referee, Steve Allen, Richard Pogge, Joseph Harrington, Roberto Assef, Andrew Becker, Mercedes Lopez-Morales, Christopher Campo, Drake Deming, Ryan Hardy, Heather Knutson, Eric Agol, and Joshua Winn for useful discussions, and Wayne Landsman and Jörn Wilms for managing the IDL astronomy libraries.

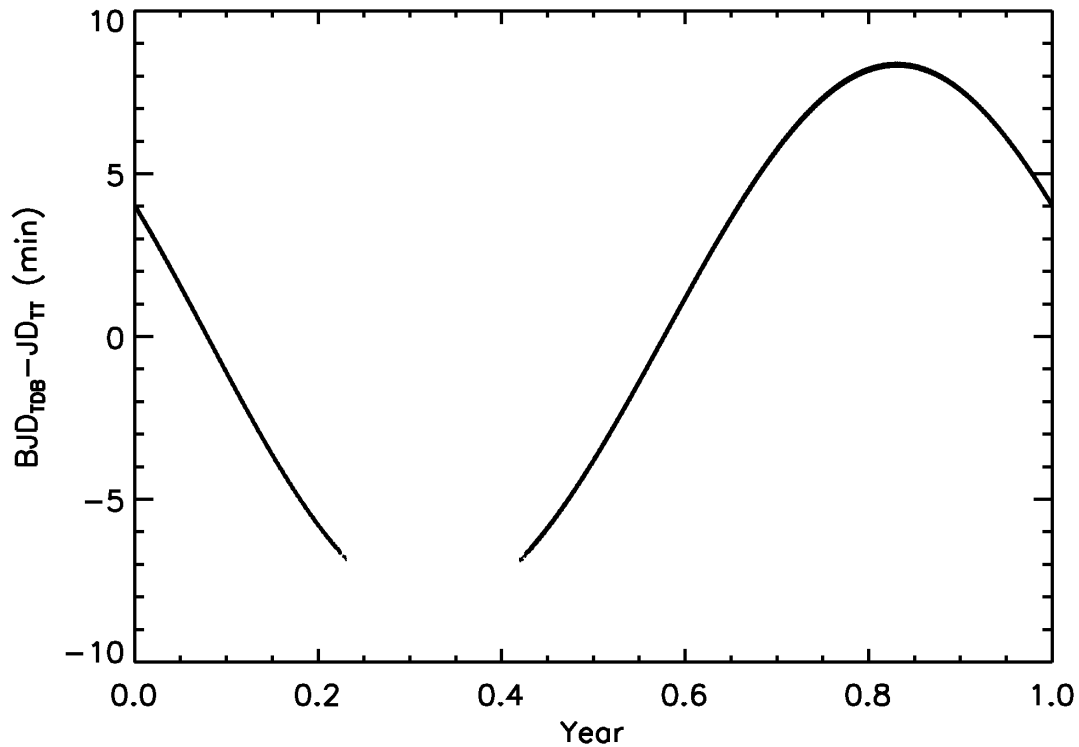


Fig. 3.1.— Difference between BJD_{TDB} and the uncorrected JD_{TT} (see text for definitions) over the course of a year. We plot the correction for a maximally-affected object on the ecliptic for an observer at latitude of zero degrees. We exclude all points where the object has an airmass greater than three and the Sun is higher than -12° in order to highlight observing biases.

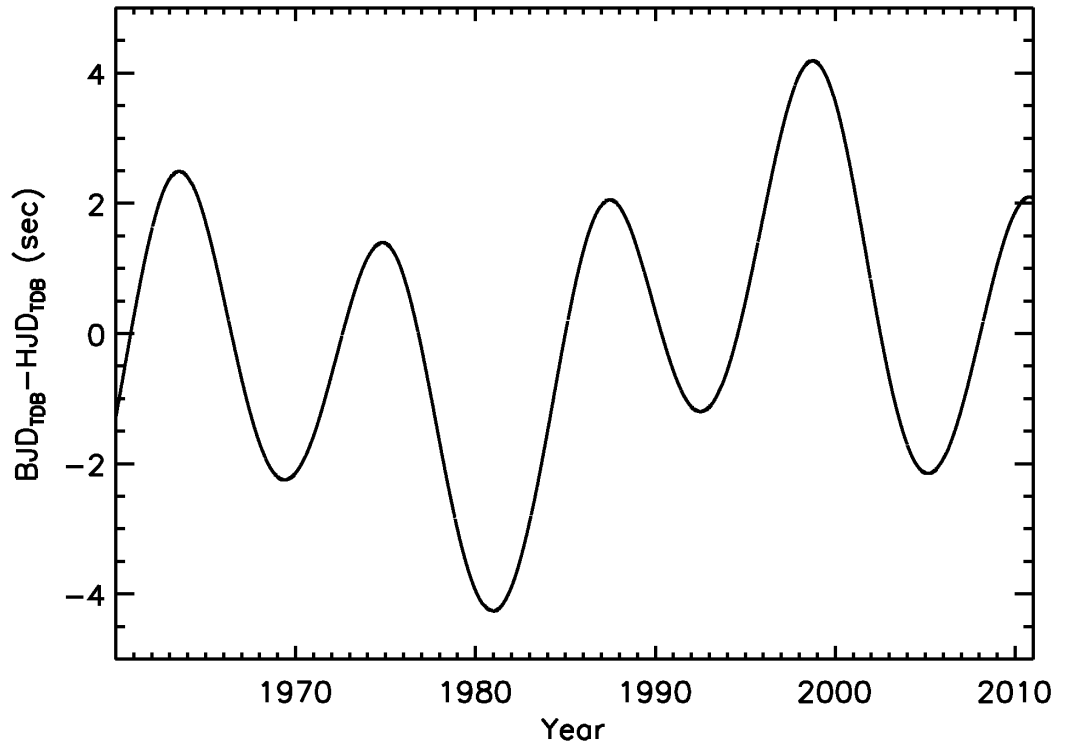


Fig. 3.2.— Difference between the BJD_{TDB} and the HJD_{TDB} for a maximally-affected object on the ecliptic. The primary periodicity is due to Jupiter and the secondary periodicity is due to Saturn.

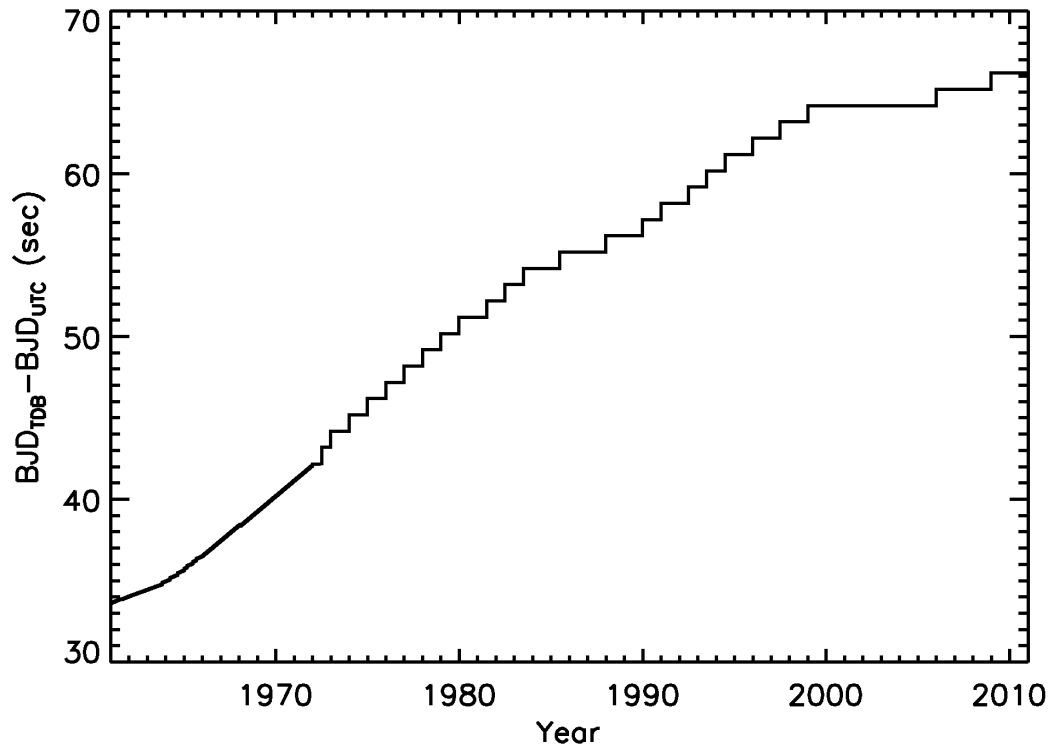


Fig. 3.3.— Difference between the uniform BJD_{TDB} and the BJD_{UTC} . It shows the discontinuities and slow drift in BJD_{UTC} due to the addition of leap seconds. Without correcting for these, relative timing between two reported values of BJD_{UTC} can only be trusted over short time scales.

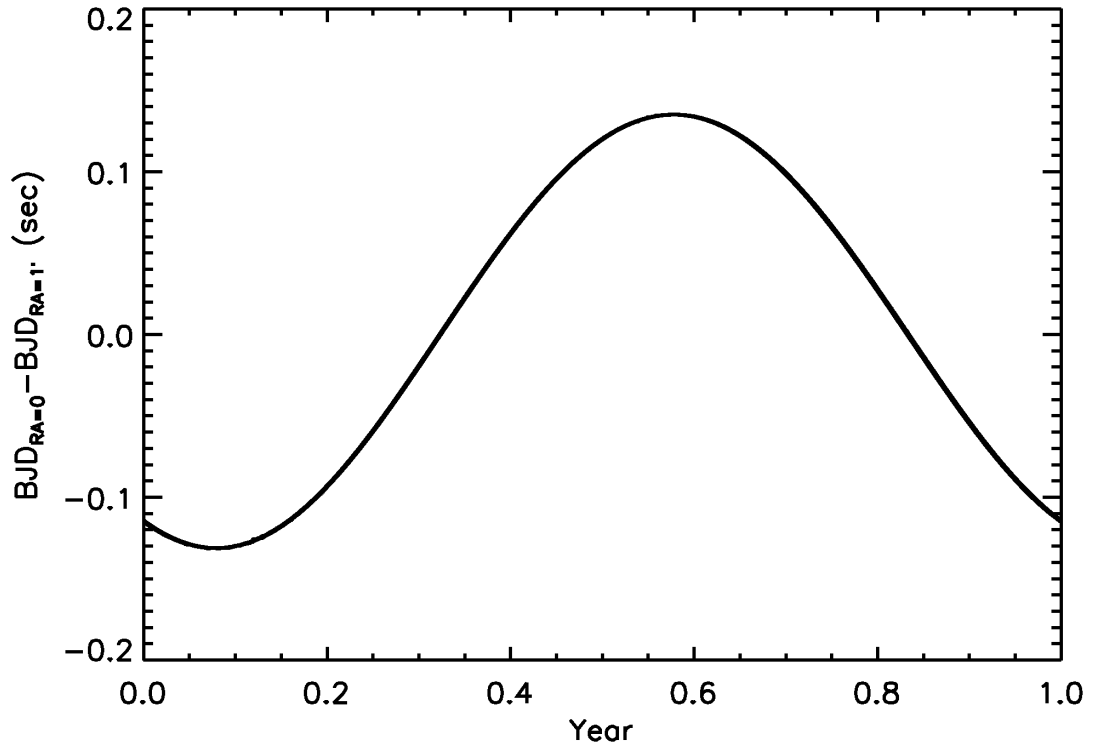


Fig. 3.4.— Difference between the BJD_{TDB} calculated for an object at $\text{R.A.} = 0^{\text{h}}$ and an object at $\text{R.A.} = 0^{\text{h}} 0^{\text{m}} 4^{\text{s}}$ observed at the same time. This difference is as large as 200 s for a 10° offset.

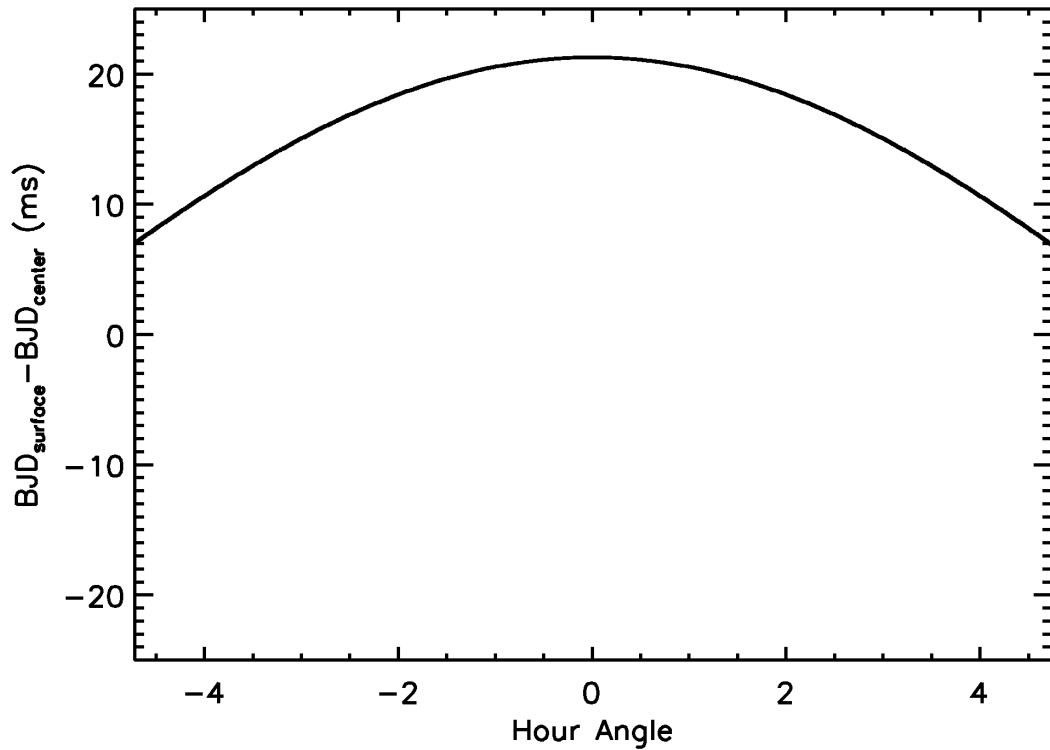


Fig. 3.5.— Difference between the BJD_{TDB} calculated at the geocenter and at the precise location of the observer on the surface of the Earth. While geometrically, this effect will oscillate between ± 21.3 ms with a period of 1 sidereal day, we exclude points when the Sun is above -12° and object is at $z > 3$, which introduces a large observational bias.

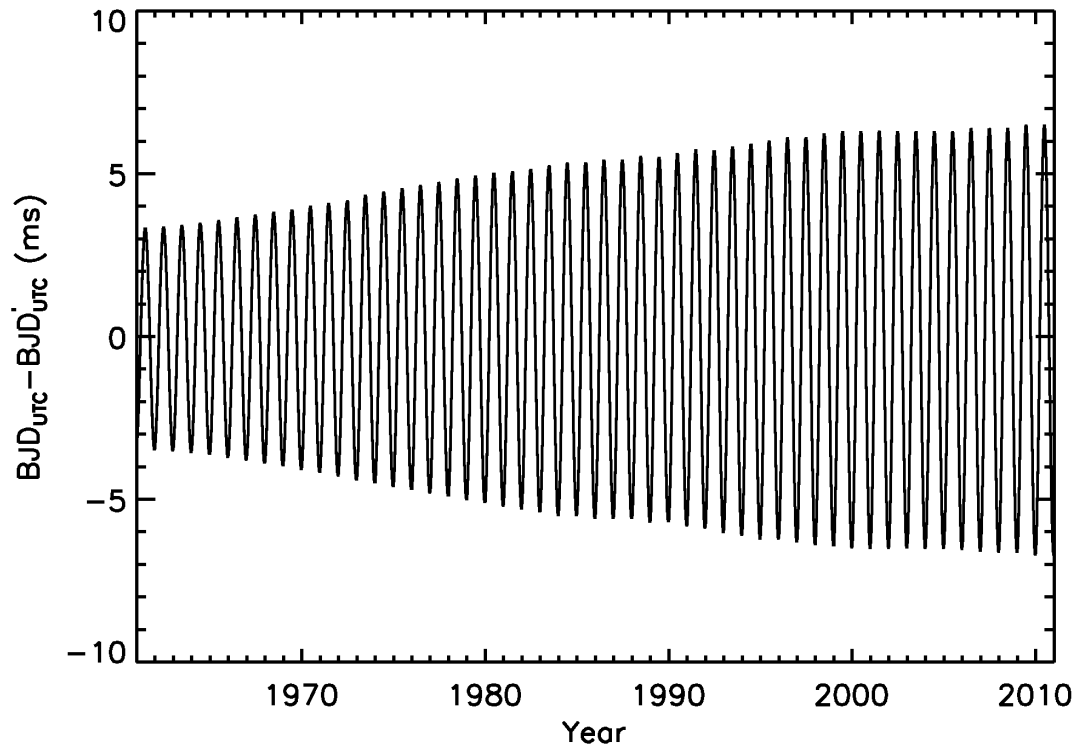


Fig. 3.6.— Difference between the correctly calculated BJD_{UTC} and the commonly calculated BJD'_{UTC} using the positions of the Earth, planets, and Sun delayed by $32.184 + N$ s. It shows that the difference can safely be ignored for ~ 15 ms precision, and therefore the approximate BJD_{TT} or BJD_{TDB} can be recovered from currently published BJD_{UTC} simply by adding $32.184 + N$ s.

