Title

Dynamical Masses for Brown Dwarf and Giant Planet Companions to Nearby Stars

Permalink

https://escholarship.org/uc/item/74s795s2

Author

Brandt, Gregory Mirek

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY of CALIFORNIA
Santa Barbara

Dynamical Masses for Brown Dwarf and Giant Planet Companions to Nearby Stars

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Physics

by

G. Mirek Brandt

Committee in charge:

Professor Timothy D. Brandt, Chair

Professor Omer Blaes

Professor Lars Bildsten

December 2022

The dissertation of G. Mirek Brandt is approved:

Professor Omer Blaes

Professor Lars Bildsten

Professor Timothy D. Brandt, Chair

September 2022

Dynamical Masses for Brown Dwarf and Giant Planet Companions to Nearby Stars

by

G. Mirek Brandt

To my family, including Joe and Linda Clark, and of course, Copper.

Acknowledgements

Personal Acknowledgements:

I thank Professor Timothy Brandt for his close guidance, patience, and mentorship. I am grateful he took me on as his student. He has always encouraged me to do my best, yet in a way that also prioritized my personal well-being. He created an immensely welcoming and nurturing research environment. I am overwhelmingly lucky to have been advised by Dr. Brandt.

I thank Professor Lars Bildsten for opening the doors to his group meetings to me, where I learned a vast amount of stellar physics. I thank him immensely for mentoring me for nearly a year, following a summer Paxton fellowship. I thank Bill Paxton for his generous support for the fellowship.

I am extremely grateful for the support of the National Science Foundation (NSF) Graduate Research Fellowship (grant no. 1650114). The support of the NSF was integral to the work both in this thesis and my Ph.D. as a whole.

I thank my collaborators over the years. The work in this thesis, and my published works, would not be what they are without the generous contributions, insight, and time of many wonderful collaborators. In no particular order, I thank Gabriel-Dominique Marleau, Daniel Michalik, Minghan Chen, Trent J. Dupuy, Sunny Wong, Curtis McCully, Marshall Johnson, Thayne Currie, Brendan P. Bowler, Yiting Li, Zhexing Li, and Kyle Fransen.

I thank Josiah Schwab and Emily Martin for immensely helpful career advice leading up to the end of my graduate school experience.

I owe a monumental thanks to Las Cumbres Observatory for hosting me as an intern early in my graduate school career, and for providing such an amazing environment and opportunity to learn software development. I appreciate immensely the free espresso from their lounge that I liberally took advantage of. I cannot understate the impact on my career that LCO provided through the opportunity to co-develop the NRES data analysis pipeline. I give immense gratitude to Curtis McCully, Lisa Storrie-Lombardi, Todd Henderson, Daniel Harbeck, Tim Brown, Marshall Johnson, and Sarah Rettinger. I want to give a special thanks to Curtis McCully for patiently teaching me and building my foundation of software engineering, and for pushing me to always write better code.

Finally, I thank my friends, both old friendships and those that formed newly during graduate school. Thank you.

Technical Acknowledgements and Attributions: The work in this thesis made use of data from the European Space Agency (ESA) mission *Gaia* (https://www.cosmos.esa.int/Gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/Gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

I wish to thank the teams behind the calibrated HIRES (Tal-Or et al., 2019) and HARPS (Trifonov et al., 2020) catalogs. These catalogs were immensely helpful to the work in this thesis.

The work in this thesis made use of the following Python packages: astropy (Astropy Collaboration et al., 2013; Price-Whelan et al., 2018), scipy (Virtanen et al., 2020), numpy (Oliphant, 2006; van der Walt et al., 2011), pandas (Wes McKinney, 2010; pandas development team, 2020), htof (Brandt et al., 2021c; Brandt & Michalik, 2020), `orvara` (Brandt et al., 2021f), and REBOUND (Rein & Liu, 2012).

Many of the orbits presented in this work can be found at whereistheplanet.com. I thank Jason Wang, Matas Kulikauskas, and Sarah Blunt for building and hosting this wonderful web service, and for making their orbital posteriors open and freely available.

Curriculum Vitæ

G. Mirek Brandt

Education

2022	Ph.D., Physics, University of California, Santa Barbara, CA
2020	M.A., Physics, University of California, Santa Barbara, CA
2018	B.S., Physics, Oregon State University, Corvallis, OR
2018	B.S., Mathematics, Oregon State University, Corvallis, OR

Select Publications

Statistical properties of Hipparcos 2, caveats on its use, and a recalibration of the Intermediate Astrometric Data
G. Mirek Brandt, Daniel Michalik, Timothy D. Brandt
Submitted to RASTI, Aug 2022.

htof: A new open-source tool for analyzing Hipparcos, Gaia, and future astrometric missions
G. Mirek Brandt, Daniel Michalik, Timothy D. Brandt, Yiting Li, Trent J. Dupuy, Yunlin Zeng
The Astronomical Journal, 162, 230 (2021)

Improved Dynamical Masses for Six Brown Dwarf Companions Using Hipparcos and Gaia EDR3
G. Mirek Brandt, Trent J. Dupuy, Yiting Li, Minghan Chen, Timothy D. Brandt, Tin Long Sunny Wong, Thayne Currie, Brendan P. Bowler +3 co-authors
The Astronomical Journal, 162, 301 (2021)

The First Dynamical Mass Measurement in the HR 8799 System
G. Mirek Brandt, Timothy D. Brandt, Trent J. Dupuy, Daniel Michalik, Gabriel-Dominique Marleau
The Astrophysical Journal, 915, L16 (2021)

Precise Dynamical Masses and Orbital Fits for β Pic b and β Pic c
G. Mirek Brandt, Timothy D. Brandt, Trent J. Dupuy, Yiting Li, Daniel Michalik
The Astronomical Journal, 161, 179 (2021)

Automatic Echelle Spectrograph Wavelength Calibration
G. M. Brandt, T. D. Brandt, C. McCully
The Astronomical Journal, 160, 25 (2020)

Dynamical Mass of the Young Substellar Companion HD 984 B
Kyle Franson, Brendan P. Bowler, Timothy D. Brandt, Trent J. Dupuy, Quang H. Tran, **G. Mirek Brandt**, Yiting Li, Adam L. Kraus
The Astronomical Journal, 163, 50 (2022)

Limits on the Mass and Initial Entropy of 51 Eri b from Gaia EDR3 Astrometry
Trent J. Dupuy, **G. Mirek Brandt**, Timothy D. Brandt
Monthly Notices of the Royal Astronomical Society, 509, 4411 (2022)

The Gliese 86 Binary System: A Warm Jupiter Formed in a Disk Truncated at ≈2 AU
Yunlin Zeng, Timothy D. Brandt, Gongjie Li, Trent J. Dupuy, Yiting Li, **G. Mirek Brandt**, Jay Farihi, Jonathan Horner +7 co-authors
[arXiv:2112.06394]

Precise Masses and Orbits for Nine Radial Velocity Exoplanets
Yiting Li, Timothy D. Brandt, **G. Mirek Brandt**, Trent J. Dupuy, Daniel Michalik, Rebecca Jensen-Clem, Yunlin Zeng, Jacqueline Faherty +1 co-author
The Astronomical Journal, 162, 266 (2021)

14 Her: a likely case of planet-planet scattering
Daniella C. Bardalez Gagliuffi, Jacqueline K. Faherty, Yiting Li, Timothy D. Brandt, Lauryn Williams, **G. Mirek Brandt**, Christopher R. Gelino
The Astrophysical Journal, 922, L43 (2021)

orvara: An Efficient Code to Fit Orbits using Radial Velocity, Absolute, and/or Relative Astrometry
Timothy D. Brandt, Trent J. Dupuy, Yiting Li, **G. Mirek Brandt**, Yunlin Zeng, Daniel Michalik, Daniella C. Bardalez Gagliuffi, Virginia Raposo-Pulido
The Astronomical Journal, 162, 186 (2021)

A Dynamical Mass of 70 ± 5 *Jupiter Masses for Gliese 229B, the First T Dwarf*
Timothy D. Brandt, Trent J. Dupuy, Brendan P. Bowler, Daniella C. Bardalez Gagliuffi, Jacqueline Faherty, **G. Mirek Brandt**, Daniel Michalik
The Astronomical Journal, 160, 196 (2020)

Abstract

Dynamical Masses for Brown Dwarf and Giant Planet Companions to Nearby Stars

by

G. Mirek Brandt

Brown dwarfs are sub-stellar objects that span the enormous range in mass from $\sim 15 M_{\mathrm{Jup}}$ to $\sim 80 M_{\mathrm{Jup}}$. The internal physics and temporal evolution of these objects are remarkably similar across the entire range of mass, albeit massive brown dwarfs cool to a given luminosity more slowly than lighter ones. Giant planets (masses between ~ 1 and $\sim 15 M_{\mathrm{Jup}}$) behave very similarly. Electron degeneracy causes these objects to all have similar radii, almost independent of both mass and temperature. These objects do not fuse hydrogen. Thus, the luminosity of a brown dwarf simply decreases over time due to the steady radiative loss of the brown dwarf's internal energy. Most observational evidence points to these objects forming "hot" (so-called hot-starts), whereby most the energy of their formation is retained and so their internal energy is set at birth and is solely a function of their initial mass, much like how the number of grains in an hourglass is fixed. Under the hourglass analogy, the mass is the total amount of sand trapped the hourglass and the observed luminosity is akin to the number of grains remaining in

the top reservoir. Adding in knowledge of the energy-loss (sand-grain) rate from cooling models, one can calculate the age of the companion (i.e., when the hourglass was flipped). Thus, brown dwarfs have a unique advantage compared to e.g., rocky planets. If one has a precisely measured luminosity and a stellar age, one can predict from the cooling-models what the mass of that object ought to be. Conversely, if one has a precisely measured mass and luminosity, one can determine when the brown dwarf was born and hence age date the host system and every planet therein. Unfortunately, both those above scenarios are subject to uncertainties in, and slight differences between, the various models of brown dwarf evolution. The goal is not to use cooling models to infer the third quantity when two are known, but rather to measure all three: the mass, luminosity, and the age of the brown dwarf, *independent* of evolutionary models. Then one can actually test and inform models of brown dwarf evolution. The work in this thesis moves towards that end goal: expanding the sample of brown dwarfs and giant planets where all three quantities are measured so that we may advance our understanding of sub-stellar evolution across the entire mass range of $\sim 1M_{\rm Jup}$ to $\sim 80M_{\rm Jup}$.