CHAPTER 4

EXOFAST

4.1. INTRODUCTION

The goal of this chapter is to provide a fast, easy-to-use code that utilizes the MCMC technique for robust analysis of photometric and radial velocity data for planets. Providing a general code that can fit any conceivable planetary phenomenon without modification is not practical. Rather, we wrote a modular, easily-extensible framework with a relatively straight-forward but powerful example implementation for exoplanets that fits a single-planet system which has either or both RV and primary transit data. This framework and example implementation can be easily adaptable to add additional effects as the data are able to constrain them (e.g., TTVs, secondary eclipses, etc), impose different priors, or even analyze completely different problems (e.g., Supernovae, Cepheids, etc.).

While IDL is a proprietary language that is generally slower than lower-level languages like C and Fortran, we chose to use this language because of the large library of existing code, the ease of development, and the fact that well-written

IDL is comparable in speed to higher level languages for most (i.e., non-serial) applications. While an MCMC code is necessarily serial (i.e., one cannot calculate an arbitrary step in the Markov Chain without first calculating the step before it), the vast majority of the time spent is during each model evaluation, which has been carefully vectorized whenever possible.

We describe our specific procedure to fit RV data (§4.3), including further exposition on the nature of eccentricity bias and a novel method to reduce it (§4.3.2). We discuss our procedure to fit transit data in §4.4, and the combined RV and transit data set simultaneously in §4.5. The outputs of our code are shown in §4.6, validation in §4.7, and our online interfaces to the most useful codes in §4.9. Along with these online interfaces, all of the source code described here is available online¹.

Appendix A demonstrates that the linear and/or quadratic limb darkening coefficients can be fit analytically, which can reduce the dimensionality of a non-linear solver, thereby drastically increasing its speed. Specific improvements to the Mandel & Agol (2002) code to calculate the quadratically limb-darkened flux during transit, including over a factor of 100 improvement in speed are described in appendix B, which cuts the execution time of typical combined RV and transit fit from 30 hours to 2.5 hours. Appendix C warns against the use of the poor-quality random number generator built into IDL and provides an alternative at a moderate increase in the

¹<http://astrutils.astronomy.ohio-state.edu/exofast/>

overall execution time. Lastly, appendix D discusses the program's performance and identifies areas of future improvement.

4.2. OVERVIEW OF THE PROBLEM

4.2.1. FINDING THE BEST FIT

Given a data set, D , with errors, σ , we would like to generate a model M from a set of model parameters that describe the data. If we assume the errors are Gaussian, we can use the standard χ^2 goodness of fit statistic,

$$\chi^2 = \sum_{i=1}^n \left(\frac{D_i - M_i}{\sigma_i} \right)^2 \quad (4.1)$$

and the subscript i corresponds to each data point, to determine the likelihood of a model, given the data is

$$P(D|M) = L = e^{-\chi^2/2}. \quad (4.2)$$

Therefore, finding the maximum likelihood is equivalent to finding the model with the lowest χ^2 . When the model is linear (i.e., can be written as a simple linear combination of known quantities with unknown coefficients, or parameters), the χ^2

can be minimized analytically and exactly to find each of the coefficients (see Gould 2003).

However, when the model is non-linear, such as for transits and radial velocities, we must solve for the parameters numerically. Unfortunately, there are no generic algorithms to minimize the χ^2 for a global parameter space – often, various tricks are required that are specific to the particular problem at hand. We will discuss the tricks specific to exoplanets in §4.3 and §4.4, but otherwise we assume we have a reasonably well-behaved region of parameter space with a smoothly varying χ^2 surface. In such a circumstance, there are many routines that can robustly find a local minimum. Amoeba is a popular non-linear solver that uses the downhill simplex method to find local minima (Nelder & Mead 1965). We simply give it a starting point and stepping scale (which is approximately the range of parameter space it will consider), and amoeba will crawl through parameter space (like an amoeba), to find the minimum, using the χ^2 at each step to inform its next step. This routine is very robust at finding a local minimum.

IDL comes with its own built-in `AMOEB` routine, but we discovered a bug that truncates the stepping scale to floating point precision, regardless of the data type initially given. This is detrimental when fitting parameters that require double precision (e.g., JD), since the model will simply oscillate about the minimum. We provide a debugged version of this code, which now forces all scales to be double precision, as a new code in this suite, `EXOFAST_AMOEB`.

Another popular non-linear algorithm to find the local minimum is Levenberg-Marquardt (LM) (Levenberg 1944; Marquardt 1963), which uses numerical derivatives to predict the minimum more precisely after each evaluation, and therefore requires fewer evaluations of the χ^2 statistic (which can be very costly for complicated models). The downside is that, if the χ^2 surface is not smooth, the numerical derivatives may be a poor predictor of the minimum and the fit will not be as robust.

Craig Marquardt wrote an extremely versatile and widely-used IDL implementation of the LM algorithm, called `MPFIT`². As expected, we found our debugged version of `AMOEBA` to be more robust than `MPFIT`, routinely finding as good or lower values of χ^2 , but about 10 times slower.

Once we have the best fit, we check for self-consistency by ensuring that the χ^2 per degree of freedom, χ_ν^2 , is 1. Many times, when fitting a data set, the $\chi_\nu^2 \neq 1$, which could be for a variety of reasons. It may be that the model does not properly describe the data (e.g., there is another planet that is not modeled), but it is often difficult to know for sure without having a good understanding of the instrumental setup.

Assuming the model properly describes the data, the next thing to do is eliminate outliers. If that does not help, the last option is to scale the errors. There

²<http://www.physics.wisc.edu/~craigm/idl/down/mpfit.pro>

are two ways in which one can scale the errors – to add a constant “red noise” term in quadrature with each of the other errors, or to scale each of the errors by a constant multiplicative factor, $\sqrt{\chi_\nu^2}$. Which to use depends on the nature of unaccounted error, and unfortunately, if we knew precisely what it was, we probably would have included it in our error estimate to begin with. As a general rule, if there is an unaccounted for systematic (e.g., stellar jitter), the error should be added in quadrature. The practical difference is that adding errors in quadrature will tend to even out all of the errors, whereas multiplying will preserve the relative errors.

Fortunately, when we were fitting the brown dwarf in (Lee et al. 2010), we found no statistical difference between the two methods, so we always choose to scale by $\sqrt{\chi_\nu^2}$ because it is easiest.

If we consider data sets from different sources, they are likely to have different systematics, so it is a good idea to fit each desperate data set independently and scale the errors to ensure they are self-consistent before attempting to combine them. If there are large inconsistencies between the data sets, it is indicative of serious problems with either the model (neglected effects) or data (systematic errors) and a simultaneous fit should not be trusted. If they are consistent, we quickly find the best fit to the combined data set using a local minimum solver.

4.2.2. FINDING THE UNCERTAINTIES: MCMC

Since there is zero phase space at precisely the best fit, it tells us nothing of the uncertainties. Recently, Bayesian inference using a Markov Chain Monte Carlo (MCMC) technique has become a widely-favored method to characterize the uncertainties because (among many other reasons) it naturally accounts for non-Gaussian errors and covariances between parameters, is based solely on the data acquired and the priors imposed, and it is trivial to derive values, uncertainties, and covariances for quantities based on the fitted parameters.

Bayes' theorem is a robust way of incorporating new data to revise initial beliefs. It states that the probability of our model given the data, $P(M|D)$, is simply

$$P(M|D) = \frac{P(D|M)P(M)}{P(D)}, \tag{4.3}$$

where $P(D)$ is a normalization constant that can be ignored and $P(M)$ is our model priors.

This provides a natural framework in which to characterize the probability that our model is correct, called Markov Chain Monte Carlo (MCMC). We simply

calculate the χ^2 given a set of parameters, randomly step to a different set of parameters, and calculate the χ^2 again. We calculate the likelihood ratio,

$$L_1/L_2 = \exp^{(\chi^2(M_1) - \chi^2(M_2))/2}, \quad (4.4)$$

and draw a random number uniformly distributed between 0 and 1. If the random number is greater than the likelihood ratio, the model is rejected and we do not step there. If the random number is less than the likelihood ratio, we accept the new model. Note, when $\Delta\chi^2 < 0$ (i.e., the new model is a better fit), L_1/L_2 is always greater than 1, and so the step will always be accepted.

We repeat this process, stepping to a new region of parameter space until we have a smooth distribution of values for each parameter. The resultant distribution of steps is proportional to the probability of each parameter, naturally resulting in a robust best-fit value (e.g., the median value) and the error distribution (e.g., the 68% confidence interval). We do not consider the absolute normalization ($P(D)$), which one would need to do in order to consider the likelihood of models with a different parameterization.

One of the most problematic aspects of using MCMC is determining an appropriate stepping scale. Using the proper length scale is key to speedy convergence. If the scale is too large, very few trial links will be chosen, and many

models will be calculated unnecessarily. If the scale is too small, too many links will be accepted and the resultant chain will not be “well-mixed.”

The ideal step mimics the posteriori probability distribution precisely (Gelman et al. 2003). This means, in the case of Gaussian, uncorrelated errors, the ideal step size for each parameter is the 1-sigma uncertainty. Of course, that is exactly what we are trying to calculate. Fortunately, this only influences the convergence time, not our conclusions, so an approximate guess is perfectly adequate. Starting at the best-fit value, we use our program `EXOFAST_GETMCMCSCALE` to calculate the scale of each parameter for the Markov Chain. It starts with a very small positive step and calculates the $\Delta\chi^2$. It then doubles the step size until the $\Delta\chi^2 > 1$, at which point it performs a binary search with the values that bracket $\Delta\chi^2 = 1$. It repeats the process in the negative direction, and uses the average of the two as the step size.

Unfortunately, poorly-behaved χ^2 surfaces are common and present various special cases that must be dealt with in order to be robust. A successful binary search requires that the $\Delta\chi^2$ be monotonically increasing away from the best fit, which is not generally true. If the maximum and minimum bounds of the binary search are equal without finding $\Delta\chi^2 = 1$, we use the step size that was closest along the way. If a parameter is explicitly bounded (i.e., $\chi^2 = \infty$), and $\Delta\chi^2 < 1$ until that boundary, we use the step size in the opposite direction as our step size. If a parameter is periodic and poorly constrained, it is possible that no value corresponds to $\Delta\chi^2 = 1$. In these cases, we set the step size to half the period.

When the parameters have Gaussian, uncorrelated errors, this algorithm yields roughly the optimal acceptance rate of our Markov Chains of $\sim 44\%$ for a single parameter asymptoting to $\sim 20\%$ for more than 5 parameters (Gelman et al. 2003). Even when the parameters are highly correlated, we find the acceptance rate is not far from optimal.

Then, roughly following the guidelines set forth by Ford (2006), we begin a series of ten independent Markov Chains, each with starting values randomly 5 sigma from their best-fit values. We take Gaussian random steps in all non-linear parameters simultaneously, until the ten chains have converged.

We consider the chains to be converged when both the number of independent draws, T_z as defined by Ford (2006) is greater than 1000 and the Gelman-Rubin statistic, \hat{R}_ν is less than 1.01 for all parameters. The independent draws and Gelman-Rubin statistic are calculated in `EXOFAST_GELMANRUBIN`. This test must be passed 6 consecutive times – after passing these tests the first time, we take 1, 2, 3, 4, and 5 percent more steps and check again. If any of these subsequent tests are failed, we restart the convergence test. When all tests have been passed, we consider the chain well-mixed and stop.

The resultant parameter distributions are proportional to the probability of each parameter. We quote the median of the distribution as the best value, and the 68% confidence interval as the error.

4.3. RADIAL VELOCITY MODEL

As a planet orbits its star, it gravitationally interacts and pulls it. This motion causes a Doppler shift of the observed spectrum, which is given by the following equation:

$$RV(t) = K [\cos(\theta(t) + \omega_*) + e \cos \omega_*] + \gamma + \frac{dv}{dt}(t - t_0) \quad (4.5)$$

where K is the radial velocity semi-amplitude, $\theta(t)$ is the true anomaly as a function of time, ω_* is the argument of periastron of the star’s orbit measured from the ascending node to periastron³, e is the orbital eccentricity, γ is the systemic velocity (or often just an arbitrary instrumental offset), dv/dt is a systematic slope either due to an additional planet in the system or systematics in the data, and t_0 is an arbitrary zero point for the slope (we use the error-weighted mean of input times).

³Throughout this paper, when we reference the argument of periastron, we refer to the argument of periastron of the star’s orbit. Since we measure the radial velocity of the star’s orbit, this is the often unspoken – but not completely universal – standard in the exoplanet literature. The argument of periastron for the planet, ω_P differs from ω_* by π . As this definition can be somewhat counter-intuitive, we keep the subscript “*” to make this explicit.

The true anomaly, which is the angle between periastron and the planet, measured from the barycenter of the system, is

$$\theta(t) = 2 \arctan \left[\sqrt{\frac{1+e}{1-e}} \tan \left(\frac{E(t)}{2} \right) \right] \quad (4.6)$$

where $E(t)$ is the eccentric anomaly, as given here:

$$M(t) = E(t) - e \sin(E(t)) \quad (4.7)$$

as a function of the Mean anomaly, $M(t)$. Unfortunately, this is a transcendental equation for which no analytic solution for $E(t)$ exists. It must be solved numerically from a given $M(t)$. We use a slightly improved version of a code written by Marc Buie and Joern Wilms, which uses the method by Mikkola (1987) and then refines it using a Newton-Raphson method for high eccentricities.

The mean anomaly simply describes a uniformly flowing time and can be computed from the period of the orbit, P , and the time of periastron passage (of the planet), T_P :

$$M(t) = \frac{2\pi}{P}(t - T_P). \quad (4.8)$$

In many instances, the reverse calculation is also required. That is, we would like to know the time, t , that the planet will be at a given true anomaly. For instance, there are many special times in an orbit that one may be interested in knowing, like the time of periastron, T_P , the time of primary transit center, T_C ⁴, the time of secondary eclipse center, T_S , the time when the star is at its ascending node T_A (when the RV is at a maximum), the time when the star is at its descending node, T_D (when the RV is at a minimum), or the time of the L_4 and L_5 star-planet Lagrange point passage, T_{L4} and T_{L5} , respectively, which would tell us when to look for transiting Trojan planets (Ford & Gaudi 2006). Each of these times correspond to a particular value of the true anomaly, given here:

$$\begin{aligned}
\theta(T_P) &= 0 \\
\theta(T_C) &= \pi/2 - \omega_* \\
\theta(T_S) &= 3\pi/2 - \omega_* \\
\theta(T_A) &= -\omega_* \\
\theta(T_D) &= \pi - \omega_* \\
\theta(T_{L4}) &= 11\pi/6 - \omega_* \\
\theta(T_{L5}) &= 7\pi/6 - \omega_*
\end{aligned} \tag{4.9}$$

⁴The use of this terminology does not mean that the object transits – this is simply the predicted transit center if the inclination were favorable. This may more appropriately be called the time of (inferior) conjunction.

Fortunately, this is much easier. We simply invert equation 4.6,

$$E(t) = 2 \arctan \left[\sqrt{\frac{1-e}{1+e}} \tan \left(\frac{\theta(t)}{2} \right) \right], \quad (4.10)$$

plug $E(t)$ into equation 4.7, and solve equation 4.8 for t . Our routine for the reverse correction, `EXOFAST_GETPHASE` has keywords that can calculate the phases of each of the true anomalies described in equation 4.9 or for an arbitrary true anomaly.

While the ω_* is completely degenerate for a planet in a circular orbit, we follow the convention that $\omega_* = \pi/2$ when $e = 0$. This has the virtue that $T_P = T_C$.

4.3.1. RV PARAMETERIZATION

First, we start with an initial guess of each of the input parameters and generate a model described above. Then, we calculate the χ^2 and follow the general outline described in §4.2.2.

The choice of parameterization is extremely important because we implicitly impose priors that are uniform in each of these parameters. If these priors are not physically motivated, they may introduce biases in what we measure. Second, the choice of parameterization is important because highly-correlated parameter sets converge very slowly. The particular parameterization we use differs from the method outlined in Ford (2005) a few key ways. First, we agree on using $\log P$,

$\log K$, and γ , but the notion of imposing strict bounds outside of what we think is possible on each of these parameters to make them “normalizable” or “proper” priors (i.e., the integral over the allowed states is finite) is misleading.

On page 62, Gelman et al. (2003) states that we can obtain a proper result from an improper prior as long as the posterior distribution (i.e., our parameter distribution) is normalizable, but he warns that such “distributions must be interpreted with care – one must always check that the posterior distribution has a finite integral and a sensible form.” Artificially imposing boundaries outside of what can be reasonably expected does not free us from this responsibility. In either case, if our posterior distribution is unexpected, we must investigate why. Further, by not imposing strict bounds, we simplify the code and ensure that we do not exclude solutions that may actually be allowed. In practice, this should make no difference as long as the bounds chosen were sufficiently large so as not to bias the result.

T_C vs T_P vs M_0

Ford (2005) suggested parameterizing a time in the orbit with a mean anomaly, M_0 , at some arbitrary zero-point, T_0 . However, M_0 is trivially related to T_P ,

$$T_P = T_0 - \frac{M_0 P}{2\pi}. \tag{4.11}$$

So, modulo comparatively minor covariances with P , it is just as covariant with e and ω_* . Imagine, when e is very nearly zero, the periastron is poorly defined, and may swing wildly from point to point in the orbit at each step in the chain. This means that T_P , and therefore M_0 would also want to swing wildly or the model would be out of phase with the data (and the step would be rejected).

Ford (2006) suggested many other alternative parametrizations to aide convergence, including $\omega_* + M_0$ for low-eccentricity orbits, or T_P and $\omega_* + T_0$ for high eccentricity orbits. However, when we consider the shape of the RV curve, there are 4 more obvious (and static) features in the RV curve – T_C , T_S , T_A , and T_D . As ω_* changes, the true anomaly of these points change too (see equation 4.9) and they perfectly cancel. Indeed, this happens for any time that corresponds to a true anomaly of $C - \omega_*$, where C is a constant, and we have discovered empirically that all such times are all about as well-constrained.

One might be tempted to try to define such parameters in the mean anomaly, as Ford (2006) did, thereby removing the dependence on (and prior imposed by) the period of the orbit. However, the mean anomaly at these special times are precisely determined by ω_* through equations 4.10 and 4.8, so stepping in $M(t)$ for these special cases would be a redundant constraint – and we would have no constraint on the particular time that corresponds to a phase on the orbit.

The price for not stepping in the mean anomaly is that each of these times impose a prior that is linear in P (that is, the Jacobian from transforming from M_0 to t is $\partial t/\partial M_0 = P/2\pi$). While this is typically a small effect because the period is much better known than any of these times, Ford (2006) shows that we can easily compensate for it by weighting the stepping probability by P_{i-1}/P_i .

Since all such times are roughly equivalent, we choose to use T_C for its practical uses – T_C is the time we want to look for transits (and is the reason we first considered this family of parameterizations).

T_C could take on any value that differs by integer multiples of the period and it would have no effect on the derived model. However, the derived uncertainty of T_C , and the covariance with P is strongly dependent on the choice of the zero-point (see Figures 4.1 and 4.2). It is usually best-constrained and least covariant closest to the error-weighted mean of the input times, and that is what we use. Further, while this is very unlikely to happen, if the chain jumped from T_C to $T_C \pm P$ and was not adjusted back, it would calculate an equivalent χ^2 and the fit would have difficulty converging (and if it did, its median value would be meaningless).

4.3.2. ECCENTRICITY BIAS

Another departure from the standard RV fit is that we choose use e and ω_* directly, rather than $e \cos \omega_*$ and $e \sin \omega_*$. We discuss this choice in detail in the following section.

PRIOR

In the absence of highly-constraining data, what we assume as our prior can have a significant impact on the measured parameters of the system. Further, when the eccentricity of a planetary orbit is intrinsically zero, no data set is highly-constraining.

While it's unclear what a good, physically-motivated prior for the eccentricity of a planet is, most people have assumed a prior that is uniform in e is best, and for our study, we do the same.

Unfortunately, as noted by (Ford 2006), uniform steps in $e \cos \omega_*$ and $e \sin \omega_*$ introduce a linear prior in e . He suggested that we can correct for this linear prior during an MCMC fit by weighting the stepping probability by the determinant of the Jacobian of the transformation between the parameters in which you step and the parameter you desire to be uniform, or e_{i-1}/e_i , where the subscript i denotes the current link in the chain.

However, we demonstrate that this method does not completely unbias the linear prior in e , particularly at low eccentricities where this bias is most important, and can significantly impact the interpretation of small eccentricities. The amount of uncorrected bias depends on the step size of e in the MCMC chain, Δe .

To find the prior distribution of e and ω_* , we ran a standard Markov Chain, but set the χ^2 to 1 as long as the chain made an allowed step. Whenever a parameter stepped out of the allowed range, we set the χ^2 to infinity. Since the stepping probability is $\exp^{-\Delta x^2/2}$, this will accept any step in bounds and reject any step out of bounds. In this way, we show what prior we have imposed on each parameter.

First, we step in $e \cos \omega_*$ and $e \sin \omega_*$, solve for e and ω_* at each step, and reject the step if $e > 1$. Figure 4.3 shows the normalized prior probability of each value as a result of taking 1 billion Gaussian random steps in this manner. As expected, we clearly see a linear prior in e . We also see a slight attenuation at $e = 1$, since such values near 1 are preferentially approached from below. The step size affects the attenuation depth, but not the relative probability of hitting the boundary at $e = 1$.

Next, we weight the stepping probability by e_{i-1}/e_i , as suggested by Ford (2006), such that our new jump criteria is $\frac{e_{i-1}}{e_i} \exp^{-\Delta x^2/2}$, which will preferentially reject steps to higher eccentricity and is supposed to correct for the linear prior in e . As shown in figures 4.3 and 4.4, it works well when $e \gg \Delta e$, but at $e \lesssim \Delta e$, we see the correction is quite poor. Since Δe ideally corresponds to the 1 sigma

error in e , and the bias is proportional to Δe , we should then expect this bias to be proportional to the ultimate uncertainty stated in e . When the eccentricity is intrinsically zero, this poorly-corrected region is precisely where the chain will spend the majority of its time, so we can expect a large bias to remain.

To correct for this, we try stepping directly in e and ω_* , rejecting steps outside $0 < e < 1$. If ω_* steps outside the range of $-\pi < \omega_* < \pi$, we add or subtract 2π until it falls within range. This is a significant improvement over previous results, but we still see clear evidence for a bias away from $e = 0$ and $e = 1$, since they are preferentially approached from one side.

To avoid this bias, we simply allow the chain to move to a negative eccentricity. While formally undefined, we were motivated to do this by considering the definition of eccentricity, $e = (r_a - r_p)/(r_a + r_p)$, where r_a is the distance to the focus at apoapsis and r_p is the distance to the same focus at periapsis. It is obvious, then, that a negative eccentricity would imply that we have incorrectly labeled our periapsis and apoapsis. So when $e < 0$, we redefine a new eccentricity, $e' = -e$. Since that flips our periapsis and apoapsis, we also need to change the argument of periastron accordingly: $\omega'_* = \omega_* + \pi$.

We see that this method imposes the desired prior that is uniform in e . In the end, however, we cannot allow negative eccentricities in our final distribution because it will create a discrete degeneracy between e and ω_* , and $-e$ and $\omega_* + \pi$.

In a properly sampled Markov Chain, the median value of e would always be 0, and for statistically significant eccentricities, the chain is likely to get stuck at either the negative or positive eccentricity and never sample the other. We avoid this by explicitly changing our values of e and ω_* following the above prescription as soon as e steps to a negative value. In the case of a zero eccentricity, then, we expect to see half a Gaussian distribution, peaked at $e=0$.

The boundary at $e = 1$ still imposes a bias for all methods, though this is unlikely to affect the results of any planet measured to date. Further, planets with $e = 1$ are not stable, and thus are not expected.

We note that the distribution of ω_* is uniform for each method. However, if we were to reject steps outside of $\pm\pi$, rather than rescaling it, we would impose a bias similar to that seen at $e = 1$ seen in all cases. If we were to allow ω_* to vary freely without regard to the periodic boundary condition, the chain would be very unlikely to converge.

LUTZ-KELKER BIAS

Separate from any bias imposed by the prior, there is a subtly different, Lutz-Kelker-type bias (Lutz & Kelker 1973) that also conspires to inflate the eccentricity of intrinsically circular orbits, as pointed out by Laughlin et al. (2005). The premise behind this bias is that, for any positive-definite parameter, there is zero

phase space at exactly zero and therefore any real errors will necessarily conspire to make the parameter appear to have a non-zero (positive) value.

As mentioned in the previous section, the distribution of e for a perfectly circular orbit, should be a half Gaussian distribution peaked at $e = 0$. The median value, which is usually taken as the best value, of such a distribution is 0.67σ . Any errors will tend to broaden this Gaussian and therefore increase the median value. This is the nature of the Lutz-Kelker bias.

Unfortunately, there is nothing that can be done about this bias, except to be aware of it when interpreting small but nominally significant eccentricities.

SIMULATED DATA

We simulated data sets for a hot Jupiter ($\gamma = 500$ m/s, $P = 3.223$ days, $\omega_* = 53^\circ$, $K = 500$ m/s, $dv/dt = 0$, fixed) with 25 evenly spaced data points over a span of 100.123 periods. We injected Gaussian random error with a 1-sigma width of 50 m/s, then fit it as described in §4.3.3 with each of the four eccentricity priors. For each eccentricity ranging from 0 to 0.3 in steps of 0.03, we repeated this 100 times, adding a different Gaussian random noise (still with a 1-sigma width of 50 m/s) each time. The error in the eccentricity was approximately 0.03 (in the $e = 0.3$ case) so our steps in eccentricity roughly correspond to 0 to 10 sigma significant eccentricity.

We calculated the median value for eccentricity of fit and average them together in order to reduce the statistical error and flesh out the bias. This mean of medians is plotted as a function of intrinsic eccentricity in figure 4.5 for each of the four priors. A non-biased result would fall on the dotted line. As expected from the Lutz-Kelker bias, all of the methods over-estimate the lowest significance eccentricities, but the linear prior routinely over-estimates it on average by 2 sigma, and it was not uncommon for the eccentricity to be overestimated by 3 sigma, even with the corrected linear prior. The uniform prior only over-estimates it by 1 sigma on average. Further, the uniform prior is more accurate for smaller eccentricities, allowing us to be more confident when claiming small but significant eccentricities.

Figure 4.6 shows the parameter distributions for e in all 100 fits for each of the four priors in the $e = 0$ case. Again, we see the superiority of the uniform prior in recovering the true eccentricity, and this suggests that the median is not a good summary of such distributions. A far more appropriate quantity to quote for insignificant eccentricities ($\lesssim 3\sigma$) would be the 95% upper limit or the mode. A problem with quoting the mode, however, is that the number of steps required to get a smooth parameter distribution (and therefore a well-defined mode) as seen in figure 4.6 is much larger than the number of steps required for convergence. Therefore, the mode is highly susceptible to noise in the parameter distribution. Regardless, it is always a good idea to inspect the parameter distributions of parameters, and especially for e .

We have shown that with great care, we may be able to trust low-significance eccentricities. However, this bias is prevalent in the literature to varying degrees, so it would be wise not to trust any eccentricity quoted with less than 5σ -significance.

A common reason cited for stepping in $e \cos \omega_*$ and $e \sin \omega_*$ is that the covariances are smaller and therefore the convergence time is much faster than stepping in e and ω_* . To test this, we recorded the time each fit took until the chain was well-mixed, according to the criteria outlined in §4.2.2. The mean execution time as a function of eccentricity for each of the four priors is plotted in figure 4.7. Because we wanted to primarily test the effect on the parameterization, the innermost loop of our simulation cycled through each of the priors. Therefore, to the extent the run times were affected by other activities on the computer, there is no reason to expect one of the priors to have been preferentially affected. The outermost loop cycled through eccentricity, so it is easier to believe that a systematic effect could influence the relative execution times of small to high eccentricities, though the behavior of faster convergence at higher eccentricities follows our expectation. We expected to see that on average, the uniform prior would converge slower, and we wanted to measure by how much. However, the uniform prior actually converged fastest for all values of eccentricity. In the $e \geq 0$ case (the conventional way to step in e) we see that it takes nearly twice as long as stepping in $e \cos \omega_*$ and $e \sin \omega_*$, correcting for the linear prior, explaining the conventional wisdom.

This clearly demonstrates that it is both faster and more accurate to step in e and ω_* directly while allowing $e < 0$. However, there are practical difficulties in stepping in angles such as ω_* . In particular, we must be careful to rescale it whenever it crosses a boundary – otherwise, it would be free to jump to equivalent, widely-separated minima. Further, we may get unlucky and find that our best fit overlaps with the boundary or, when it is poorly constrained, it is possible for a significant amount of power to span the entire range of 2π . Either case could drastically bias the median value, which is required to calculate the convergence criteria and is often used as the best-fit value. In order to account for this, we first center the distribution about the mode, then calculate the median. Lastly, our scheme of finding the appropriate step size would fail if the angle is so poorly-constrained that no value of ω_* produced a $\Delta\chi^2 = 1$. In such cases, we set the step size equal to π . Each of these must be handled as special cases in our code for angular parameters, which can be burdensome. We make this relatively straight-forward by allowing an array of indices that specifies which parameters are angular and will therefore be subject to these special rules.

4.3.3. RADIAL VELOCITY BEST FIT

Finding the global minimum to a radial velocity data set can be extremely complicated because of the vast volume of parameter space. Fortunately, this process can be greatly simplified if the orbit is circular, in which case the $e \cos \omega_*$ term drops

out and T simplifies to the the mean anomaly, $2\pi(t - T_P)/P$. Then, the RV equation can be re-written as

$$RV(t) = A \cos\left(\frac{2\pi t}{P}\right) + B \sin\left(\frac{2\pi t}{P}\right) + \gamma + \frac{dv}{dt}(t - t_0), \quad (4.12)$$

where A and B are arbitrary coefficients that can be solved analytically along with γ and dv/dt (see Gould 2003). This leaves only one non-linear parameter, P , which can be quickly stepped through while the others are solved analytically at each step.

$$K = \sqrt{A^2 + B^2} \quad (4.13)$$

$$T_P = \frac{P}{2\pi}(\arctan(A, B) + \pi/2).$$

where $\arctan(A, B)$ represents $\arctan(B/A)$ ⁵.

This is the basis for the Lomb-Scargle periodogram (Lomb 1976; Scargle 1982). Typically, a scan through period will reveal several peaks in likelihood, due to the planetary orbit and aliases thereof, which gives a good starting point for more rigorous, fully non-linear local minimization routines. We choose the best periods (the maximum likelihood in the LS periodogram) and their corresponding

⁵Mathematically, the arc tangent ranges from $-\pi/2$ to $\pi/2$, but when the sign of the numerator and denominator are known independently, the precise inverse mapping to the full range $-\pi$ to π can be determined. This notation is commonly used in computer programming, and is sometimes called `atan2`.

analytically fit parameters as a starting point to an AMOEBA code which fits a full Keplerian orbit (i.e., including eccentricity).

Sometimes, because of aliases, multiple periods will be similarly good fits, which should be investigated individually. However, often, one particular period will be a clear best, and AMOEBA will give us a robust fit. Given adequate phase coverage over one or more complete orbits, this method can robustly fit most single-planet systems without any special effort. Those with very high eccentricities or poor phase coverage can often be fit, but sometimes need slight adjustments to the default period range or number the number of minima to explore. For example, fitting the data from Winn et al. (2009) for HD 80606b ($e \sim 0.93$) required searching the 100 likeliest peaks in the Lomb-Scargle Periodogram, whereas the default is 5.

4.4. TRANSIT MODEL

A primary transit occurs when a planet passes in front of its star and blocks a portion of its light for a period of time. We monitor the star's brightness during this time by taking repeated exposures, comparing the target star's brightness to an ensemble of comparison stars, and so we expect the errors to be Gaussian. If the star were uniformly bright, the relative flux we would see during transit would be:

$$F = F_0(1 - \lambda^e) \tag{4.14}$$

where analytic equations for λ^e are defined in Mandel & Agol (Eq. 1, 2002) (for all special cases) and are functions of the transit geometry, $R_P/R_* \equiv p$, the radius of the planet in stellar radii, and $r/R_* \equiv z$, the projected distance from the center of the planet to the center of the star, in stellar radii.

4.4.1. LIMB DARKENING

Unfortunately, a star is not uniformly bright – its apparent brightness falls toward the limb of the star, which is an effect called limb darkening. For main sequence stars, the intensity of the star, I , is well-described as functions of $\mu = \cos \theta_*$, where θ_* is the angle between the observer and the normal vector on the surface of the star – that is, $\mu = 1$ at the center of the star, where the normal vector points directly at the observer, and $\mu = 0$ at the limb of the star, where the normal vector is perpendicular to the observer’s line of sight.