In the decade preceding the launch of the *Gaia* satellite, directly imaged brown dwarfs and giant planets (which orbited main-sequence stars) were rare. Of order twenty were known, and only a small fraction had an independently measured mass. Masses measured via observing the orbit of the companion are referred to as dynamical masses, because the mass measurement relies only on Newtonian dynamics. These are the gold-standard of independent mass measurements ("independent" because they do not rely on cooling/evolutionary models).

In order to: 1. test models of brown dwarf and giant planet evolution; 2. provide independent age estimates of a stellar system; and 3. test models of brown dwarf atmospheres; one needs not only more brown dwarfs, but a dynamical mass for each. On one front, *Gaia* enabled the discovery and subsequent confirmation of many more brown dwarfs companions (of-order 100 in only 4 years; by 2025, *Gaia* alone is expected to detect 20,000-70,000 companions across a broad range of mass). The work in this thesis has worked precisely towards the second front of producing precise mass measurements, independent ages, and tests of brown dwarf evolutionary models. The key is leveraging the synergy between the *Hipparcos* and *Gaia* astrometric missions, which allows us to measure a precise mass for companions around stars by measuring the reflex motion of the star in the sky as the unseen, orbiting companion tugs the star.

In Chapter 2, we resolve a tension between the dynamical mass of the giant planet β Pictoris b and the model-predicted mass based on its luminosity and host-star age. We showcase how one can combine the radial-velocities of the star, direct imaging, and *Hipparcos* and *Gaia* reflex motion (via the Hipparcos-Gaia Catalog of Accelerations), to obtain an excellent dynamical mass measurement. We utilize the Markov-Chain-Monte-Carlo code, `orvara`, and the astrometry parsing code `htof`, to constrain orbits using those three separate sources of data. In Chapter 3, we use those same methods to obtain the mass, for the first time ever, of the planet HR 8799 e. From the mass, we obtain a measure of the planet's age and thereby the age of the star. HR 8799 is a star whose cluster membership is uncertain, and so an independently measured age is one way to finally assign the birthplace of the HR 8799 system. In Chapter 4, we use all the aforementioned techniques to update the masses and orbits of six brown dwarfs, including the cornerstone and puzzling “over-massive” system, Gl 229. This thesis concludes with a detailed recalibration of *Hipparcos*, in Chapter 5, to enable better mass measurements in future works. In addition, Chapter 5 addresses a long standing question surrounding overfitting in the second *Hipparcos* data reduction.

Contents

Chapter 1

Introduction

1.1 Giant Planets and Brown Dwarfs

Brown dwarfs (BDs) are substellar objects with masses below the hydrogen-fusion mass limit of 75–80 $M_{\rm Jup}$ (Burrows et al., 2001; Dupuy & Liu, 2017). Sufficiently massive to fuse deuterium but not hydrogen, they cool as they age. A BD has a convective interior coupled to an atmosphere that contains chemically diverse clouds with detailed interactions and opacities (Marley & Robinson, 2015). The atmosphere modulates the BD's cooling, affecting its present-day spectrum, effective temperature, and luminosity (e.g., Saumon & Marley, 2008).

A rich variety of atmospheric and evolutionary models have been constructed that predict the radii, spectra, and luminosities of BDs as functions of their age and composition (e.g., Allard et al. 2001a; Baraffe et al. 2003; Saumon & Marley 2008; Spiegel & Burrows 2012; Phillips et al. 2020). BDs all have similar Jupiter-sized radii after initial contraction finishes. However, a fundamental degeneracy exists whereby older and more massive BDs can have similar temperatures and luminosities to younger and less massive BDs (Bildsten et al., 1997; Marleau & Cumming, 2014). This degeneracy between age, luminosity, and mass has to be broken to test evolutionary models. Possessing all three allows one to constrain BD properties and/or the physics of their cooling. Independent measures of planet age, luminosity, and mass for young ($\lesssim$500 Myr) giant planets, like those in β Pictoris or HR 8799, allows one to potentially constrain their initial entropy at formation (Marley et al., 2007a; Marleau & Cumming, 2014).

Direct-imaging instruments such as Subaru/CHARIS (Groff et al., 2013, 2015), VLT and SPHERE (Beuzit et al., 2019), Gemini/GPI (Macintosh et al., 2014), Keck/NIRC2 (McLean & Sprayberry, 2003; Johansson et al., 2008), and recently interferometers like VLT/GRAVITY (e.g., Lagrange et al., 2020; Nowak et al., 2020) allow the measurement of spectra and luminosities with newfound precision for BD companions to nearby stars. Because their host stars are bright and nearby, they often have well-measured metallicities and age indicators. However, these BDs are typically on wide orbits with long orbital periods. Radial velocity (RV) time series typically cover a small fraction of the orbit. Thus, the most difficult quantity to measure is usually the mass of the BD.

Masses based only on Newton's laws (dynamical masses) allow us to calibrate and benchmark models of BD and sub-stellar evolution. Brown dwarf modeling couples a simple, fully convective interior model to a complex atmosphere with clouds, hazes, and molecular chemistry (Burrows et al., 1997; Ackerman & Marley, 2001). The ability of model spectra and luminosities to reproduce observations depends on the correct treatment of opacity sources, and on the initial thermal state of the brown dwarf (Allard et al., 2001b; Saumon & Marley, 2008). These brown dwarf models are the high-mass versions of the models used to infer the properties, and formation conditions, of giant exoplanets (Marley & Robinson, 2015).

For all brown dwarfs and giant planets, combining a well-measured age with an independently measured mass and luminosity can constrain the companion's initial supply of thermal energy and provide clues to its formation mechanism (Marley et al., 2007b; Fortney et al., 2008; Spiegel & Burrows, 2012; Marleau & Cumming, 2014). The age of the companion typically comes from knowledge of the host-star age, e.g., from gychronology (Soderblom, 2010; Ahuir et al., 2020) or kinematic traceback (e.g., with β Pictoris or HR 8799; Nielsen et al., 2014; Marois et al., 2008). The luminosity of the companion comes directly from high-contrast images of the system. However, an independently measured mass was, until now, elusive.

Until the *Gaia* astrometric mission, independently measured masses were rare for largely three reasons. First was because of the low-yield rate of blind direct imaging surveys (Brandt, 2018). Second, is the fundamental mismatch between companions that are easiest to image, and those where we have radial velocities (RVs) across the entire orbit. Directly imaged companions are preferentially on wide, long-period orbits (because it is easier to image them). Thus it is extremely difficult to have coincidentally had ~ 30 year (or longer) RV coverage, covering that entire orbit. Absolute astrometry can break the need for full-orbital coverage of RVs, but the only existing space-astrometry mission was *Hipparcos* (ESA, 1997; van Leeuwen, 2007a) (launched in 1989, spanning only 4 years); and it lacked the precision and time-baseline that is needed for most long-period companions. This limited precise mass measurements to a narrow range of orbital periods and to the high-mass regime only (e.g., see Sahlmann et al., 2011), or to the few brown dwarfs where absolute astrometry was not needed: those whose orbital separation was large enough to be imaged yet also small enough such that RVs covered a significant fraction of the orbit (e.g., HR 7672; Crepp et al., 2012a).