There are many different ways to characterize the intensity of the star. We will discuss the most commonly used laws for transiting planets. The linear limb darkening law,

$$\frac{I(\mu)}{I(1)} = 1 - u_0(1 - \mu). \tag{4.15}$$

was the first obvious choice, but it became clear it was insufficient for stellar theory (e.g. Klinglesmith & Sobieski 1970). It was not until the high-precision transit light curves from *Hubble Space Telescope* that it was shown to be inadequate for transit light curves (Brown et al. 2001). For the precision of many ground based transits, the linear law is still sufficient, but many have adopted a quadratic limb darkening law of the form:

$$\frac{I(\mu)}{I(1)} = 1 - u_1(1 - \mu) - u_2(1 - \mu)^2. \quad (4.16)$$

Mandel & Agol (2002) state the the quadratic limb darkening law is accurate for planetary radii for $10^{-4}(p/0.1)^2$, a precision that has never been achieved from the ground, but significantly better (20 parts per million) is achieved routinely by *Kepler* (Koch et al. 2010), for which they must use a non-linear limb darkening law of the form,

$$\frac{I(\mu)}{I(1)} = 1 - a_1(1 - \mu^{1/2}) - a_2(1 - \mu) - a_3(1 - \mu^{3/2}) - a_4(1 - \mu^2). \quad (4.17)$$

Since the transit flux given a quadratic limb-darkening law is widely applicable and significantly faster to compute than the non-linearly limb darkened transit flux, we limit our discussion to the quadratic limb darkening law. Note that in the

discussion that follows, we can reproduce the linear law precisely by fixing u_2 to be zero.

Mandel & Agol (2002) gives the quadratically limb-darkened flux during transit as:

$$F = F_0 \left(1 - \frac{(1 - u_1 - 2u_2)\lambda^e + (u_1 + 2u_2) \left[\lambda^d + \frac{2}{3}\Theta(p - z) \right] - u_2\eta^d}{1 - u_1/3 - u_2/6} \right) \quad (4.18)$$

where F_0 is the baseline flux and λ^d , and η^d are given in Mandel & Agol (Table 1, 2002) for all possible geometries. Like λ^e , λ^d and η^d only depend on p and z .

As a side note, we demonstrate that both the quadratic and linear limb darkening coefficients can be fit analytically in Appendix A.

4.4.2. PLANETARY PATH

Given these analytic expressions, generating a model lightcurve becomes a matter of computing z for all times, which is similar to computing the RV – we

calculate the true anomaly in the same way (i.e., solving Kepler's equation). Then, the three-space coordinates of the planet as seen from Earth are

$$\begin{aligned}
 r &= \frac{a}{R_*} \frac{(1 - e^2)}{1 + e \cos \theta(t)} \left(1 + \frac{M_P}{M_*} \right) \\
 X &= -r \cos(\theta(t) + \omega_*) \\
 Y &= r \sin(\theta(t) + \omega_*) \cos i \\
 Z &= r \sin(\theta(t) + \omega_*) \sin i
 \end{aligned}
 \tag{4.19}$$

where a is the semi-major axis of the planet's orbit, R_* is the stellar radius, M_* is the stellar mass, M_P is the planetary mass, and M_* is the stellar mass. Some have opted to mix ω_* and ω_P at this point – while using ω_P may make more intuitive sense, we feel the consistent use of one value for the argument of periastron reduces the chance of accidentally misapplying one or the other, and ω_* is already widely in use.

The latter term in r (i.e., $1 + M_P/M_*$) adjusts the origin to be at the stellar center rather than the barycenter and can be ignored when $M_* \gg M_P$. Z is along the line of sight, where $+Z$ is toward the observer, and the X-Y plane is the plane of the sky. Neither transits nor RV can constrain the longitude of the ascending node, Ω , which is the angle from North to the ascending node, measured counterclockwise

(i.e., the rotation of the X-Y plane), but for completeness, the orientation with respect to an observer on Earth is

$$\begin{aligned}
 X' &= X \cos \Omega - Y \sin \Omega \\
 Y' &= X \sin \Omega + Y \cos \Omega \\
 Z' &= Z
 \end{aligned}
 \tag{4.20}$$

where $-X'$ is East and $+Y'$ is North. For concreteness, we assume $\Omega = 0$, which means $X = X'$, $Y = Y'$, and $Z = Z'$. This implies that, during primary transit, the planet moves from $-X$ to $+X$ and at $X = 0$, Y is equal to the impact parameter, b . Finally,

$$z = \sqrt{X^2 + Y^2},
 \tag{4.21}$$

where z is in units of stellar radii, exactly as expected by the Mandel & Agol (2002) code. Be sure to note the sign of Z – transits and occultations occur when $z < 1 + p$, but it is a primary transit when $Z > 0$, and a secondary eclipse when $Z < 0$. This calculation is done in our program `EXOFAST_GETB` and includes the general handling of Ω , which would be useful if one would like to include astrometric measurements.

4.4.3. INPUTS

We also need to define the parametrization of the transit light curve, which is much less obvious and has been done many different ways in the literature. Most, however, have agreed upon T_C , F_0 , and $\log P$, e , and ω_* , and the quadratic limb darken parameters u_1 and u_2 .

The remaining parameters, which describe the shape of the transit, have no universally-accepted parameterization, likely because each parameterization has its own advantages and disadvantages. Seager & Mallén-Ornelas (2003) suggested T_T , the total duration of the transit, T_F , the duration of the flat part of the transit, and δ (for non-grazing transits, $\delta = p^2$), the depth of the transit.

Carter et al. (2008) suggested that a less-correlated parameterization would be τ , the duration of ingress or egress, T , the duration from mid-ingress to mid-egress, and δ .

A more physically-motivated parametrization, which we advocate, is $\log(a/R_*)$, $\cos i$, and p . The advantage to this parameterization is in the intuitive physical priors we have imposed – our experience with eccentricity bias in §4.3.2 has made us leery of non-physical priors.

Further, when fitting both photometry and RV, this parameterization is more easily portable. The disadvantage is that they are further removed from what is

actually measured (the shape of the transit) and the covariances between these parameters can be large (and therefore the convergence time may be long).

4.4.4. OTHER BIASES

If we are not careful, biases similar to the eccentricity bias can exist for all bounded parameters, which includes $\cos i$ (in the case of transiting planets when we cannot distinguish between $\pm \cos i$), or p . In order to eliminate the boundary at $\cos i = 0$, which attenuates the prior probability near edge-on inclinations, we allow $\cos i$ to be negative. Similar to negative e , this creates a discrete degeneracy between $\pm \cos i$, which we avoid by explicitly changing $\cos i$ to its absolute value at each step. Again, there is an unavoidable Lutz-Kelker bias for this positive-definite quantity.

Fortunately, unlike eccentricity, where we expect Hot Jupiters to be tidally circularized and therefore e to be exactly 0, there is no reason to expect a planet to be exactly edge on. Therefore, we are significantly less likely to encounter such a bias. Further, the theoretical interpretation of such a system does not qualitatively change if we measure a small inclination, whereas a small non-zero eccentricity for a planet that is supposed to be tidally circularized requires exotic explanations, such as anomalous values of the tidal Q factor or additional bodies perturbing the system.

We can avoid a bias in p altogether by thinking about how our model would behave if negative values were allowed. While completely unphysical, it is obvious

that a negative planetary radius would add flux during transit. To that end, we allow negative planetary radii, calculate the flux decrement as if it were positive, and then add the flux to the baseline rather than subtract it. This avoids both the prior bias, since there is no boundary, and the Lutz-Kelker bias, since a negative value has a well-defined likelihood. If the best-fit (median) p is negative, it is likely there is no transit at all. If, however, we see a small tail at negative values but the result is statistically significant, we can be more confident that the non-zero measurement is not a result of a bias in fitting. This will be particularly useful when measuring small transits and secondaries eclipses with low significance.

There is no obvious way to implement a similar scheme when δ is used instead of p , since in order to determine the model (Equation 4.18), we must take the square root, which would be imaginary if $\delta < 0$.

4.4.5. TRANSIT BEST FIT

Transits are first identified in transit surveys using relatively small telescopes that monitor large areas of the sky at once (e.g., HAT, WASP). Transit candidates are first identified with a Box Least-Squares (BLS) algorithm (Kovács et al. 2002), which extracts the duration, depth, period, and T_C of the transits. We cannot improve upon this method to find the global minimum. However, for most high-precision transits that will be fit in the manner outlined in this code, these

quantities will be already roughly known, greatly simplifying the problem to finding a local minimum around relatively well-behaved region of parameter space.

With good starting values for P and T_C , we can begin with fairly generic guesses for the rest of the parameters and standard local minimization routines like AMOEBA work well to find the local minimum. Once found, we ensure the fit is self consistent by scaling the errors by $\sqrt{\chi_\nu^2}$, which forces $\chi_\nu^2 = 1$.

4.5. RADIAL VELOCITY AND TRANSIT

The models for the two simultaneously are the same, but the advantage of fitting radial velocity and transit data simultaneously is that they both constrain many of the same parameters, which improves the quality of both fits and ultimately gives us a clearer picture of the system as a whole. Further, we can include additional affects with no penalty, such as the reflex motion of the star (Eq. 4.19) or light travel time delay in the planetary transit. More important, covariances between parameters in the different data sets may be unintuitive and non-negligible. Fitting the two data sets separately assumes the covariances between the parameters in the two data sets is zero, whereas a simultaneous fit naturally takes all of this into account.

4.5.1. INPUTS

The disadvantage to a simultaneous fit is having to rethink the parameterization of the problem, since the overlapping constraints are not always intuitive.

$\log P$, e , ω_* , and T_C trivially overlap between the two data sets. $\log K$, γ , dv/dt , are still mostly independent parameterizations for RV, and F_0 , $\cos i$, and p are mostly independent for photometry.

The only non-intuitive common constraint is on the semi-major axis, a between the RV and transit data. Essentially, linking the two data sets requires that we know the M_* and R_* . Of course, the transit geometry itself scales with R_* , for which we have no constraint. That is, we must determine that elsewhere in order to derive the physical parameters we care about (Seager & Mallén-Ornelas 2003). Since we must do this anyway, it behooves us to do it during each step in the Markov chain and use all of the information of the two data sets to their full advantage while simultaneously exploring the covariances between M_* and R_* and the other parameters.

4.5.2. STELLAR MASS AND RADIUS

Torres et al. (2010) empirically determined a relation to derive the masses and radii of stars based on a large sample of well-measured stars, calibrated via the stars'

effective temperatures, T_{eff} , surface gravities, $\log g$, and metallicities, $[\text{Fe}/\text{H}]$ (see their Table 4), which we can measure given a high-quality spectrum.

Therefore, in order to combine the simultaneous constraints from RV and transits, we add the T_{eff} , $\log g$, $[\text{Fe}/\text{H}]$ to our input parameters and calculate the M_* and R_* using `MASSRADIUS` at each step in the MCMC chain. The covariance between these parameters is large, so ideally one should take this into account or fit the spectrum simultaneously with the transit and RV. However, since neither the spectrum nor the covariances are typically given in the literature, we simply step in these parameters and add a penalty term to the χ^2 of $((X_i - X)/\sigma_X)^2$ where X is the best fit value, σ_X is its quoted uncertainty, and X_i is its value in the current step of the Markov Chain.

Typically, the uncertainties in T_{eff} , $\log g$, and $[\text{Fe}/\text{H}]$ dominate the resultant errors in the stellar mass and radius. However, it is important to note that the Torres et al. (2010) relation has uncertainties that correspond to 6.4% errors in mass and 3.2% errors in radius. A proper handling of this scatter is beyond the scope of our paper. Other methods, such as isochrone fitting, also do not take into account the uncertainties in the models.

Enoch et al. (2010) produced a similar relation based on the Torres et al. (2010) relations with the stellar density, ρ_* , replacing $\log g$. The motivation was based on the fact that the stellar density is directly constrained by the light curve (Seager &

Mallén-Ornelas 2003) and the measurement of $\log g$ is imprecisely determined from spectra.

However, one can calculate the density from the stellar mass and radius, and the Enoch et al. (2010) relations are not self consistent – that is, the input density does not agree (precisely) with density derived from the output mass and radius. While the inconsistency is small, and well within the model errors, we derive the same constraint on the stellar density using the Torres et al. (2010) relations in a self-consistent manner from the output mass and radius based on $\log g$.

Another advantage of using these as free parameters is that we can uniquely determine the limb darkening parameters by interpolating the tables from Claret & Bloemen (2011) for each value of $\log g$, T_{eff} , and $[\text{Fe}/\text{H}]$ (and the observed bandpass) without having to fit them explicitly, which we do in QUADLD. This means that, if the data are of a sufficient quality to constrain the limb darkening parameters, they become an independent constraint on $\log g$, T_{eff} , and $[\text{Fe}/\text{H}]$ (and therefore the mass/radius of the star). Of course, the risk is that if the Claret & Bloemen (2011) models are not accurate, they will bias the mass/radius measurement of the star.

The least-intuitive prior is our uniform step in these parameters. To understand it, we uniformly distributed T_{eff} between 3026 and 6933, $\log g$ between 1.5 and 5.0, and $[\text{Fe}/\text{H}]$ between -1.0 and 0.56, which are the ranges of values for all exoplanet hosts in the exoplanets.org database. We then calculated the mass and radius with

the Torres et al. (2010) relation. The histograms of masses and radii can be seen in figures 4.8 and 4.9. Ideally, our prior would resemble the underlying distribution of planet hosts, which is poorly understood.

4.5.3. PLANETARY MASS

Once we know the mass of the primary, the radial velocity and transit parameters constrain the mass of the secondary precisely:

$$K = \left(\frac{2\pi G(M_* + M_P)}{P} \right)^{1/3} \frac{M_P \sin i}{M_* \sqrt{1 - e^2}} \quad (4.22)$$

While typically, it is reasonable to assume that $M_* \gg M_P$, in which case the equation simplifies to

$$K = \left(\frac{2\pi G}{PM_*^2} \right)^{1/3} \frac{M_P \sin i}{\sqrt{1 - e^2}}, \quad (4.23)$$

there is no need to make this approximation when the planet also transits (i.e., we know i). Equation 4.22 can be either be solved for M_P analytically by re-writing as a quartic:

$$M_P^4 + M_* M_P^3 - \left(\frac{K M_* \sqrt{1 - e^2}}{\sin i} \right)^3 \frac{P}{2\pi G} = 0 \quad (4.24)$$

or numerically with a binary search. While the analytic solution to the quartic is sufficiently complicated that its benefit over the numerical solution is not generally obvious, since IDL is comparatively slow at (serial) binary searches, we opted to solve the quartic analytically. Generally, the solutions of a quartic are complex and may be negative, but since we know the solution corresponds to a physical mass, we may just consider the positive, real solution, which simplifies the calculation. This calculation is done in `KTOM2`.

By solving exactly for the secondary mass, we do not implicitly assume that the primary is stationary, and so our method also works for eclipsing binaries.

4.5.4. ROEMER DELAY

Having the semi-major axis in physical units gives the system an absolute scale, which allows us to make other minor corrections without introducing any additional parameters, such as the Roemer delay in the target system. Similar to the correction from JD to BJD, the light may be early or delayed due to the light travel time from the moving source. Rather than correct the as-seen flux as suggested by Loeb (2005), it is more straight-forward to transform the time stamp into the target barycentric frame (Eastman et al. 2010b), which can be done with our code `BJD2TARGET`. This effect is most important when trying to reconcile observations for which the information comes from different points in the planetary orbit, such as

primary transits and radial velocities or multiple transiting planets. However, even in the case of a single observation of a transiting Hot Jupiter ($a \sim 0.05AU$), the ingress is delayed by about 0.15 seconds.

The information from the RV comes from the center of the star, which is also offset from the Barycenter. For Hot Jupiters, this offset is trivial (~ 10 ms), but grows with the semi-major axis and the mass ratio of the system.

However, as noted by Eastman et al. (2010b), this should only be done during the calculation of the model for consistency. When quoting the parameters, the most-relevant and most-precisely-known times will be the times in the Solar System Barycentric frame (i.e., BJD_{TDB}). We convert back to this frame before quoting the final parameters to avoid confusion using `TARGET2BJD`.

Finally, we note that we do not take into account γ when calculating the Roemer delay. Since RVs are often arbitrarily normalized, we usually do not always have a good measurement of the absolute systemic velocity. More important, the only practical effect such a correction would have would be to change the period by a factor of γ/c . While this $\lesssim 0.1\%$ difference is already statistically significant in many systems, the practical implications of such a difference will be completely obscured by the errors in the other parameters for the foreseeable future, and quoting this corrected period would only make deriving the ephemeris more complicated.

So now, with a mere 13 parameters, γ , dv/dt , T_C , $\log P$, e , ω_* , K , $\cos i$, F_0 , R_P/R_* , $\log g$, T_{eff} , and $[\text{Fe}/\text{H}]$, we can describe the entire system in a self consistent manner, including often neglected effects.

4.5.5. COMBINED BEST FIT

Since we have a very good guess for the starting values for each parameter from the individual fits, no global fit is required. Assuming they are consistent with one another, we only need a slight refinement to find the best fit of the combined data sets. Therefore, using the starting values for the best fits we found in both individual cases, we run a quick `AMOEBA` minimization to find the local minimum of the combined data set, and we are finally ready to use our MCMC code, `EXOFAST_MCMC` to determine the model uncertainties and covariances.

4.6. OUTPUTS

`EXOFAST_MCMC` outputs a three dimensional array, the number of parameters by the number of links in the chain by the number of chains. After we ensure the chains are well mixed, we discard the first 10% of every chain to eliminate any bias from the starting values, then reform the array into a two dimensional array – the number of parameters by the number of steps times the number of chains. This

array is then the input to several other useful routines to synthesize the chain into easily understood results.