1.2 Dynamical Mass Measurements of Brown Dwarfs and Giant Planets

Dynamical mass measurements of $\sim 10 M_{\rm Jup}$ objects on mid-length (~5 year) to long (~100-300 year) orbits became feasible thanks to absolute astrometry from the *Gaia* (Gaia Collaboration et al., 2016; Lindegren et al., 2018) mission, which began operation in 2016. Tantamount to this was the act of coupling the *Gaia* astrometry with the 25-year-old preceding mission, *Hipparcos*. This endeavour was spear-headed by Brandt (2018), Kervella et al. (2019), and Fontanive et al. (2019). The Brandt (2018) realization of this endeavour was the Hipparcos-Gaia Catalog of Accelerations (hereafter, HGCA, Brandt, 2018, 2021). The catalog gives extremely high precision measurements of the acceleration of the host star in the plane of the sky. The magnitude of that acceleration is, in-part, proportional to ratio of the companion mass to the mass of the star. This additional constraint allowed one to finally solve precisely for the masses of brown dwarfs whose RV and direct imaging covered only a small fraction of the orbit. The HGCA also extended the time-baseline of absolute astrometry to 25-years (a factor of ∼6 larger than either *Hipparcos* or *Gaia* on their own). The extended time baseline and improved astrometric precision solved the three major problems outlined in the previous section. Long period (e.g., over 200 years in the case of Gl 229 B; Brandt et al., 2021e) and light-weight (less than $\sim 10 M_{\rm Jup}$) companions could now have highly precise dynamical masses derived; and direct-imaging surveys turned their eyes preferentially to systems that were accelerating in the HGCA, greatly boosting the yield of newly discovered planets (e.g., Currie et al., 2021)

Since *Gaia*'s second data release and the catalog of Brandt (2018), a number of authors derived masses and orbits for a litany of nearby brown dwarfs and giant planets (Brandt, 2018; Brandt et al., 2018; Calissendorff & Janson, 2018; Dupuy et al., 2019a; Fontanive et al., 2019; Kervella et al., 2019; Feng et al., 2019; Maire et al., 2020a; Currie et al., 2020; Chen et al., 2022). The work presented in this thesis is towards that same end-goal. This thesis can be divided into two distinct, yet related, endeavours. One is measuring the dynamical masses of more brown dwarfs and giant planets, in order to constrain and understand formation scenarios. The other avenue is to improve the precision of *Hipparcos*, and fold these improvements back into mass and orbit studies of brown dwarfs to improve mass constraints.

1.2.1 Revised Orbits and Masses for the β Pictoris Planets

β Pictoris b (β Pic b) is a massive planet ($\sim 10 M_{\mathrm{Jup}}$; Snellen & Brown, 2018; Dupuy et al., 2019a; Nielsen et al., 2020) orbiting the nearby star β Pic A (parallax 51.44 mas). β Pic b is one companion that is of special importance for calibrating and understanding planet formation. It was among the first exoplanets to be directly imaged (Lagrange et al., 2010). Since then, it has been observed dozens of times, resulting in photometry spanning the near-infrared (Quanz et al., 2010; Bonnefoy et al., 2011; Currie et al., 2011; Bonnefoy et al., 2013; Males et al., 2014), low-resolution spectroscopy (Chilcote et al., 2015, 2017), and even medium-resolution spectroscopy (Snellen et al., 2014; Gravity Collaboration et al., 2020). Part of the system's importance derives from the well-measured age of β Pic A. The host star is the defining, highest-mass member of the β Pictoris moving group (Barrado y Navascués et al., 1999; Zuckerman et al., 2001).

In Chapter 2, we resolve a recent observational tension in the β Pic system. The tension arose when the picture was complicated and enriched by the discovery of a second companion, β Pic c, orbiting roughly 3 AU from the host star (Lagrange et al., 2019a; Nowak et al., 2020), interior to the $\approx$10 AU orbit of β Pic b. Nowak et al. (2020) and Lagrange et al. (2020) fit a two-planet Keplerian model to the β Pic system. When these authors adopted an uninformative prior on the mass of β Pic b, their best-fit dynamical mass measurements were 3.2 M_{Jup} (Lagrange et al., 2020) and $5.6 \pm 1.5 M_{\mathrm{Jup}}$ (Nowak et al., 2020). As the authors noted, such low masses are incompatible with cooling models. Cooling models predict much higher masses that are needed to produce the observed flux (e.g., Baraffe et al. 2003; Spiegel & Burrows 2012). Nowak et al. (2020) ultimately adopted a prior of $15 \pm 3\ M_{\mathrm{Jup}}$ while Lagrange et al. (2020) used a prior of $14 \pm 1\ M_{\mathrm{Jup}}$. With these priors, the posterior masses are shifted to near $\sim$10 M_{Jup}. The necessity to use such an informative prior indicates a tension between the dynamical constraints and model predictions from spectral analyses, as noted by Nowak et al. (2020). In Chapter 2, we present precise masses and orbits of β Pic b and β Pic c without the need for informative priors on the planets' masses. Our inferred masses are compatible with a range of cooling model predictions and incorporate the new GRAVITY relative astrometry.

1.2.2 The First Dynamical Mass Measurement in the Four-Planet System HR 8799

The HR 8799 system is most commonly thought to be an $\approx$ 40-Myr-old member of the Columba association (Zuckerman et al., 2011), but Lee & Song (2019) recently suggested membership with the younger β Pic Moving Group (BPMG). Given a $\approx$40-Myr age, hot-start evolutionary models predict masses of $7{\pm}2\,M_{\mathrm{Jup}}$ for the innermost three planets, well below the deuterium-fusion mass boundary ($\approx$12-13 M_{Jup}; Spiegel et al., 2011). Lower masses ($\lesssim$5–7 M_{Jup}) improve the system's dynamical stability (Fabrycky & Murray-Clay, 2010; Sudol & Haghighipour, 2012), but resonant locking could render the system stable at higher planet masses (Götberg et al., 2016; Wang et al., 2018; Goździewski & Migaszewski, 2018, 2020). Higher masses would imply either an older age for the system or entropy loss during formation (a colder start), although there is a limit to the entropy that can be lost at formation (Marleau & Cumming, 2014). Conversely, lower masses would suggest a younger age and therefore favor membership with the BPMG. Dynamical mass measurements can conclusively test such hypotheses. The inferred atmospheric properties and chemical abundances of the planets also depend on their assumed masses (e.g., Wang et al., 2020; Mollière et al., 2020). In Chapter 3, we derive for the first time, the mass of the inner planet HR 8799 e. We make only mild assumptions about the mass ratio relative to the other planets, so the mass measurement is largely anchored in dynamics alone.

1.2.3 Improved Dynamical Masses for Six Brown Dwarf Companions

The recent *Gaia* data release, *Gaia* EDR3, yields proper motions that are on average a factor of 3–4 times more precise than those from *Gaia* DR2 (Gaia Collaboration et al., 2021; Lindegren et al., 2020; Brandt, 2021). This precision improvement allows for even stronger mass constraints for most directly imaged sources. In Chapter 4, we use the improved astrometry from *Gaia* EDR3, as published and calibrated in the EDR3 version of the *Hipparcos*-*Gaia* Catalog of Accelerations (HGCA; Brandt 2021), to produce the most precise orbits and companion masses to date for six systems: Gl 229, Gl 758, HD 13924, HD 19467, HD 33632, and HD 72946. We use the Markov-Chain Monte-Carlo (MCMC) orbit-fitting code `orvara` (Brandt et al., 2021f).

1.3 Recalibrating Past Astrometry to Improve Future Mass Measurements

Astrometric missions like *Hipparcos* or *Gaia* measure absolute positions, parallaxes, and motions for the stars that they survey (Perryman, 2012). These missions observe a star many times and fit a model to that star's motion in the plane of the sky, also known as the star's ephemeris. A star's apparent motion is most simply modeled as parallactic motion (motion induced by the earth's orbit) together with a constant velocity; encoded in five parameters. The *Hipparcos* (ESA, 1997; van Leeuwen, 2007a) and *Gaia* (Gaia Collaboration et al., 2016, 2018; Lindegren et al., 2018; Gaia Collaboration et al., 2021) catalogs report the five best-fit astrometric parameters: two position parameters at a reference epoch, two proper motions, and the parallax. The five-parameter astrometric model, however, is only an approximation: even a single star with a constant 3D velocity exhibits apparent accelerations when its motion is projected into spherical coordinates (perspective acceleration; Michalik et al., 2014; Lindegren & Dravins, 2021). Stars may also accelerate due to visible and/or unseen companions, or even due to their orbits through the Galaxy.

The influence of unseen companions may be modeled as a higher-order astrometric fit with seven or nine parameters, or as a full Keplerian orbit. The first full orbital fits of *Hipparcos* astrometry were done by the *Hipparcos* team themselves (Lindegren et al., 1997). Work over the following decades built on this work to include brown dwarfs and giant planets (e.g., Zucker & Mazeh, 2001; Sozzetti & Desidera, 2010; Reffert & Quirrenbach, 2011; Sahlmann et al., 2011; Snellen & Brown, 2018). The most recent additions to this field are covering harder to detect and less massive exoplanets, leveraging the long-term proper motion anomalies between *Hipparcos* and *Gaia* (Brandt, 2018; Brandt et al., 2018; Calissendorff & Janson, 2018; Dupuy et al., 2019a; Fontanive et al., 2019; Kervella et al., 2019; Feng et al., 2019; Maire et al., 2020a; Currie et al., 2020; Brandt et al., 2021a).

Because of the importance of *Hipparcos* within the field of constraining the masses and orbits of companions, we present a recalibration in Chapter 5 of *Hipparcos* 2. This yields a catalog which is improved, with inflated formal errors to minimize the effects of overfitting, making the catalog much more suitable for use in orbital fits.

Chapter 2

The Masses and Orbits of β Pic b and β Pic c

2.1 The Necessity for a Reanalysis of the System Orbits

In this chapter, we present a comprehensive orbital analysis to the exoplanets β Pictoris b and c that resolves previously reported tensions between the dynamical and evolutionary mass constraints on β Pic b. We use the MCMC orbit code `orvara` to fit fifteen years of radial velocities and relative astrometry (including recent GRAVITY measurements), absolute astrometry from *Hipparcos* and *Gaia*, and a single relative radial velocity measurement between β Pic A and b. We measure model-independent masses of $9.3^{+2.6}_{-2.5}\,M_{\rm Jup}$ for β Pic b and $8.3 \pm 1.0\,M_{\rm Jup}$ for β Pic c. These masses are robust to modest changes to the input data selection. We find a well-constrained eccentricity of 0.119 ± 0.008 for β Pic b, and an eccentricity of $0.21^{+0.16}_{-0.09}$ for β Pic c, with the two orbital planes aligned to within $\sim$0.5°. Both planets' masses are within $\sim 1\,\sigma$ of the predictions of hot-start evolutionary models and exclude cold starts.

We validate our approach on N-body synthetic data integrated using `REBOUND`. We developed and implemented a three-body approximation into `orvara`. We show that `orvara` can account for three-body effects in the β Pic system down to a level $\sim$5 times smaller than the GRAVITY uncertainties. Systematics in the masses and orbital parameters from `orvara`'s approximate treatment of multiplanet orbits are a factor of $\sim$5 smaller than the uncertainties we derive here. Future GRAVITY observations will improve the constraints on β Pic c's mass and (especially) eccentricity, but improved constraints on the mass of β Pic b will likely require years of additional RV monitoring and improved precision from future *Gaia* data releases.

The contents of this chapter have been published in the Astronomical Journal, under the title "Precise Dynamical Masses and Orbital Fits for β Pic b and β Pic c" (Brandt et al., 2021a).

2.2 Orbit-Fitting Methods and Data

2.2.1 Observational Data

The available data for the β Pic system comprise more than 15 years of radial velocities (RVs) of β Pic A and relative astrometry for β Pic b, three epochs of relative astrometry for β Pic c, a single RV of β Pic b relative to β Pic A, and absolute astrometry of β Pic A from *Hipparcos* and *Gaia*. In this section we summarize each of these.

There are several sources and numerous measurements of relative astrometry for β Pic b, and three recent measurements for β Pic c. We use all relative astrometry, which is comprised of measurements from NICI on Gemini-South (Nielsen et al., 2014), NACO on the VLT (Currie et al., 2011; Chauvin et al., 2012), MagAO on Magellan (Nielsen et al., 2014), GPI on Gemini South (Wang et al., 2016; Nielsen et al., 2020), SPHERE on the VLT (Lagrange et al., 2019b), and GRAVITY on the VLT (7 measurements of β Pic b, 3 of c) (Lagrange et al., 2020; Nowak et al., 2020). This corresponds to the Case 6 relative astrometry data set of Nielsen et al. (2020) plus recent observations by GRAVITY.

GRAVITY measurements of β Pic b clustered near 2020 disagree internally by $\sim 2\sigma$ and result in an unacceptable reduced χ^2 (nearly 3) on Position Angle (PA) in the final fit (see Section 2.3.2). We therefore inflate the errors on the seven GRAVITY measurements of β Pic b by a factor of two to make the PA reduced χ^2 acceptable and bring the PA of the 2020 measurements into internal agreement.

We use the RVs of β Pic A as presented in Vandal et al. (2020), which are corrected for pulsations via a Gaussian Process. We add the five new RVs presented in Lagrange et al. (2020) that are not in the data set of Vandal et al. (2020).[1] We also use the single measurement of the relative RV of β Pic b and β Pic A from Snellen et al. (2014).

We use the absolute astrometry of the *Hipparcos*-*Gaia* Catalog of Accelerations (HGCA, Brandt, 2018). These astrometric measurements adopt the *Gaia* DR2 parallax values as priors to all *Hipparcos* data. *Gaia* is usually much more precise than *Hipparcos*, but β Pic A is at the saturation limit of *Gaia* (G-band magnitude of 3.7). This strongly impacts the astrometric performance of *Gaia* (Lindegren et al., 2018). Thus, the formal parallax uncertainties of the two missions are comparable (assuming a substantial error inflation to the parallax of the *Hipparcos* re-reduction in line with the HGCA's inflation of proper motion errors). Because the HGCA adopts *Gaia* parallaxes as a prior, we take the *Gaia* DR2 parallax value of 50.62 ± 0.33 milli-arcseconds (mas) (Lindegren et al., 2018) as our prior for the orbital fit. This value is consistent to within 1% with the *Hipparcos* values (ESA, 1997; van Leeuwen, 2007a). Regardless, the precise distance to β Pic does not drive our results.

The HGCA argues for a factor of $\sim$2 inflation for all *Gaia* DR2 proper motion errors. We further inflate the HGCA *Gaia* DR2 proper motion errors on β Pic by another factor of 2 (a net factor of $\sim$4 over the *Gaia* DR2 errors). This is due to systematics in the astrometric fit for very bright stars and is justified by the black histogram (worst 5%

---

[1] The RVs between Vandal et al. (2020) and Lagrange et al. (2020) agree within the errors for the epochs mutual to the two data sets.

of stars) in Figure 9 of Brandt (2018). The *Gaia* DR2 proper motion has a negligible impact on our results with this large error inflation.

2.2.2 The orvara MCMC Orbit Fitting Code.

We use `orvara` (Brandt et al., 2021f) along with `htof` (Brandt et al., 2021c; Brandt & Michalik, 2020) to fit for the motion of the β Pic system. `orvara` fits one or more Keplerian orbits to an arbitrary combination of RVs, relative, and absolute astrometry. For the present analysis, we added the ability to fit the single relative RV measurement by Snellen et al. (2014). `orvara` treats the full motion of the system as a linear combination of Keplerian orbits: an orbit between β Pic b and the combined β Pic A/c system, and a second Keplerian orbit between β Pic A and β Pic c. When computing relative astrometry between β Pic A and β Pic c, `orvara` neglects interactions with β Pic b. For relative astrometry between β Pic A and β Pic b, `orvara` computes the displacement of β Pic A from its center of mass with β Pic c and adds this to the displacement of β Pic b from the center of mass of the β Pic A/c system. In other words, `orvara` only adds astrometric perturbations due to inner companions, not due to outer companions. For RVs and absolute astrometry of β Pic A, `orvara` adds the perturbations from the two Keplerian orbits. The perturbation from planet c on the relative RV measurement is negligible.

`orvara` uses `htof` to derive positions and proper motions from synthetic epoch astrometry relative to the system's barycenter. `htof` uses the known *Hipparcos* observation times and scan angles and the predicted *Gaia* observation times and scan angles[2] (with dead times removed) and solves for the best-fit position and proper motion relative to the barycenter. `orvara` then compares these positions and proper motions to the equivalent values in the HGCA.

`orvara` marginalizes out the RV zero point, the parallax, and the barycenter proper motion. We fit a total of 16 parameters to the system using Markov Chain Monte Carlo (MCMC) with `ptemcee` (Foreman-Mackey et al., 2013; Vousden et al., 2016). These are the six Keplerian orbital elements for each of planets b and c, the mass of each companion, the mass of β Pic A, and a RV jitter to be added in quadrature with the RV uncertainties. We adopt uninformative priors on all parameters: uniform priors on all parameters except for RV jitter (a log-uniform prior) and inclination (a geometric prior).

2.2.3 Validation of the 3-body Approximation

Given the approximate treatment of the three-body problem in `orvara`, we test its fidelity on data integrated forward using `REBOUND` (Rein & Liu, 2012). We initialize a 1.8 $M_\odot$ star with two planets of 9 and 8 $M_{\rm Jup}$; we give these planets the best-fit orbital elements of β Pic b and c, respectively, found by Lagrange et al. (2020). We then integrate the system forward to produce synthetic RVs and relative astrometry for both companions with `REBOUND`. We take 52 measurements of relative astrometry for each planet distributed over 17 years, each of which has the 100 μas precision typical of GRAVITY (Gravity Collaboration et al., 2020; Lagrange et al., 2020). We fit 52 RV points, each with an