Perhaps the most powerful aspect of MCMC is the fact that, with the parameter chains in hand, it is trivial to generate uncertainties and covariances for derived quantities – we simply perform the appropriate transformation to the entire chain. We can easily concatenate derived parameters (e.g., T_P , $e \cos \omega_*$, etc.) to the 2D parameter array before using any of the following analysis tools.

In particular, because of the high covariances among T_C , T_P , T_S , e , and ω_* , estimating the errors of derived times after the fact without knowing the covariances is unreliable. Generally, quoting all potentially relevant times directly from the fit is a good practice.

4.6.1. PLOTS

We include several useful visualization tools to present the parameter distributions and the model fit in an easily digestible form. The histogram of parameter values is often a useful indicator of how well-mixed the chains are. We provide a tool, `EXOFAST_PLOTDIST` which accepts the two dimensional parameter array and creates histograms of each value.

It is also quite useful to check the parameters for covariances which may influence the fit. However, because the chains are very large, simple scatter plots

are unwieldy and become unreadable. A contour plot makes much more sense to see the covariances between parameters, but IDL's built-in contour function is not well suited for this task. Our code, `EXOFAST_ERRELL`, draws custom constant probability contours of two-parameter distributions, or returns the paths of the contours for easy manipulation (e.g., over-plotting contours from two different fits to compare, as in Figure 4.2). It can also give the indices of points outside a particular contour. Our plotting code uses this contour routine to plot the covariances of every parameter against every other parameter.

The plots generated are not publication quality – writing such a routine in a general way is not practical. However, the routine may be tweaked relatively simply or used as a template to generate better-quality plots specific to the input data and what is to be illustrated.

Table 4.10 shows each of the 33 parameters and derived quantities for the simulated data set of combined RV + transit fit. Our plotting code plots each parameter distribution, 8 to a page as shown in figure 4.10. It can optionally place a vertical line to mark a single value for each parameter. In this case, we plotted the input value of the simulated data.

Our code also shows the covariances for every possible combination of N parameters, which means $N(N - 1)/2$ plots, 16 per page (528 plots in this case). Values marked with lines in the 1D parameter distributions are shown as black dots

in the 2D covariances, and the numerical value of the covariance is shown at the top of each plot. Again, an example page is shown in figure 4.11.

4.6.2. PARAMETERS AND UNCERTAINTIES

An additional output to `EXOFAST_PLOTDIST` is the median values and 68% confidence interval for each of the N parameters, in an 3 by N element array. The first column is the median value, the second is the upper error bar, and the third is the lower error bar. This array can be used on its own, but it is designed to be an input to `EXOFAST_LATEXTAB`. This program rounds the two-sided errors to two significant digits each and rounds the median value to the higher precision error. When the error is symmetric, we use the \pm symbol; otherwise, we list both errors.

Table 4.10 shows an example table generated automatically from one of the fits of simulated data mentioned in §4.7. Many may find it easier to edit the title, label, caption, and parameter names in LATEX rather than as input in IDL, but a fast and accurate way to go from parameter arrays to properly-formatted latex source code is very convenient. Even if it is not intended to be inserted into a paper, it is often more readable as LATEX code than simply printing the output array.

4.6.3. MODELS

The provided χ^2 routines described above have a global debug flag. If set, it will plot the data and the model each time it is called. As the name implies, it is useful for debugging if the fit is not working. It is also instructive – if, for example, the debug flag is left on during an AMOEBA fit, it will essentially play a movie of the routine settling into the best fit. In this manner, you can gain an intuitive feel for how the algorithm works (or why it’s failing). Obviously, this slows down the fit by orders of magnitude – it is not intended to be used during a standard fit. Lastly, once the best fit is found, we can call the χ^2 routine one final time with the debug flag to generate a plot of the best fit model with the data.

The routines also have a “PS” flag, which can make a postscript output. An example of the best-fit transit generated from this output is shown in figure 4.12 and the best-fit RV is shown in figure 4.13.

4.7. VALIDATION

4.7.1. KNOWN SYSTEMS

We fit systems as diverse as HD189733, a planet in a circular orbit with a period of 2.2 days, and HD80606, the most eccentric planet known ($e \sim 0.93$) with the longest period of all confirmed transiting planets, 111 days. Using the RV portion of

this code, we refit the RV data presented in Winn et al. (2009) and recovered their best values.

4.7.2. NEW SYSTEMS

The radial velocity portion of this code has been used to fit the radial velocity data in Fleming et al. (2010) and Lee et al. (2010), though we used slightly modified χ^2 function to incorporate data from multiple sources.

4.7.3. SIMULATED DATA

As described in §4.3.2, we ran many simulations of the RV. With the exception of the eccentricity bias discussed in detail there, all measured values were statistically consistent with the input values. Further, all 4,400 of these simulations were robustly fit without any intervention.

4.8. ADDITIONAL UTILITIES

4.8.1. READEXO

READEXO reads the exoplanets.org database (Wright et al. 2011) into an IDL structure. Each header entry, described on their website⁶, becomes a unique tag

⁶<http://exoplanets.org/exohelp/exoHelp.html>

name for the structure. The code will automatically adapt to the addition of rows and columns into the exoplanets.org database, and a flag can be set to automatically update the local copy. This code is useful for more general manipulation and comparison of all exoplanetary data, and integration into software pipelines (e.g., using the literature values of a known planet as the starting seed for a fit). However, we caution strongly that the selection effects inherent to this sample of planets is very poorly understood and care should be taken not to over-interpret results that these tools make trivial. If you use this code, please also cite their paper and acknowledge their work.

4.8.2. SYS-REM

We implement the systematics removal algorithm (Sys-Rem), as outlined in Mazeh et al. (2007). Most implementations of SYSREM require an input of the airmass at each observation as a starting seed. While custom seeds are allowed, we have seen no starting seed consistent with airmasses that creates a different result than a seed of all ones. We make that our default seed and do not require a separate seed.

4.9. ONLINE TOOLS

4.9.1. EPHEMERIDES

There are now so many transiting planets that, on any given night from any given location on Earth, we are very likely to be able to observe at least one of them (Eastman et al. 2010a). It is useful, then, to quickly determine which those are in order to plan observations or fill gaps in observing schedules. We present a code that calculates which planets are transiting or eclipsing from a particular observatory on a particular date. This code makes use of the exoplanets.org database. We also created a web interface to this code⁷, which updates its local copy of the exoplanets.org database daily. For the predicted secondary eclipse times of non-circular orbits and especially the transit times of RV planets, the precision is likely to be much worse than what we could derive from the original data because these times were derived from P , e , ω_* , and T_P (or T_C) without the covariance matrix. These could be greatly improved if in the future, published results included T_C and T_S directly from the fits, and these were included in the exoplanets.org database.

⁷<http://astrutils.astronomy.ohio-state.edu/exofast/ephem.shtml>

4.9.2. TRANSIT FITTING

We provide an intuitive online interface to the basic features of our local transit fitter⁸, which can fit and remove an arbitrary number of systematics. Since the fit is only local, it often needs relatively good starting values to fit correctly. These can be entered manually and checked (without fitting) or they can be pulled from the exoplanets.org database by selecting the appropriate planet from a pull down menu. This pull down menu will update daily as new planets are added to the exoplanets.org database. When fitted, both the data (normalized to 1 and with systematics removed) and the fit will be plotted and the parameters will be output to the screen.

Unfortunately, since these codes are run on the server side, it is not practical to support the full MCMC implementation of the code, so only the best fit values and very rough estimates of the errors (what would be used as the MCMC steps) are reported. This online fitter is not intended for research-quality fits.

4.9.3. RV FITTING

Our global, single planet RV fitter is also online⁹. It will report the χ^2 and parameters for the top N peaks in the Lomb-Scargle periodogram within a

⁸<http://astrutils.astronomy.ohio-state.edu/exofast/transit.shtml>

⁹<http://astrutils.astronomy.ohio-state.edu/exofast/rv.html>

user-specified range (default is 1 day to the range of input times), where N is a user selectable value. It will plot the best fit and data, and can be fit with a slope and/or forced to be circular. When the phase coverage is good and there is only one planet present, we have found this routine to be very robust. Like the transit fitter, since this is run on the server side, we do not implement the full MCMC fit, and only provide rough estimates of the errors – this is not intended for publication quality fits.

4.9.4. QUADRATIC LIMB DARKENING

Our last online code linearly interpolates the Claret & Bloemen (2011) quadratic limb darkening tables for given values of $\log g$, T_{eff} , and $[\text{Fe}/\text{H}]$ in any of the following bands: Johnson/Cousins U, B, V, R, I, J, H, and K; Sloan u' , g' , r' , i' , and z' ; *Spitzer* 3.6 μm , 4.5 μm , 5.8 μm , and 8.0 μm ; Kepler; CoRoT; and Stroemgren u, v, b, and y ¹⁰. This can also draw the stellar parameters from the exoplanets.org database, if available.

4.10. SUMMARY AND SUGGESTIONS FOR FUTURE EXTENSIONS

The true power of this code is its extensibility. The core MCMC code can be run with any χ^2 routine to fit an arbitrary model to an arbitrary data set.

¹⁰<http://astrutils.astronomy.ohio-state.edu/exofast/quadld.shtml>

It is relatively straight-forward to modify the provided χ^2 routines to include a different prior or add other effects. The major complication is in either finding the appropriate step size for each parameter in the MCMC chain or in robustly finding the initial best fit from which the program can calculate the appropriate step size.

Therefore, this code naturally lends itself to extension. One could include additional parameters to encapsulate variations in central transit time, duration, or depth, add parameters to describe the secondary eclipse, fit multi-planet RV systems, fit multi-band transit data, or fit an entirely new model to completely different data. Use of the core MCMC routine, `EXOFAST_MCMC` is analogous to IDL's built-in fitting functions `AMOEB` and `LMFIT`.

We would like to thank Robert Siverd for useful discussions about the eccentricity prior. This research has made use of the Exoplanet Orbit Database and the Exoplanet Data Explorer at exoplanets.org. We would also like to thank Wayne Landsman for curating the IDL astronomy library, from which we use many functions (Landsman 1993).

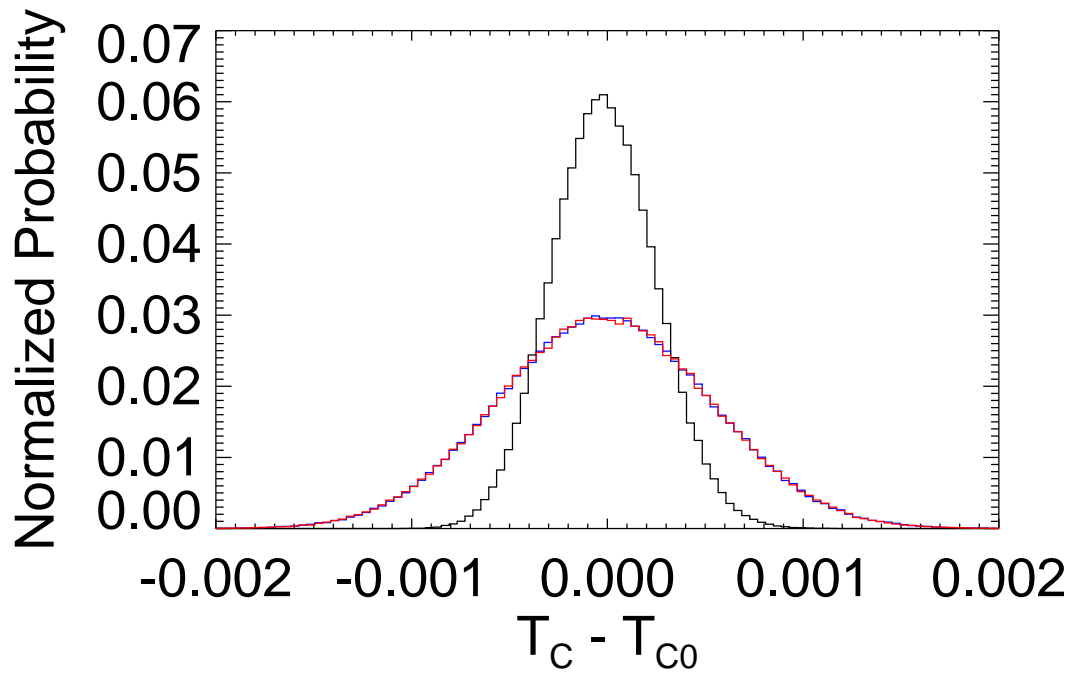


Fig. 4.1.— The parameter distributions for T_C (black), $T_C + P$ (red) and $T_C - P$ (blue), where T_C is the value closest to the error-weighted mean of the input times for the simulations described in §4.7. While the central value does not change, the uncertainties are minimized at the error-weighted mean.

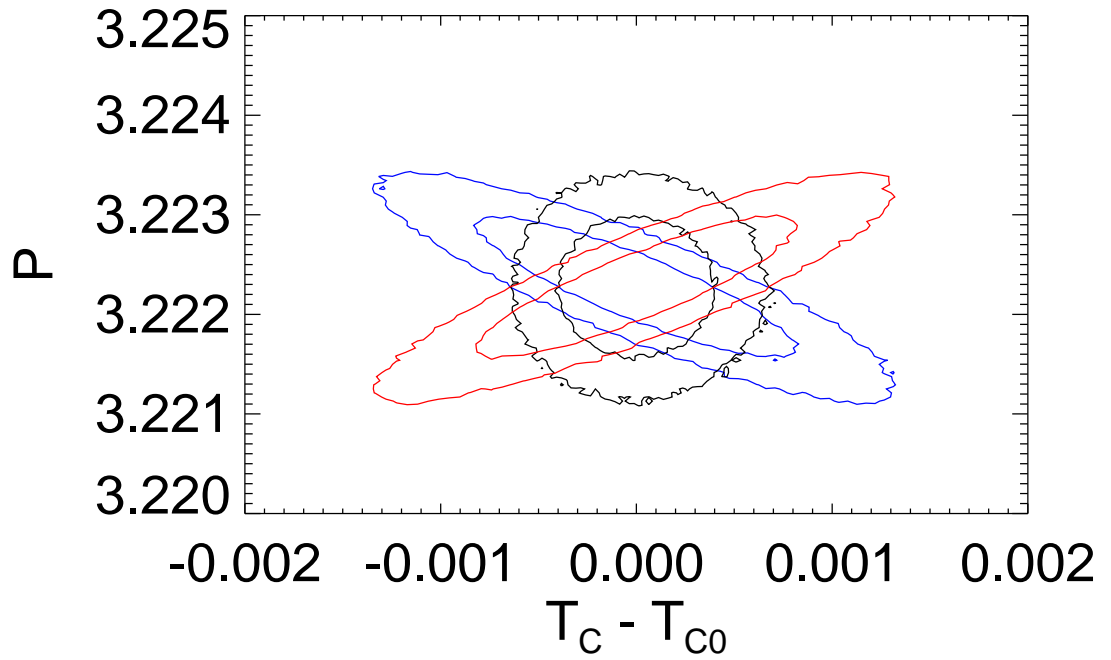


Fig. 4.2.— The covariance between P and T_C (black), $T_C + P$ (red) and $T_C - P$ (blue), where T_C is the error-weighted mean of the input times for the simulations described in §4.7. While the central values do not change, the covariances are minimized at the error-weighted mean.

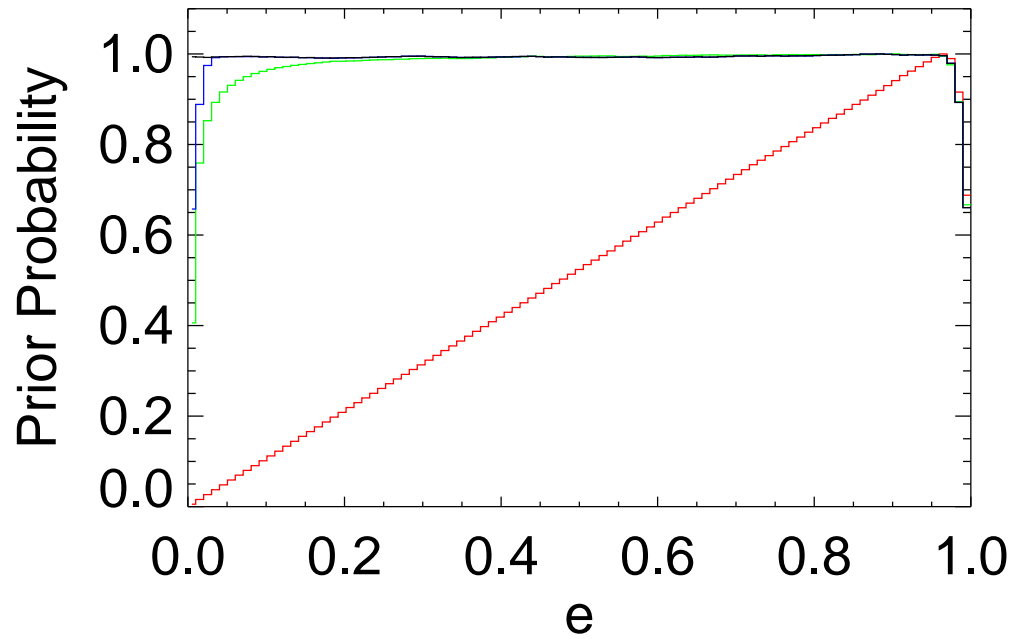


Fig. 4.3.— The prior distributions of e and ω while stepping in $e \cos \omega$ and $e \sin \omega$ (red), $e \cos \omega$ and $e \sin \omega$ while correcting for the prior as Ford (2006) proposed (green), uniform in e and ω bounded such that $0 \leq e < 1$ (blue), and uniform in e and ω bounded such that $-1 < e < 1$ (black).