---

[2] https://gaia.esac.esa.int/gost/

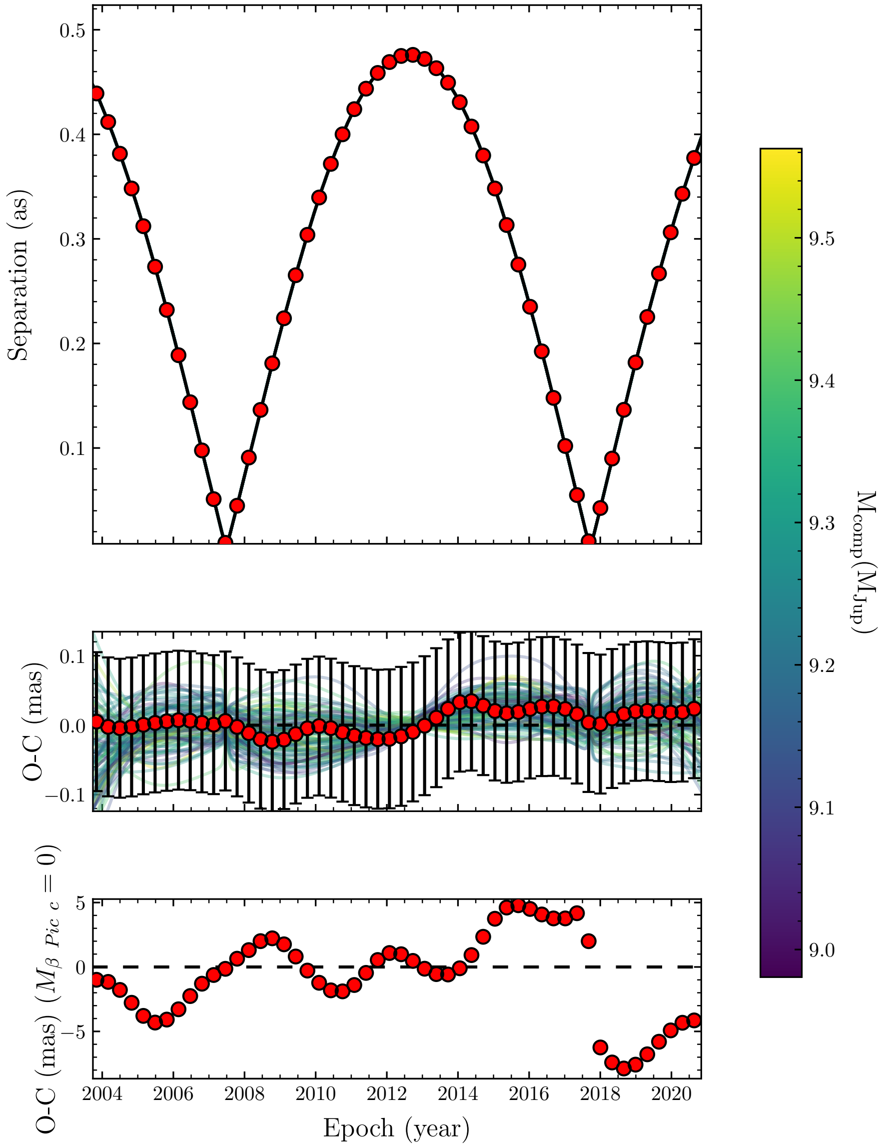

Figure 2.1 `orvara` can fit the 3-body system of β Pic to several factors below the GRAVITY precision (assumed to be 0.1 mas). Top panel: the observed separation for the fictitious β Pic b analog with evenly spaced observations with the precision of GRAVITY as presented in Lagrange et al. (2020). Black is the best fit orbit found by an `orvara` MCMC analysis. Middle panel: observed data minus the best fit orbit (O-C). The remaining variations are from the mutual tugs of β Pic c on β Pic b. These variations are at the level of ∼0.02 mas – a factor of 5 below the 0.1 mas precision of the GRAVITY-like data. Bottom panel: the O-C if the approximate 3-body compensation is turned off in `orvara`. The synthetic data here are generated with a 9 $M_{\rm Jup}$ β Pic b and the best-fit mass is 9.2 $M_{\rm Jup}$.

uncertainty of 1 m/s. We add a single relative RV between β Pic b and A (the synthetic analog to the relative RV of Snellen et al. (2014)) with an uncertainty of 1 km/s. We then fit these synthetic data with `orvara`.

`orvara` is able to fit all data satisfactorily. Figure 2.1 shows that the unmodeled three-body effects are a factor of ∼5 below the level detectable by GRAVITY (see middle panel). The superposition of the two Keplerian orbits shows up in the relative astrometry of β Pic b, where synthetic GRAVITY observations clearly detect the orbit of β Pic A about its center of mass with β Pic c (bottom panel of Figure 2.1). Unmodeled RV residuals are well below 1 m/s (the reduced χ^2 of the RV fit is 0.05). We derive masses that agree well, but not perfectly, with the input masses: the derived masses of β Pic A and β Pic b are each ∼3% larger than their true values. These systematics are a factor of ∼5 lower than the uncertainties we derive for β Pic c in the following section and are negligible for β Pic b.

`orvara` returns two body elements for each planet about the star. In the three body system that we initialized in `REBOUND`, the two-body input orbital elements (semi-major axis, eccentricity, etc.) cease to have a strict meaning unless a primary is specified (e.g., the barycenter or β Pic A). However, we still expect the recovered orbital elements to roughly be equal to those that were used as inputs. We expect the semi-major axes to be close but not exactly equal to the inputs, because, e.g., the input semi-major axis of β Pic b was defined relative to the barycenter of β Pic A and c – yet β Pic b will orbit the total system barycenter during integration. Likewise, we expect the argument and time of periastron to be biased slightly. Elements like the inclination i and PA of the ascending node Ω should be returned *exactly* – the 3-body interactions should not rotate the orientation of either orbit over a $\sim$20 year integration.

We find that `orvara` recovers i and Ω exactly; with a residual less than 10^{-3} of a degree (nearly equal to the formal error) on both. Although unexpected, we recover the eccentricity exactly: the residual is less than 10^{-4} and the formal error is $2 \cdot 10^{-4}$. The three elements recovered with biases follow. The argument of periastron and mean longitude at the reference epoch are recovered to within 0.2 degrees. The semi-major axes of both planets are recovered to within 0.1 A.U.

We conclude that our approximation to the three-body dynamics is more than sufficiently accurate for the β Pic system: the biases induced in the parameters inferred from the test data are much smaller than the formal errors on the measured parameters. Our accounting of only inner companions when perturbing relative astrometry recovers the masses to within a few percent. Figure 2.1 shows that a full N-body integration of the β Pic system will remain unnecessary even with future GRAVITY relative astrometry.

2.3 Results

We infer masses and orbital parameters using a parallel-tempered MCMC with 15 temperatures; for each temperature we use 100 walkers with one million steps per walker.[3] Our MCMC chains converged after 40,000 steps; we conservatively discard the first 250,000

---

[3] `orvara` completes this million-step MCMC in roughly 4 hours on a 4 GHz AMD Ryzen desktop processor.

as burn in and use the remainder for inference.[4]

We check convergence informally by confirming that we obtain the same posterior distributions, for every parameter, from any several percent portion of our chains. Next, the acceptance fraction of the coldest chain is satisfactory (∼0.15). Lastly, multiple MCMC analyses starting with different, and in many cases poor, initial guesses converge to the same posterior distributions. We quantitatively confirm convergence with the Gelman-Rubin Diagnostic (GRD) (Gelman & Rubin, 1992; Roy, 2019). Perfect convergence for a parameter is suggested if the GRD is 1, and a common threshold adopted for convergence is 1.1 Roy (2019). Our chains have GRD values better than 1.0001 for all parameters, although one should keep in mind that the GRD was designed for chains with independent walkers.

2.3.1 Orbital Analysis of the β Pic System

Table 2.1 lists the six Keplerian orbital elements for both β Pic b and β Pic c, along with the other five fitted parameters.

---

[4] The chains and input data are available by request.

Table 2.1. Posteriors of the β Pic system from an `orvara` MCMC analysis.

Parameter	Prior Distribution	Posteriors $\pm 1\sigma$	
Stellar mass	Uniform	$1.83 \pm 0.04\,M_\odot$	
Parallax (ϖ)	$50.62 \pm 0.33\,\mathrm{mas}$ (*Gaia* DR2)	$50.61 \pm 0.47\,\mathrm{mas}$	
Barycenter Proper Motions^{b}	Uniform	$\mu_\alpha = 4.80 \pm 0.03\,\mathrm{mas\,yr^{-1}}$ & $\mu_\delta = 83.87 \pm 0.03\,\mathrm{mas\,yr^{-1}}$	
RV Zero Point	Uniform	33 ± 13 m/s	
RV jitter	Log uniform over $[0, 300\,\mathrm{m/s}]$	$50 \pm 8\,\mathrm{m/s}$	
Parameter	Prior Distribution	Posterior on β Pic b $\pm 1\sigma$	Posterior on β Pic c $\pm 1\sigma$
Semi-major axis (a)	Uniform	10.26 ± 0.10 A.U.	$2.738^{+0.034}_{-0.032}$ A.U.
Eccentricity (e)	Uniform	0.119 ± 0.008	$0.21^{+0.16}_{-0.09}$
Inclination (i)	$\sin i$ (geometric)	88.94 ± 0.02 degrees	89.1 ± 0.66 degrees
PA of ascending node (Ω)	Uniform	211.93 ± 0.03 degrees	$211.1^{+0.3}_{-0.2}$ degrees
Mean Longitude at t_{ref} (λ_{ref})	Uniform	-36.7 ± 0.9 degrees	-50^{+13}_{-14} degrees

Table 2.1 (cont'd)

Parameter	Prior Distribution	Posteriors $\pm 1\sigma$	
Planet Mass (M)	Uniform	$9.3^{+2.6}_{-2.5} M_{\mathrm{Jup}}$	$8.3 \pm 1.0 M_{\mathrm{Jup}}$
Argument of Periastron (ω)	(derived quantity)	$22.6^{+2.8}_{-2.9}$ degrees	$119^{+30}_{-7.0}$ degrees
Periastron Time (T_0)	(derived quantity)	2456656^{+61}_{-64} BJD	2455789^{+95}_{-63} BJD
Period	(derived quantity)	8864^{+118}_{-113} days	1222^{+18}_{-17} days
		$24.27^{+0.32}_{-0.31}$ years	$3.346^{+0.050}_{-0.045}$ years
`orvara` Reference Epoch (t_{ref})	2455197.50 BJD	$\cdots$	$\cdots$

Note. — Orbital elements all refer to orbit of the companion about the barycenter. The orbital parameters for β Pic A about each companion are identical except $\omega_A = \omega + \pi$. We use ± when the posteriors are Gaussian. In the case of non-Gaussian posteriors we denote the value by $\mathrm{median}^{+u}_{-l}$ where u and l denote the 68.3% confidence interval about the median. The reference epoch t_{ref} is not a fitted parameter and has no significance within the fit itself, it is the epoch at which the Mean Longitude (λ_{ref}) is evaluated.