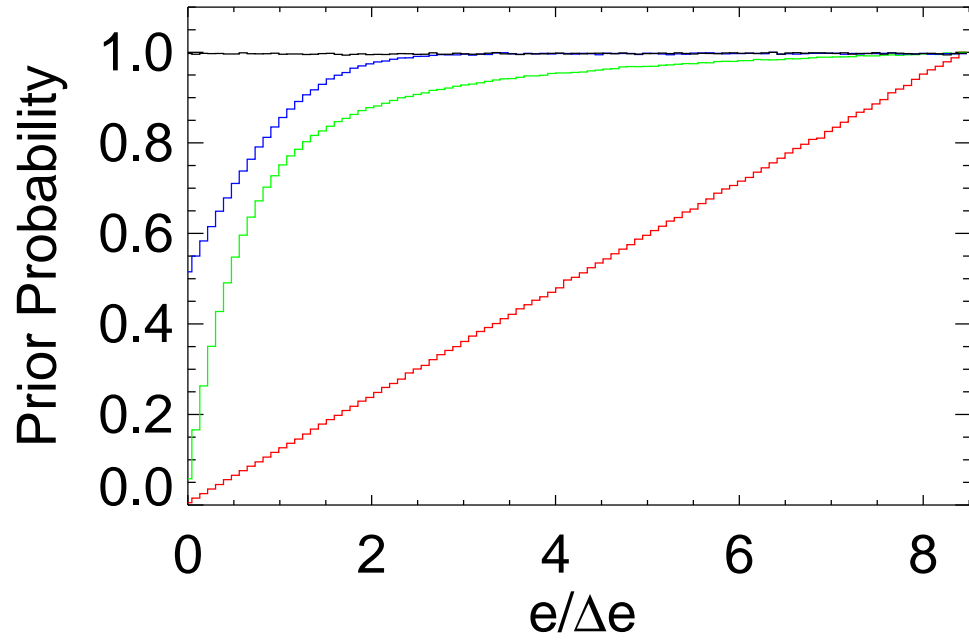


Fig. 4.4.— A zoom of figure 4.3 from $e=0$ to 0.1, re-normalized, with the same color coding. The bottom axis has been converted to units of the step size of e used in the Markov Chain, which should ideally correspond to its 1-sigma uncertainty. This shows that even after attempting to correct for the linear prior, a relatively significant prior remains that biases the measured eccentricity away from zero.

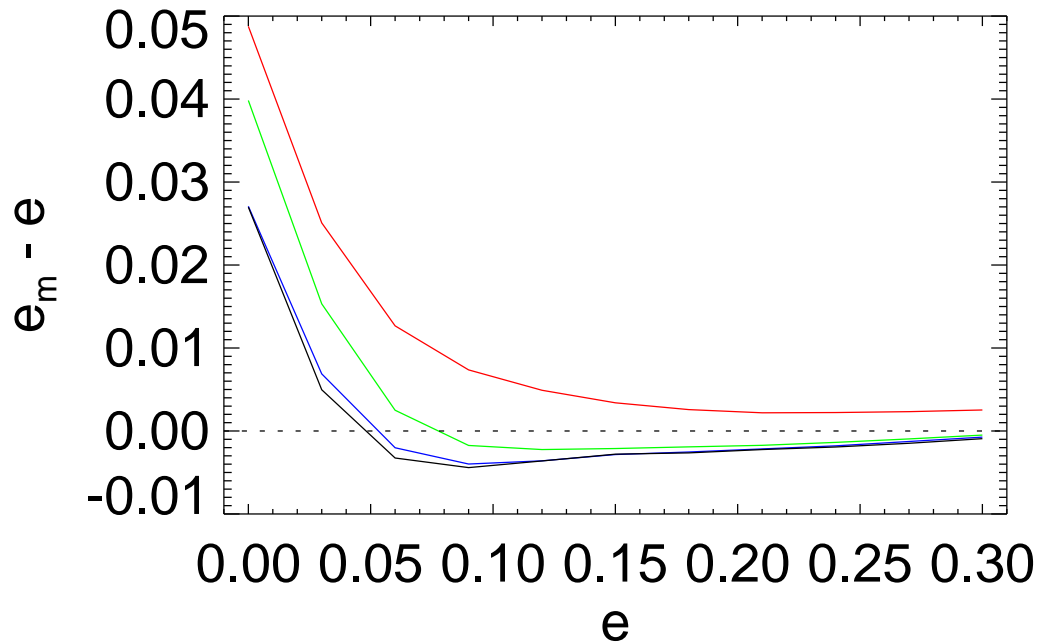


Fig. 4.5.— The difference between the measured and actual eccentricity as a function of the intrinsic eccentricity for the 4 different priors, using the same colors as figure 4.3. The values are the average of the median values for all 100 fits. The statistical errors for each of the 100 fits was roughly 0.03. The small, systematic offset between the measured and intrinsic eccentricities is due do the fact that the ensemble of 100 fits we used all had the same “random” offsets, and therefore the same skewed eccentricity. This offset would shrink with additional random trials.

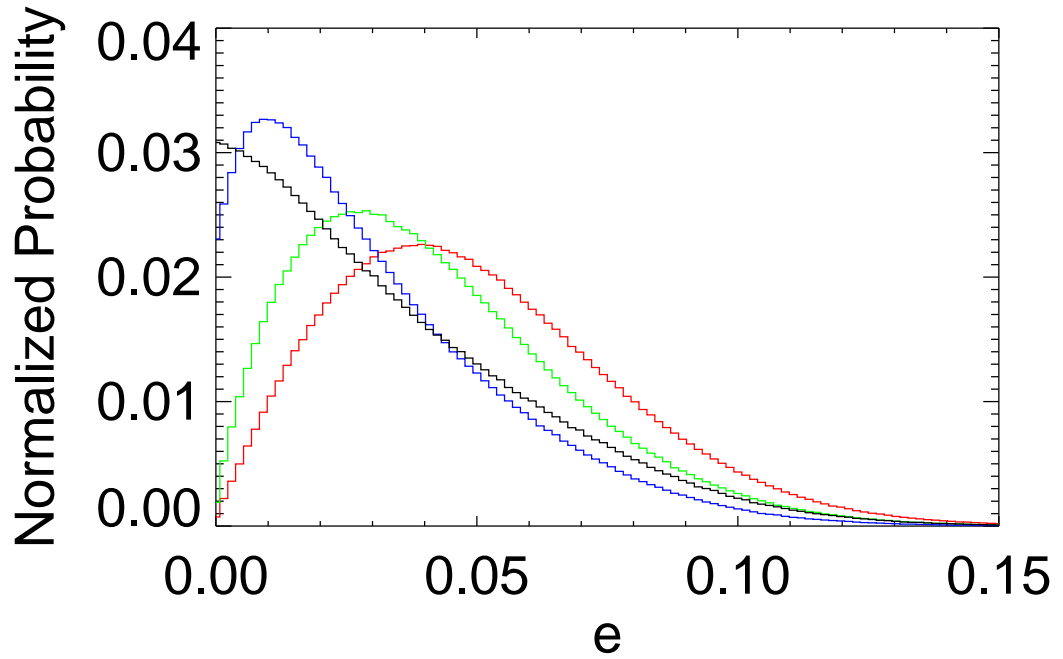


Fig. 4.6.— The parameter distributions for all 100 $e = 0$ fits for each of the four priors, using the same colors as figure 4.3. We can clearly see the truly uniform prior in e is a substantial improvement over all other methods.

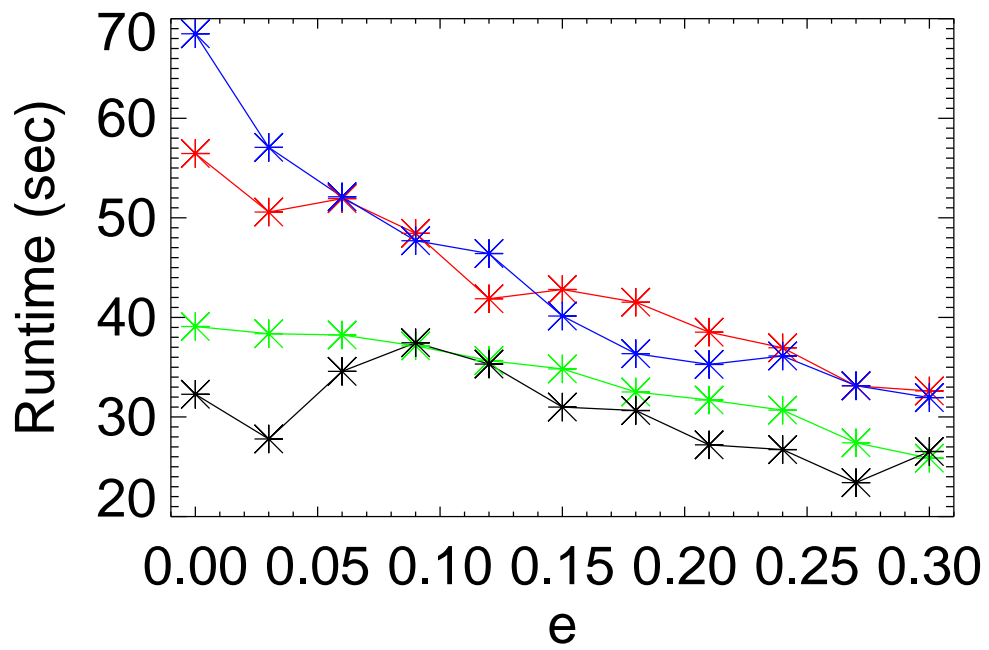


Fig. 4.7.— The mean execution time for a single fit as a function of eccentricity for the four priors. See figure 4.3 for the colors.

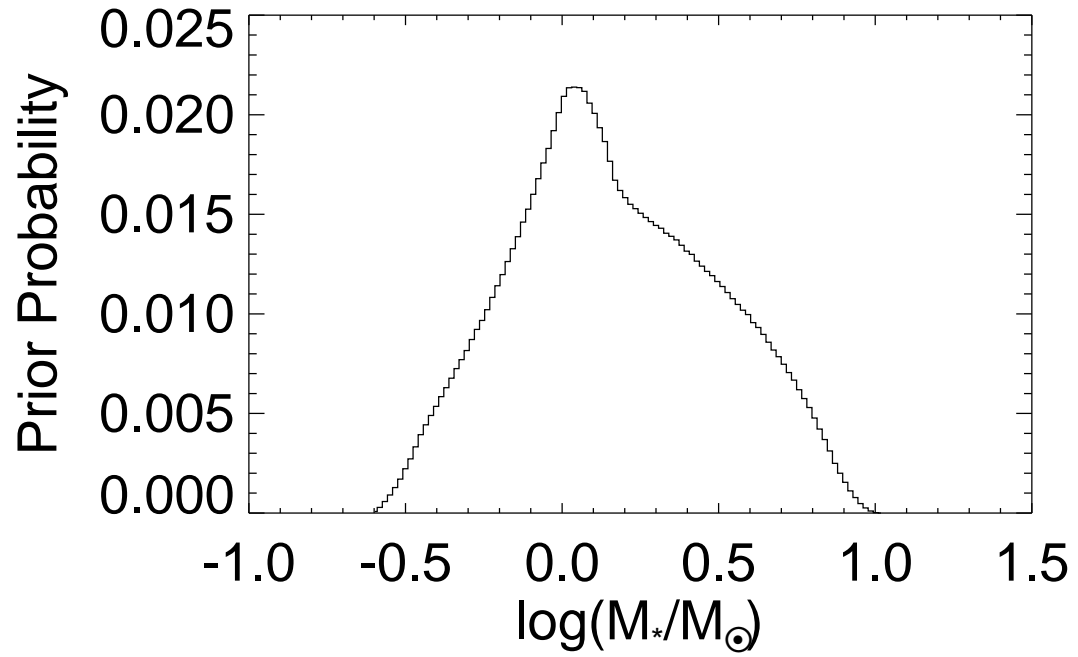


Fig. 4.8.— The prior probability distribution of stellar masses, from stepping uniformly in T_{eff} , $\log g$, $[\text{Fe}/\text{H}]$, throughout the range of exoplanet host values found in the exoplanets.org database.

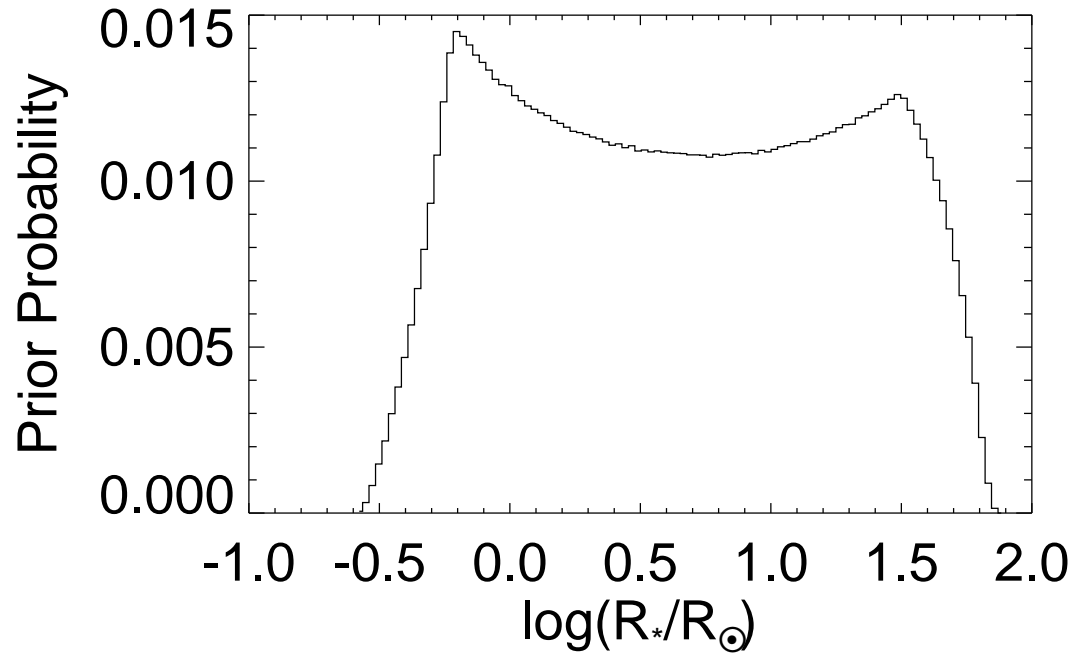


Fig. 4.9.— The prior probability distribution of stellar radii, from stepping uniformly in T_{eff} , $\log g$, $[\text{Fe}/\text{H}]$ and using the Torres et al. (2010) relation in the range of exoplanet host values found in the exoplanets.org database.

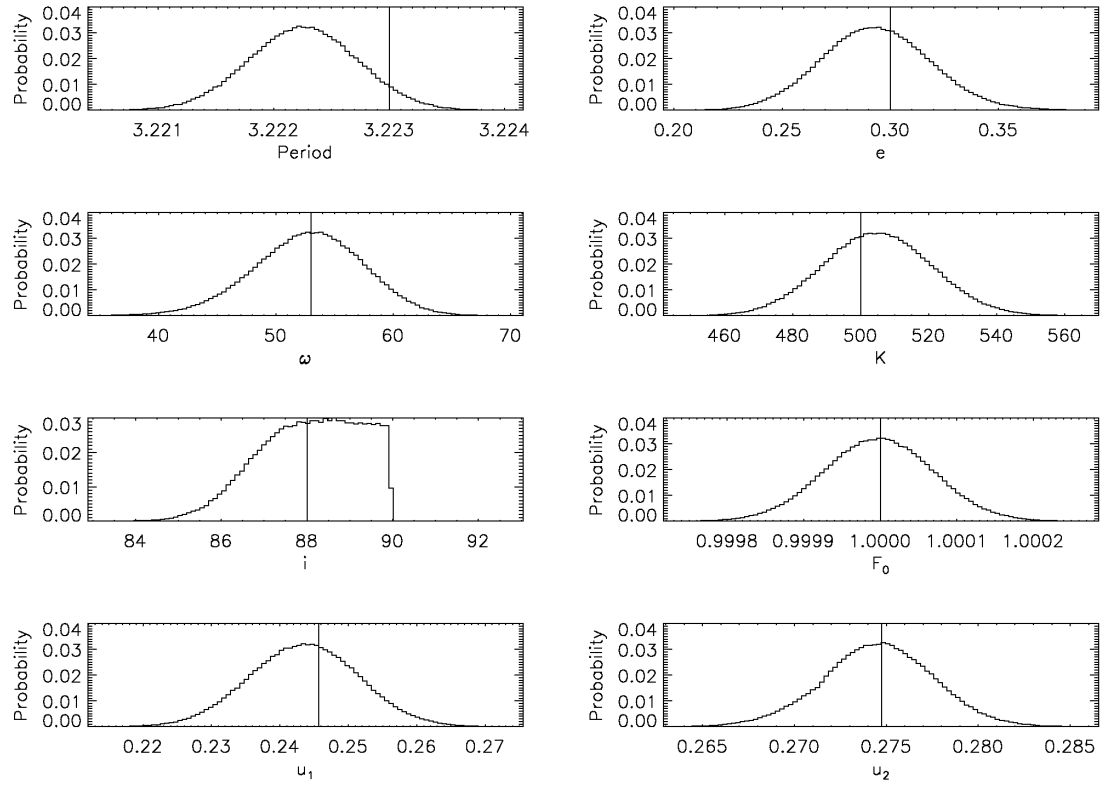


Fig. 4.10.— The parameter distributions for a subset of parameters in the combined RV + transit fit, as generated by EXOFAST_PLOTDIST. The line represents the input value for the simulated data.