[a] The posterior on the eccentricity of β Pic c is not Gaussian. However, eccentricities below 0.1 and above 0.7 are strongly disfavored (See Figure 2.3).

[b] μ_α and μ_δ refer to the proper motions in right-ascension and declination, respectively.

Every fitted element of β Pic b results in a nearly Gaussian posterior (see Figure 2.2). The elements of β Pic c are also well-constrained except for eccentricity and the mean longitude at the reference epoch λ_{ref}. The mean longitude at the reference epoch is poorly constrained because of the poor constraint on the eccentricity, which results from having only three relative astrometric measurements closely spaced in time. We show the variances and covariances between the fitted parameters in Figure 2.3 for β Pic c as a corner plot. There is a modest covariance between semi-major axis and eccentricity resulting from the short time baseline of relative astrometry on β Pic c.

The best-fit orbit and nearby (in parameter space) orbits agree well with all data: the pulsation-corrected RVs, the Snellen et al. (2014) relative RV, the relative astrometry from VLT/NACO, Gemini-South/NICI, Magellan/MagAO, Gemini-South/GPI, and GRAVITY, and absolute astrometry from the HGCA.

Figure 2.4 shows the agreement between the calibrated *Hipparcos* and *Gaia* proper motions from Brandt (2018) and the best fit orbit. The sum of the χ^2 of the fits to both proper motions is very good (nearly 1, see Table 2.4). There are six measurements, but the unknown barycenter proper motion removes two degrees of freedom. The reflex motion of β Pic c with a period of ∼three years is clearly seen, as well as the long term oscillation from the ∼24 year orbit of β Pic b. Here the constraining power of the *Hipparcos* proper motion is visible: the *Hipparcos* proper motion is much more precise than that of *Gaia* DR2 for β Pic b and can exert a sizable tug on the mass and mass uncertainty of β Pic b.

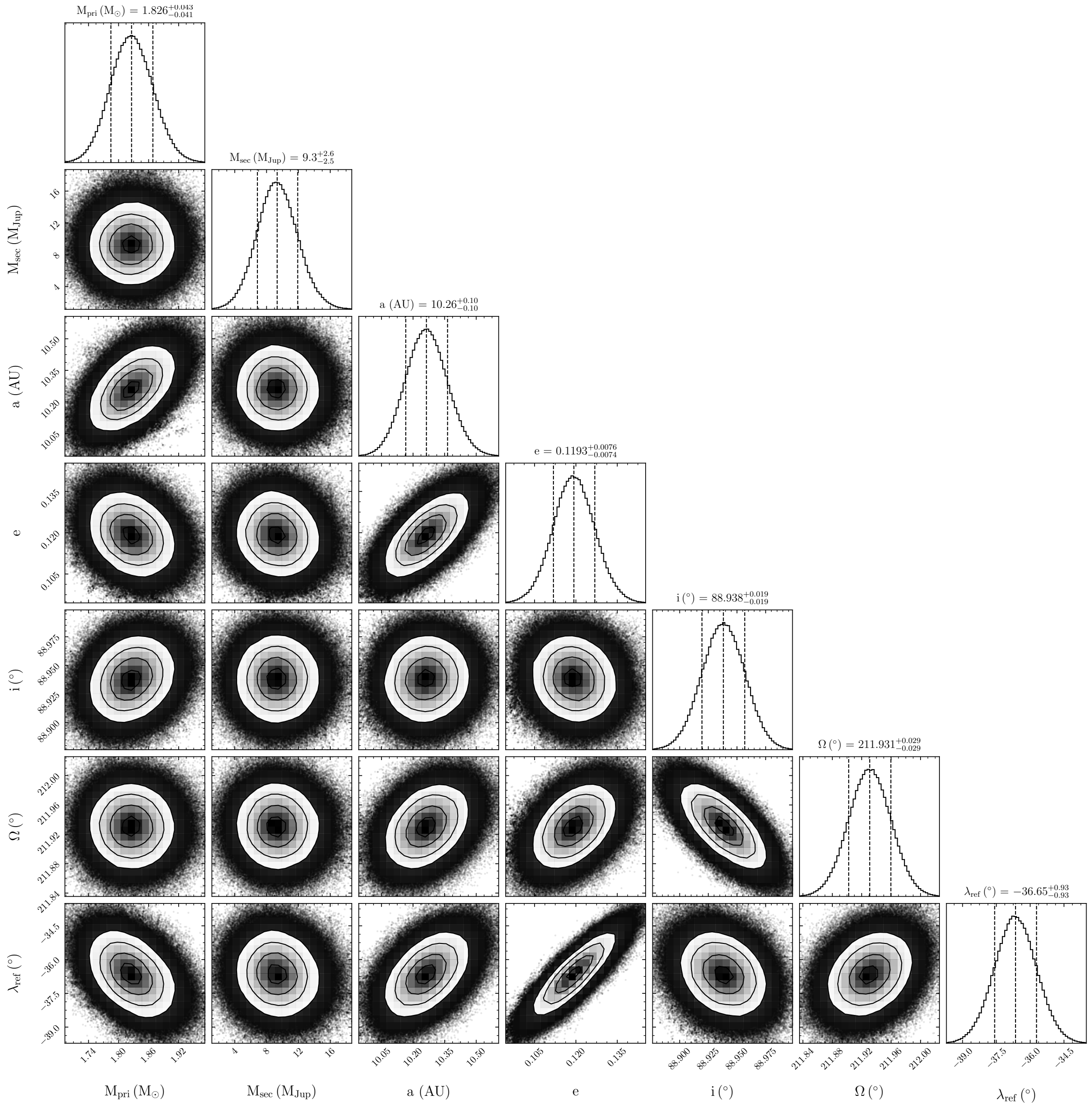

Figure 2.2 Best fit orbital elements for β Pic b from the `orvara` MCMC chain. Orbital elements are with respect to the star. The elements, in the same order as plotted, are: the primary mass in solar masses, M_{pri}; the planet mass in Jupiter masses, M_{sec}; the semi-major axis in A.U., a; the eccentricity, e; the inclination in degrees, i; the PA of the ascending node in degrees, Ω; and the mean longitude at the reference epoch (2455197.50 BJD) in degrees, λ_{ref}. In the 1D histograms, the vertical-dashed lines about the center dashed lines give the 16% and 84% quantiles around the median. In the 2d histograms, the contours give the 1-σ, 2-σ, and 3-σ levels.

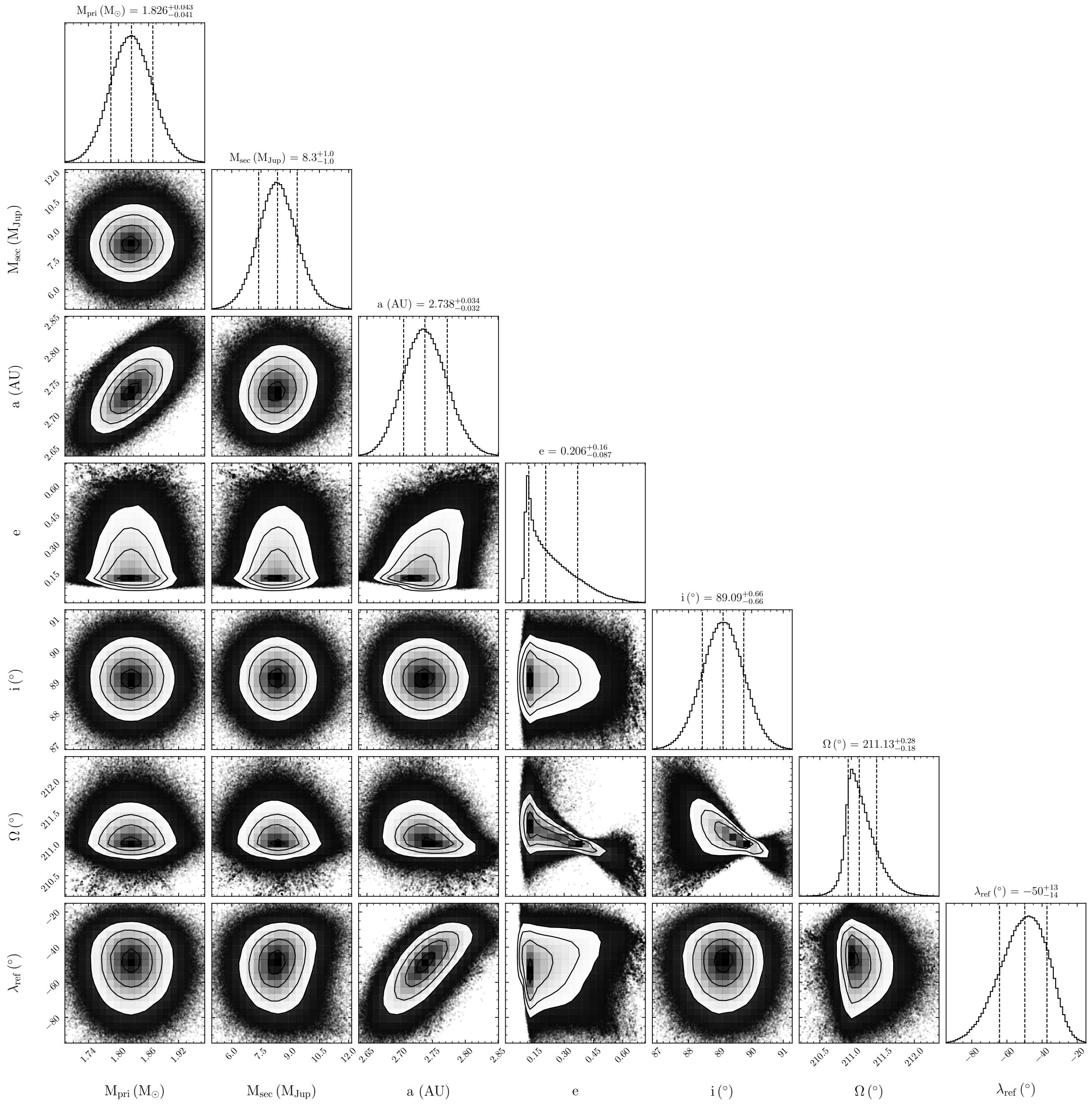

Figure 2.3 Best fit orbital elements for β Pic c. See Figure 2.2 for the description.