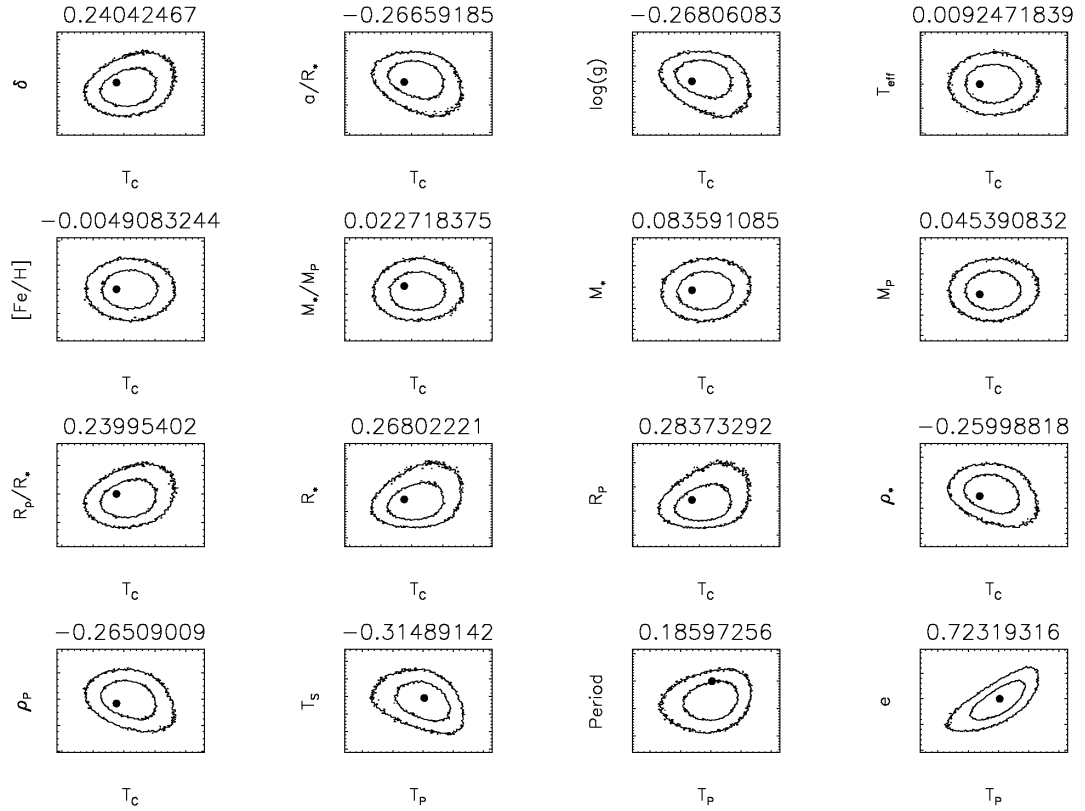


Fig. 4.11.— The covariances for a small subset of parameters in the combined RV + transit fit. The contours are the 68% and 95% probability contours, and the black dot is the input value. The number above the plot is the covariance between the parameters.

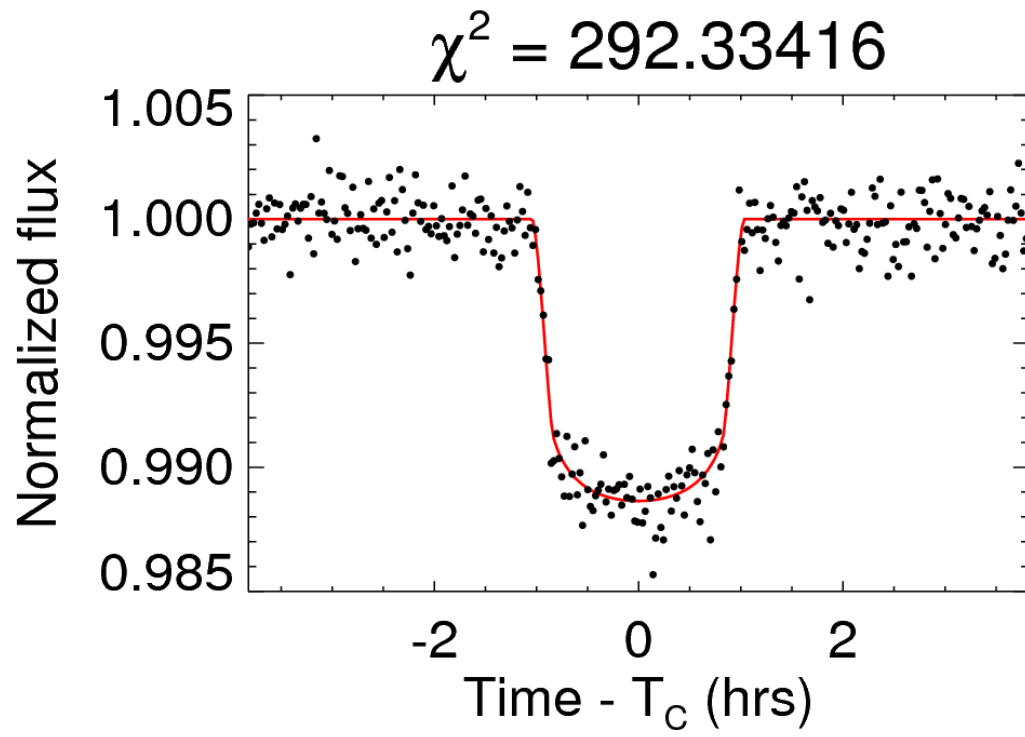


Fig. 4.12.— The best-fit transit of the simulated data, as plotted with the debug flag set in the χ^2 routine.

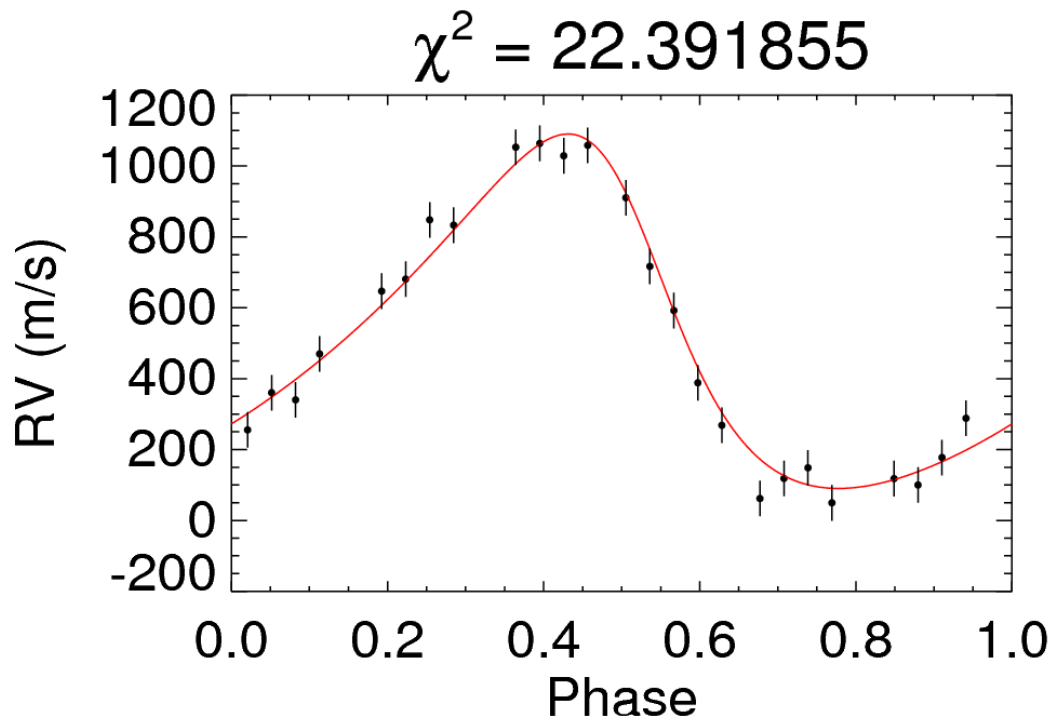


Fig. 4.13.— The best-fit RV of the simulated data, as plotted with the debug flag set in the χ^2 routine.

Parameter	Value
γ (m/s)	498 ± 11
dv/dt (m/s/day)	-0.13 ± 0.12
T_C (BJD_{TDB})	$2455563.62809^{+0.00030}_{-0.00029}$
T_P (BJD_{TDB})	$2455563.451^{+0.028}_{-0.033}$
T_S (BJD_{TDB})	2455562.387 ± 0.037
<i>Period</i> (days)	3.22227 ± 0.00047
e	$0.294^{+0.026}_{-0.025}$
ω (degrees)	$53.1^{+4.5}_{-4.8}$
K (m/s)	505 ± 16
i (degrees)	$88.2^{+1.2}_{-1.3}$
F_0	1.000001 ± 0.000071
u_1	$0.2439^{+0.0079}_{-0.0080}$
u_2	0.2747 ± 0.0030
$\log P$	$0.508163^{+0.000064}_{-0.000063}$
$e \cos \omega$	0.176 ± 0.018
$e \sin \omega$	0.235 ± 0.030

Table 4.1. An example output of EXOFAST_LATEXTAB

(cont'd)

Table 4.1—Continued

Parameter	Value
$\log K$	2.704 ± 0.014
$\cos i$	$0.031^{+0.023}_{-0.021}$
a (AU)	0.04332 ± 0.00042
δ	$0.00993^{+0.00019}_{-0.00017}$
a/R_*	$9.94^{+0.44}_{-0.51}$
$\log(g)$	$4.507^{+0.038}_{-0.046}$
T_{eff} (K)	5801 ± 53
$[Fe/H]$	-0.00 ± 0.11
M_*/M_P	$302.7^{+10.}_{-9.8}$
M_* (M_\odot)	$1.042^{+0.031}_{-0.030}$
M_P (M_\odot)	$0.00344^{+0.00013}_{-0.00012}$
R_p/R_*	$0.09967^{+0.00097}_{-0.00085}$
R_* (R_\odot)	$0.938^{+0.055}_{-0.043}$
R_P (R_\odot)	$0.0934^{+0.0061}_{-0.0045}$
ρ_* ($g\ cm^{-3}$)	$1.78^{+0.25}_{-0.26}$
ρ_P ($g\ cm^{-3}$)	$5.95^{+0.96}_{-1.0}$

APPENDIX A

ANALYTIC FIT OF THE LINEAR AND QUADRATIC LIMB DARKENING

Given the quadratic limb darkening law (eq 4.16) and the flux during transit (eq 4.18), if we make the following definitions

$$\begin{aligned}x_0 &= 1 - \lambda^e \\x_1 &= 2/3(\lambda^e - \Theta(p - z)) - \lambda^d \\x_2 &= \eta^d + \lambda^e/2 \\c_0 &= F_0 \\c_1 &= F_0 \frac{u_1 + 2u_2}{1 - u_1/3 - u_2/6} \\c_2 &= F_0 \frac{u_2}{1 - u_1/3 - u_2/6}\end{aligned}\tag{A.1}$$

the flux during transit simply becomes:

$$F = c_0x_0 + c_1x_1 + c_2x_2\tag{A.2}$$

Since x_0 , x_1 , and x_2 , are solely functions of the transit geometry (p and z), they can be calculated independent of the limb darkening and are now optional outputs of EXOFAST_OCCULTQUAD. Using these, we can analytically solve for the coefficients c_0 , c_1 , and c_2 by performing a standard linear χ^2 minimization (see Gould 2003). Then the limb darkening coefficients, u_1 and u_2 simply become:

$$\begin{aligned}
 u_1 &= \frac{c_1 - 2c_2}{c_0 + c_1/3 - c_2/2} \\
 u_2 &= \frac{c_2}{c_0 + c_1/3 - c_2/2}
 \end{aligned}
 \tag{A.3}$$

This analytic fit can greatly increase the speed of any fitting routine by reducing the dimensionality of the non-linear solver, particularly in cases where data are taken with multiple bands. However, in the low signal-to-noise regime where the data have little power to constrain the limb darkening, this analytic fit can give non-physical results, in which case it is better to fit a linear limb darkening law, fix the values at the Claret & Bloemen (2011) values, fit them non-linearly with a prior, or as we do in the RV + transit fit – allow them to vary by interpolating the Claret & Bloemen (2011) limb darkening tables with each new $\log g$, T_{eff} , and $[\text{Fe}/\text{H}]$ during the Markov chain.

These relations trivially simplify in the case of linear limb darkening, when

$$u_2 = 0:$$

$$x_0 = 1 - \lambda^e$$

$$x_1 = 2/3(\lambda^e - \Theta(p - z)) - \lambda^d$$

(A.4)

$$c_0 = F_0$$

$$c_1 = F_0 \frac{u_1}{1 - u_1/3}$$

And then,

$$u_1 = \frac{c_1}{c_0 + c_1/3}$$

(A.5)

Given the same form (x_0 and x_1 are the same in the quadratic and linear cases), one could even fit both laws to see if the improvement in χ^2 justifies the addition of the extra parameter.

APPENDIX B

EXOFAST_OCCULTQUAD

The original IDL code to analytically calculate the flux decrement during transit presented in Mandel & Agol (2002) was more than an order of magnitude faster than the previous numerical method (Charbonneau et al. 2000; Henry et al. 2000). However, the IDL version was a simple line by line translation of the Fortran code, which does not take advantage of the fact that IDL is a vector language. In our code, `EXOFAST_OCCULTQUAD`, we sped it up by a factor of 100-500 and fixed many bugs.

B.0.1. SPEED ENHANCEMENTS

The majority of the speed enhancement came from a major restructuring of the code to take advantage of IDL's much faster vector operations and careful attention to optimal case ordering.

With the aid of the IDL profiler, we determined the largest remaining bottleneck was in the calculation of the elliptic integral of the third kind, which used translations of the numerical recipes routines `rj`, `rc`, and `rk`. After testing potential

replacements, including IDL's separately-licensed routines `IMSL_ELRC`, `IMSL_ELRJ`, and `IMSL_RLRD`, we replaced the numerical recipes routines with an IDL translation of an ALGOL program published by Burlirsch (1965a,b). This iterative routine is many times faster and more robust than any other alternative we tried.

For our last minor speed enhancement, we combined the routines to calculate Hasting's polynomial approximation for the elliptic integrals of the first ($K(k)$) and second ($E(k)$) kind, which now share the computationally expensive task of computing $\log(1 - k)$. We also tested other replacements for this routine, including Fukushima's piecewise polynomial approximation, (Fukushima 2009), but none was reliably faster.

The calculation of the elliptic integral of the third kind is still the primary bottleneck, but is now comparable to the calculation of the elliptic integrals of the first and second kind, the arc cosine, logarithm, and IDL's `WHERE` function, so additional substantive speed gains will be difficult without a fundamental re-characterization of the problem.

These speed enhancements combined to make this IDL routine ~ 100 times faster than the previous IDL version. Since this calculation is a significant fraction of calculating a transit model, this improves the run time of entire program by a significant factor. We rewrote the Fortran routine with the latter two enhancements,

which is now twice as fast as its predecessor, and 2-5x faster than the current IDL version, depending on the compiler.

Lastly, we include an IDL wrapper for the Fortran version, which uses `CALL_EXTERNAL` and is a drop-in replacement for the strictly IDL version described above. After its first use, it is indistinguishable from the native Fortran in terms of speed. The first use of the IDL session is dramatically slower because it needs to compile the library, though this would likely go unnoticed save for a computer-timed test. This wrapper version is 200-500 times faster than the original IDL version, and 2-5 times faster than the current IDL code, depending on Fortran the compiler used.

The downside to this wrapper is that the syntax required by IDL is technically not legal Fortran. The makefile in the IDL example does not implement `f77` on Linux, likely because of compilation warnings. Some compilers, like `gfortran` and `f95` simply refuse to compile it. It can be compiled with `g77` or `f77`, and the warnings they issue can be safely suppressed. Only `ifort` will compile it without warnings. This `CALL_EXTERNAL` version has only been tested on a 32 bit Linux machine.

B.0.2. BUG FIXES AND CONCEPTUAL ENHANCEMENTS

In addition to being relatively slow, the numerical recipes codes to calculate the elliptic integral of the third kind (`rc`, `rj`, `rk`) were poorly-behaved where values of z were within 10^{-7} of special cases (first, second, third, and fourth contact). Very near

special cases (10^{-13}), the previous codes differed from the true value by as much as 10%. With the Burlirsch (1965a,b) algorithm to calculate the elliptic integral of the third, the differences between the calculated value and true value (calculated with arbitrary precision in Mathematica) are less than 10^{-5} . Since any model that suffered from this bug would be unfairly disfavored, the practical implication of this bug is the potential for a bias in the measured times that pushes data points away from these special cases. For a typical Hot Jupiter, 10^{-7} corresponds to ~ 90 ms in the planetary orbit. Fortunately, this is beyond the current precision attained to date. However, diabolical alignments of several data points may further skew the inferred value.

We also fixed typos in unlikely cases where $p > 1$. While these typos would have caused catastrophic failures, no planetary transit would have suffered from this bug. Furthermore, the behavior of this bug is sufficiently non-physical they would have been immediately found, so were unlikely to have done any harm. We also fixed a handful of other minor bugs, such as consistent use of double precision values and (in Fortran) functions.

The major conceptual enhancements were mentioned previously. Namely, we now allow negative values of p , as discussed in §4.4.4 and we return the coefficients required to analytically fit the quadratic limb darkening parameters as discussed in Appendix A.

Thorough testing at every special case, $\pm 10^{-13}$ of every special case, and for a thousand uniformly spaced values of z between $z = 0$ and $z = 2(1 + p)$ for hundreds of values of p ranging from 10^{-13} to 2 have been checked, and no other errors were found.

APPENDIX C

RANDOM NUMBER GENERATOR

IDL’s help states that `RANDOMU`, their built-in random number generator, is similar to Numerical Recipes `ran1`, which has a periodicity, m , of $\sim 10^9$, and should suffice for a series of random numbers of length, n , such that $n \lesssim m/20 \approx 50$ million.

In our sample fit of transit and RV data with 13 free parameters, we generate 14 random numbers per attempted step (one for each parameter and an additional one to compare to the likelihood ratio), and only accept about 10% of steps. We construct 10 simultaneous chains, and stop when they are well mixed, which roughly corresponds to 100,000 steps per chain. This means that we generate roughly 140 million random numbers – far exceeding the recommended periodicity of the random number generator.

While looking for a suitable replacement, we found that the third edition of Numerical Recipes (Press et al. 2007) warns “be cautious about any source earlier than about 1995, since the field progressed enormously in the following decade.” They also state that one should never use a random number generator with a periodicity of less than $\sim 2 \times 10^{19}$. This certainly excludes `ran1`, and even excludes

their own popularly used alternative `ran2`, published in 1992, with a periodicity of 10^{18} .

Numerical Recipes recommends a specific implementation of a random number generator, which was translated into IDL by Paolo Grigis and is available online as `PG_RAN`¹. This version is ~ 120 times slower than `RANDOMU` because random number generators are serial by nature, and so not ideally implemented in IDL. Its contribution to the overall run time of the program is about 10%, as discussed in Appendix D – a built in random number generator would essentially eliminate this contribution to the overall run time.

We did a fit like the one described above using both `RANDOMU`, `PG_RAN`, and a translation of `ran2` and saw no statistically significant difference between any of the fits, so this may be overkill. However, it has been shown numerous times that the quality of the random number generator can affect the scientific conclusions (e.g., Kalle & Wansleben 1984; Parisi & Rapuano 1985; Filk et al. 1985; Grassberger 1993), and we feel that saving $\sim 10\%$ on the execution time is not worth the risk. We therefore use `PG_RAN` by default and recommend it for general use.

For integration with our MCMC fitter described above, we provide a wrapper, `EXOFAST_RANDOM` for `PG_RAN`, which makes it a drop in replacement for `RANDOMU` when uniform or Gaussian deviates are desired. In order to do this, we added the

¹http://hea-www.cfa.harvard.edu/~pgrigis/idl_stuff/pg_ran.pro

functionality to generate random Gaussian deviates with a translation of Numerical Recipes' `gasdev` using `PG_RAN` for its uniform random inputs. If more speed is required, a simple flag can switch between the two generators or any user-written generator, as long as it has the same calling structure and the ability to return 3 dimensional uniform and Gaussian deviates.

APPENDIX D

PERFORMANCE

We ran a typical case of a fit with both transit and RV data, which was well-mixed in about 2.5 hours. During this time, we ran IDL's profiler to analyze the run time performance to see which of our routines may be a bottleneck. While the particular mix will necessarily depend on the number of data points and the details of the model, it is instructive to look at this case study in depth.

The solution to Kepler's equation is by far the largest remaining bottleneck, consuming 22% of the total computation time. Since this iterative solution is necessarily serial, re-writing it in Fortran or C and calling it via `CALL_EXTERNAL` may be a quick way of improving the run time. However, for low-eccentricity orbits, the number of iterations is small.

At 14% the random number generator takes the next biggest slice. As mentioned in Appendix C, this routine is 120x slower than its built-in counterpart, owing to its (necessarily) serial algorithm. With an optimized routine, this portion could become essentially negligible. We have made provisions to replace this with

faster routines relatively easily, but re-iterate our warning that faulty random number generators can qualitatively change the results.

The next largest contributors (8% each) are `SIN` and `COS`, which are highly optimized, built-in routines with comparable speed to low-level languages. The fact that these routines are the among the largest contributors is a sign that the code was written efficiently, and sets a floor for the likely level of improvement possible – even if translated to a lower level language. The total contributions from these and other fundamental functions (e.g., `TAN`, `SQRT`) is 22%, making substantive improvement difficult.

The next largest contributor is the conversion to the target reference frame (6.1%). It is also an iterative function, and is calculated twice at each step in the MCMC chain – one for RV and one for photometry. It is important to know that this routine calls `EXOFAST_KEPLEREQ`, `SIN`, `COS`, `TAN`, and `ATAN`, and that their contributions are not included in the 6.1% figure (so as not to double count their contribution). The total run time of this calculation, including their contributions, is 41%. This is a hefty price to pay to include an effect that is nearly negligible, and this could be skipped altogether (as has been done to date) without any significant difference. Analyzing the run time excluding this routine does not qualitatively change the top contributors.

The last major contributor, at 4.8%, is the calculation of the transit light curve (using the all-IDL version). Our efforts to optimize this calculation are described in detail in Appendix B. Like the conversion to the target reference frame, this figure counts the time only spent inside the routine. The total contribution of this routine is 13%. Had we left the routine alone, the total run time of the MCMC fit would have been completely dominated by this calculation, taking 30 hours by itself.

Using our comparison between `EXOFAST_OCCULTQUAD` and its Fortran counterpart as a guide, we can guess that an all-Fortran version would probably be around 5 times faster. Further, considering the execution time and relative contribution of the Fortran version of `EXOFAST_OCCULTQUAD`, an improvement of greater than 50 times is exceedingly unlikely.

Perhaps the easiest way of substantively improving the run time is to step more efficiently, which would mean a higher acceptance rate (we were 2x below optimal $\sim 23\%$) and therefore fewer model evaluations and faster convergence time. One obvious method is to step along the Eigen vectors (i.e., the uncorrelated parameter set), but we caution that this would impose unphysical priors which may be difficult to correct (e.g., see §4.3.2).

BIBLIOGRAPHY

- Agol, E., Steffen, J., Sari, R., & Clarkson, W. 2005, MNRAS, 359, 567
- Agol, E., & Steffen, J. H. 2007, MNRAS, 374, 941
- Alard, C. 2000, A&AS, 144, 363
- Alard, C., & Lupton, R. H. 1998, ApJ, 503, 325
- Alonso, R., et al. 2009, A&A, 506, 353
- Arias, E. F., Charlot, P., Feissel, M., & Lestrade, J. 1995, A&A, 303, 604
- Barnes, J. W., & Fortney, J. J. 2004, ApJ, 616, 1193
- Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
- Binnendijk, L. 1960, Properties of double stars; a survey of parallaxes and orbits.,
ed. Binnendijk, L.
- Brown, T. M., Charbonneau, D., Gilliland, R. L., Noyes, R. W., & Burrows, A.
2001, ApJ, 552, 699
- Bundy, K. A., & Marcy, G. W. 2000, PASP, 112, 1421

- Burlirsch, R. 1965a, *Numerische Mathematik*, 7, 78
- . 1965b, *Numerische Mathematik*, 7, 353
- Carter, J. A., Yee, J. C., Eastman, J., Gaudi, B. S., & Winn, J. N. 2008, *ApJ*, 689, 499
- Charbonneau, D., Brown, T. M., Burrows, A., & Laughlin, G. 2007, *Protostars and Planets V*, 701
- Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, *ApJ*, 529, L45
- Charbonneau, D., Brown, T. M., Noyes, R. W., & Gilliland, R. L. 2002, *ApJ*, 568, 377
- Charbonneau, D., et al. 2005, *ApJ*, 626, 523
- . 2006, *ApJ*, 636, 445
- Chromey, F. R., & Hasselbacher, D. A. 1996, *PASP*, 108, 944
- Claret, A., & Bloemen, S. 2011, *A&A*, 529, A75+
- Deming, D., Harrington, J., Seager, S., & Richardson, L. J. 2006, *ApJ*, 644, 560
- Deming, D., Seager, S., Richardson, L. J., & Harrington, J. 2005, *Nature*, 434, 740
- Duffett-Smith, P. 1989, *Practical Astronomy with your Calculator*, ed. Duffett-Smith, P.

- Eastman, J., Gaudi, B. S., Siverd, R., Trueblood, M., & Trueblood, P. 2010a, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7733, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
- Eastman, J., Siverd, R., & Gaudi, B. S. 2010b, PASP, 122, 935
- Edwards, R. T., Hobbs, G. B., & Manchester, R. N. 2006, MNRAS, 372, 1549
- Enoch, B., Collier Cameron, A., Parley, N. R., & Hebb, L. 2010, A&A, 516, A33+
- Euchner, F., Jordan, S., Beuermann, K., Reinsch, K., & Gänsicke, B. T. 2006, A&A, 451, 671
- Fairhead, L., & Bretagnon, P. 1990, A&A, 229, 240
- Filk, T., Marcu, M., & Fredenhagen, K. 1985, Physics Letters B, 165, 125
- Fleming, S. W., et al. 2010, ApJ, 718, 1186
- Ford, E. B. 2005, AJ, 129, 1706
- . 2006, ApJ, 642, 505
- Ford, E. B., & Gaudi, B. S. 2006, ApJ, 652, L137
- Ford, E. B., & Holman, M. J. 2007, ApJ, 664, L51
- Fortney, J. J., Saumon, D., Marley, M. S., Lodders, K., & Freedman, R. S. 2006, ApJ, 642, 495

- Fukushima, T. 2009, *Celestial Mechanics and Dynamical Astronomy*, 105, 305
- Gaudi, B. S., & Winn, J. N. 2007, *ApJ*, 655, 550
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. 2003, *Bayesian Data Analysis*, 2nd edn.
- Gould, A. 2003, *ArXiv Astrophysics e-prints*
- Gould, A., et al. 2009, *ApJ*, 698, L147
- Grassberger, P. 1993, *Physics Letters A*, 181, 43
- Gregory, P. C. 2005, *ApJ*, 631, 1198
- Guillot, T. 2005, *Annual Review of Earth and Planetary Sciences*, 33, 493
- Henden, A. A., & Kaitchuck, R. H. 1982, *Astronomical photometry*, ed. Henden, A. A. & Kaitchuck, R. H.
- Henry, G. W., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2000, *ApJ*, 529, L41
- Heyl, J. S., & Gladman, B. J. 2007, *MNRAS*, 377, 1511
- Hirshfeld, A., & Sinnott, R. W. 1997, *Sky catalogue 2000.0. Volume 2: Double stars, variable stars and nonstellar objects.*, ed. Hirshfeld, A. & Sinnott, R. W.
- Hobbs, G. B., Edwards, R. T., & Manchester, R. N. 2006, *MNRAS*, 369, 655
- Holman, M. J., & Murray, N. W. 2005, *Science*, 307, 1288

- Irwin, A. W., & Fukushima, T. 1999, *A&A*, 348, 642
- Jordán, A., & Bakos, G. Á. 2008, *ApJ*, 685, 543
- Kalle, C., & Wansleben, S. 1984, *Computer Physics Communications*, 33, 343
- Kipping, D. M. 2009, *MNRAS*, 396, 1797
- Kipping, D. M., & Bakos, G. Á. 2010, *arXiv:1004.3538v1*
- Klinglesmith, D. A., & Sobieski, S. 1970, *AJ*, 75, 175
- Knutson, H. A., et al. 2007, *Nature*, 447, 183
- Koch, D. G., et al. 2010, *ApJ*, 713, L79
- Kovács, G., Zucker, S., & Mazeh, T. 2002, *A&A*, 391, 369
- Landolt, A. U., & Blondeau, K. L. 1972, *PASP*, 84, 784
- Landsman, W. B. 1993, in *Astronomical Society of the Pacific Conference Series*,
Vol. 52, *Astronomical Data Analysis Software and Systems II*, ed. R. J. Hanisch,
R. J. V. Brissenden, & J. Barnes, 246–+
- Lang, D., Hogg, D. W., Mierle, K., Blanton, M., & Roweis, S. 2010, *AJ*, 139, 1782
- Laughlin, G., Marcy, G. W., Vogt, S. S., Fischer, D. A., & Butler, R. P. 2005, *ApJ*,
629, L121

- Lee, B. L., et al. 2010, in *Bulletin of the American Astronomical Society*, Vol. 41,
 Bulletin of the American Astronomical Society, 519–+
- Levenberg, K. 1944, *Quarterly Journal of Applied Mathematics*, II, 164
- Loeb, A. 2005, *ApJ*, 623, L45
- Lomb, N. R. 1976, *Ap&SS*, 39, 447
- Lorimer, D. R. 2008, *Living Reviews in Relativity*, 11, 8
- Lutz, T. E., & Kelker, D. H. 1973, *PASP*, 85, 573
- Mandel, K., & Agol, E. 2002, *ApJ*, 580, L171
- Marquardt, D. W. 1963, *Siam Journal on Applied Mathematics*, 11
- Mayor, M., & Queloz, D. 1995, *Nature*, 378, 355
- Mazeh, T., Tamuz, O., & Zucker, S. 2007, in *Astronomical Society of the Pacific
 Conference Series*, Vol. 366, *Transiting Extrapolar Planets Workshop*, ed.
 C. Afonso, D. Weldrake, & T. Henning, 119–+
- Mikkola, S. 1987, *Celestial Mechanics*, 40, 329
- Miralda-Escudé, J. 2002, *ApJ*, 564, 1019
- Nelder, J. A., & Mead, R. 1965, *The Computer Journal*, 7, 308
- Pál, A., & Kocsis, B. 2008, *MNRAS*, 389, 191

- Parisi, G., & Rapuano, F. 1985, *Physics Letters B*, 157, 301
- Pepper, J., et al. 2007, *PASP*, 119, 923
- Pont, F., et al. 2007, *A&A*, 476, 1347
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007, *Numerical recipes : the art of scientific computing*, 3rd edn., ed. Press, W. H.
- Queloz, D., Eggenberger, A., Mayor, M., Perrier, C., Beuzit, J. L., Naef, D., Sivan, J. P., & Udry, S. 2000, *A&A*, 359, L13
- Rafikov, R. R. 2009, *ApJ*, 700, 965
- Sasselov, D. D. 2003, *ApJ*, 596, 1327
- Sato, B., et al. 2005, *ApJ*, 633, 465
- Scargle, J. D. 1982, *ApJ*, 263, 835
- Seager, S., & Hui, L. 2002, *ApJ*, 574, 1004
- Seager, S., & Mallén-Ornelas, G. 2003, *ApJ*, 585, 1038
- Seidelmann, P. K. 1992, *Explanatory Supplement to the Astronomical Almanac*, ed. Seidelmann, P. K. (University Science Books)
- Shapiro, I. I. 1964, *Physical Review Letters*, 13, 789
- Standish, E. M. 1998, *A&A*, 336, 381

- Steffen, J. H., & Agol, E. 2005, MNRAS, 364, L96
- Torres, G., Andersen, J., & Giménez, A. 2010, A&A Rev., 18, 67
- Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J., Ballester, G. E., Ferlet, R., Hébrard, G., & Mayor, M. 2003, Nature, 422, 143
- Winn, J. N. 2009, in IAU Symposium, Vol. 253, IAU Symposium, 99–109
- Winn, J. N. 2010, ArXiv e-prints
- Winn, J. N., et al. 2005, ApJ, 631, 1215
- . 2009, ApJ, 703, 2091
- . 2010, ApJ, 718, 575
- Wright, J. T., et al. 2011, PASP, 123, 412
- Yee, J. C., et al. 2009, ApJ, 703, 2082