Figures 2.5 and 2.6 show the agreement in relative separation and PA from our set of relative astrometry (Case 3 from Nielsen et al. (2020) plus the seven GRAVITY measurements on β Pic b and three GRAVITY measurements on β Pic c). Figure 2.7 shows the agreement between the RVs from Vandal et al. (2020) and Lagrange et al. (2020) and the best-fit orbit. The jitter parameter found by the MCMC analysis is $50 \pm 8\,\mathrm{m/s}$. Lower masses for β Pic c slightly favor lower eccentricities. The Snellen et al. (2014) relative RV χ^2 is 1.7 (indicating a $\sim 1.3\,\sigma$ residual). However, our posteriors are completely identical within rounding if we exclude the single Snellen et al. (2014) measurement.

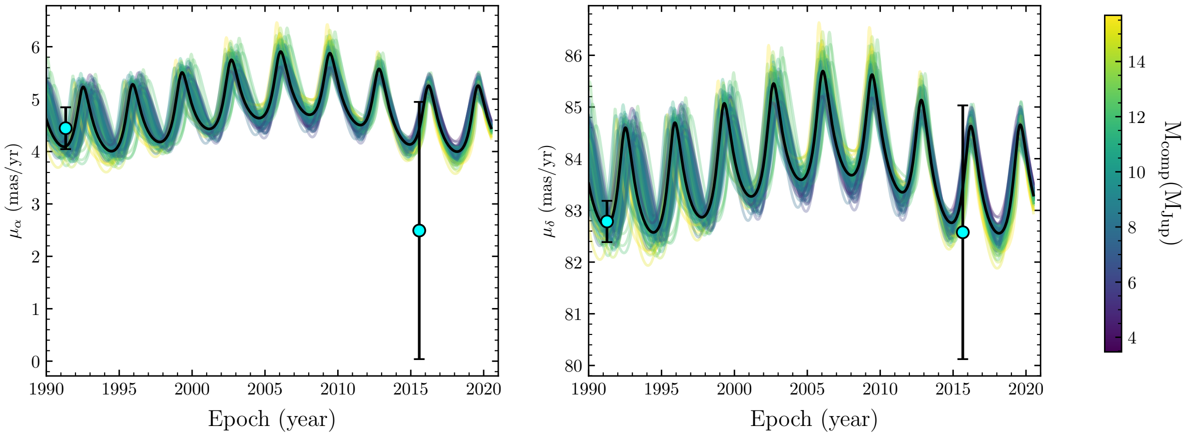

Figure 2.4 Model proper motions compared to the calibrated *Hipparcos* (dot at 1991.25) and *Gaia* proper motions (dot near 2015) from the HGCA. The *Gaia* DR2 proper motion uncertainty has been inflated by an extra factor of 2, as in Dupuy et al. (2019a), to account for additional uncertainties with stars as bright as β Pic (see Figure 9 of Brandt, 2018). The best fit orbit is shown in black. A random sampling of other orbits from the MCMC chain are shown and are color coded by the mass of β Pic b.

We display an additional corner plot in Figure 2.8 that showcases select covariances between the orbital parameters of β Pic b and β Pic c. The inferred mass of each planet is relatively insensitive to the orbital parameters of the other (see the two appropriate covariances in the left hand columns of Figure 2.8). In particular, the mass of β Pic b is nearly independent of the mass of c. However, owing to the 3-body interaction between the planets, the inferred eccentricity of β Pic b varies slightly with the eccentricity of the inner planet, β Pic c. Improved relative astrometry on β Pic b mildly improves

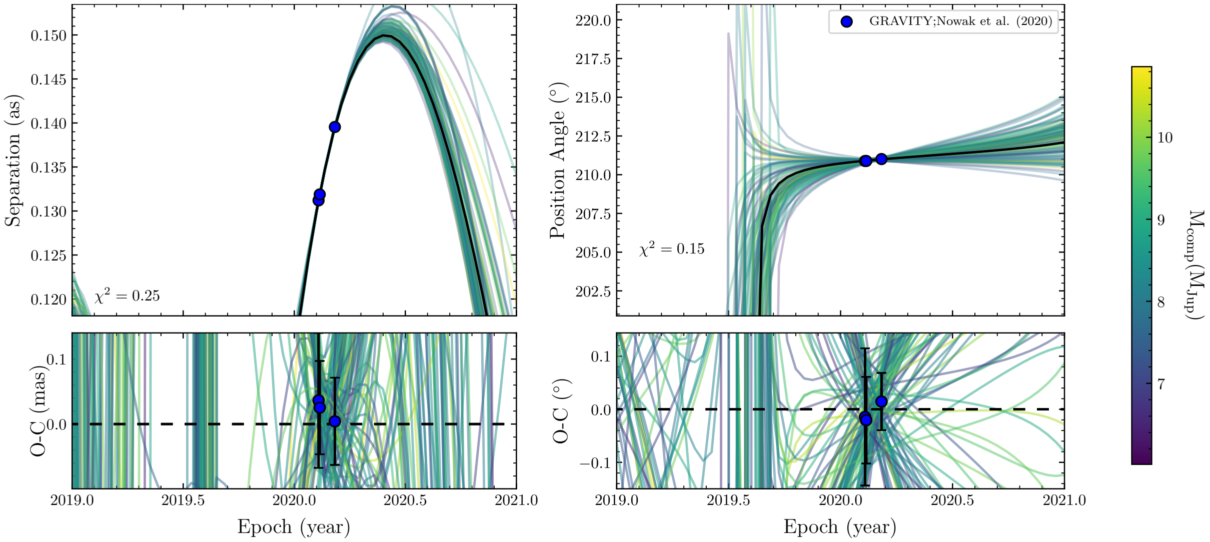

Figure 2.5 Left: relative separation of β Pic c. Right: PA of β Pic c. All three data points are from GRAVITY (Nowak et al., 2020) and are not error inflated. The best fit orbit is shown in black. A random sampling of other orbits from the MCMC chain are shown and are color coded by the mass of β Pic c.

constraints on the eccentricity of β Pic c; an identical orbital fit excluding the SPHERE relative astrometry on β Pic b results in a slightly worse eccentricity constraint on β Pic c. The inferred semi-major axis of β Pic c covaries modestly with β Pic b's eccentricity. Despite uncertainties in the eccentricity of β Pic c, we find that β Pic b and β Pic c are coplanar to within a half-degree at 68% confidence and coplanar to within one degree at 95% confidence.

We use our new constraints on the orbital parameters of β Pic b and c to predict their on-sky positions over the next 5 years at 15-day intervals. Tables 2.2 and 2.3 give a truncated version of the predicted positions of β Pic b and c. The supplementary data contain the full tables. β Pic c will be less than ∼50 mas from the star by March of 2021. β Pic c will re-emerge (once again being further than ∼50 mas from the star) in October

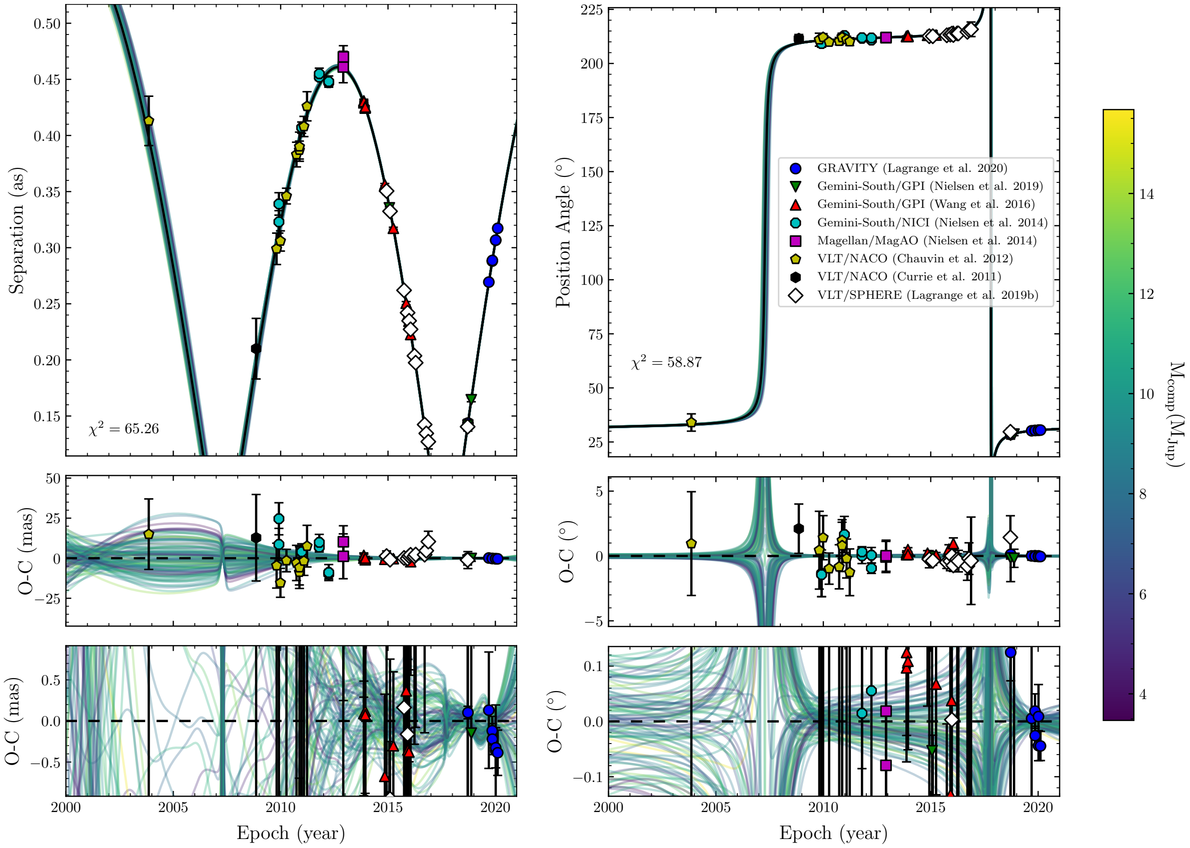

Figure 2.6 Left: relative separation of β Pic b. Right: PA of β Pic b. The GRAVITY errors have been inflated by a factor of two to make the reduced χ^2 of the fit acceptable. A random sampling of orbits from other MCMC steps are shown and are color coded by the mass of β Pic b. The best fit orbit is shown in black.

of 2021. Our predicted positions from our orbit analysis localize both β Pic b and c to within ± 40 mas, which is well within the fiber field of GRAVITY (Nowak et al., 2020), at any point over the next 5 years.

2.3.2 Assessing Consistency of Relative Astrometry

Table 2.4 shows quantitatively the goodness of the orbital fit in terms of χ^2 for PA, separation, RV, and the three proper motions. The reduced χ^2 for all the astrometry

Table 2.2. Predicted positions of β Pic b.

Date	δ mas	σ_δ mas	α mas	σ_α mas	$\rho_{\alpha\delta}$	Sep mas	σ_{Sep} mas
2020-12-30	351.9	0.5	211.6	0.3	0.951	410.6	0.1
2021-01-14	355.2	0.6	213.7	0.4	0.954	414.5	0.1
2021-01-29	358.4	0.6	215.8	0.4	0.957	418.4	0.2
2021-02-13	361.6	0.6	217.9	0.4	0.959	422.2	0.2
2021-02-28	364.7	0.7	219.9	0.4	0.961	425.9	0.2
2021-03-15	367.9	0.7	222.0	0.4	0.963	429.6	0.2
2021-03-30	370.9	0.8	224.0	0.5	0.965	433.3	0.2
2021-04-14	374.0	0.8	226.0	0.5	0.967	437.0	0.3
2021-04-29	377.0	0.8	227.9	0.5	0.969	440.6	0.3
...	...	...	...	...	...	...	...
2025-12-19	460	10	290	7	0.999	540	10

Note. — The offsets, and their errors (σ), from the star in right-ascension (α), declination (δ), and separation (Sep), are given in milli-arcseconds (mas). $\rho_{\alpha\delta}$ is the correlation coefficient between right-ascension and declination. A non-rounded, machine readable version of this table with all 122 epochs is available with the supplementary data online (or with the source TeX files if viewing this on ArXiv). This portion is shown here for guidance regarding its form and content.

Table 2.3. Predicted positions of β Pic c.

Date	δ mas	σ_δ mas	α mas	σ_α mas	$\rho_{\alpha\delta}$	Sep mas	σ_{Sep} mas
2020-12-30	−82	8	−52	5	0.911	97	5
2021-01-14	−76	9	−49	6	0.927	90	5
2021-01-29	−70	10	−45	6	0.939	83	4
2021-02-13	−60	10	−41	7	0.950	75	2
2021-02-28	−60	10	−37	8	0.958	68	3
2021-03-15	−50	10	−33	9	0.965	60	8
2021-03-30	−40	10	−29	9	0.970	50	10
2021-04-14	−40	20	−20	10	0.975	40	20
2021-04-29	−30	20	−20	10	0.978	30	10
⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
2025-12-19	60	20	40	10	0.992	80	20

Note. — See the table note of Table 2.2 for a description of the columns.

Table 2.4. The goodness of the `orvara` orbital fit to the various data in the β Pic system.

Data	Points (N)	χ^2
Separation	56	65.6
PA	56	59.0
All Astrometry	112	[a]118.8
RV	41	40.7
β Pic b – A relative RV	1	1.67
Hipparcos μ (HGCA)	2	0.33
Gaia μ (HGCA)	2	0.75
HGCA long baseline μ	2	0.001

Note. — The χ^2 quoted here include both companions and are for the maximum likelihood orbits. The χ^2 for μ includes both μ_δ and μ_α. N is the number of data points in the corresponding data set.

[a] This χ^2 is slightly less than the sum of the χ^2 in PA and separation because of the covariance between PA and separation in the GRAVITY observations.

(which takes into account the GRAVITY covariances between separation and PA) is 1.06. A good fit should have $\chi^2/N \approx 1$ where N is the number of degrees of freedom.

Nielsen et al. (2020) argued for a systematic offset between the SPHERE relative astrometry from Lagrange et al. (2019b) and the relative astrometry from Gemini-South/GPI. Nielsen et al. (2020) investigated fitting for an offset in both separation and PA within the SPHERE data. The SPHERE data do appear to be systematically offset in PA relative to the best-fit orbit (See the bottom right panel of Figure 2.6). However, a fit without the 12 SPHERE observations reduces the χ^2 in PA and separation by roughly the expected 12 points, suggesting that the data are consistent with the astrometric record. Moreover, the reduced χ^2 including SPHERE is acceptable ($\sim$65 points of χ^2 for 56 data points) and so we include SPHERE in our final analysis.

We find evidence for either an underestimate in the PA uncertainties from GPI or an offset in PA between GPI and one or more of the other astrometric data sets (see the Wang et al. (2016) GPI data in the right panel of Figure 2.6). Removing the 15 GPI relative astrometry measurements decreases the χ^2 in PA by roughly 40. Using the χ^2 survival function, a change of that magnitude corresponds to roughly 2.5 σ evidence in favor of a PA offset. However, including GPI still results in an acceptable overall χ^2 (See Table 2.4), and so we include GPI in our final fit.

Whether or not we include one, both or neither of GPI and SPHERE, our results are nearly identical. The best fit masses on both β Pic b and c shift by less than 0.5 $M_{\rm Jup}$ between all three cases, and the confidence intervals on their masses are identical to within 5%. This speaks to the constraining power of the GRAVITY measurements, and to the robustness of our results with respect to the details of how the relative astrometry is analyzed.

In Figure 2.5, the χ^2 of the β Pic c fit to the relative astrometry is much less than 1 because the relative astrometry is effectively overfit: the RVs primarily constrain the mass, phase, and semimajor axis of β Pic c while the four remaining orbital parameters have substantial freedom to fit the three relative astrometry points (6 coordinates). By contrast, β Pic b is overconstrained by the data and the reduced χ^2 of the fit is near 1. The right-hand side of the bottom-most panel for both separation and PA in Figure 2.6 show the GRAVITY points for β Pic b. GRAVITY points near the same epoch (in both PA and separation) disagree by $\lesssim 1\sigma$ after error inflation. Without error inflation, GRAVITY observations near the same epoch disagree by $\sim 2\sigma$ and the reduced χ^2 in PA of the best fit jumps to nearly 3 for β Pic b.

The three GRAVITY measurements of β Pic c do not have χ^2 or agreement issues. We leave the errors on β Pic c as they are in Nowak et al. (2020). However, inflating the errors by a factor of 2 on β Pic c does not significantly change our results: the resulting posteriors and errors are identical except for the errors on β Pic c's inclination, which are doubled.

2.3.3 N-body Simulations

We expect the orbital parameters of β Pic b and c to vary over time due to the mutual influence between these two massive planets. The evolution of the eccentricity and orbit of β Pic b depends heavily on the eccentricity of β Pic c, which is poorly constrained. In Figure 2.9, we show the evolution of the β Pic system over 0.1 million years (Myr), integrated forward using the ias15 integrator of `REBOUND` (Rein & Liu, 2012; Rein & Spiegel, 2015), assuming the median orbital parameters presented in Table 2.1 for each planet. We vary the eccentricity of β Pic c within the posterior constraints. The grey shaded region shows how the eccentricity of β Pic c and β Pic b could evolve over the next 10^5 years. The two planets exchange eccentricity with a period of ∼50,000 years. We found numerically that the system is stable and the periodic variability in Figure 2.9 repeats for at least the next 10 Myr.

2.4 Discussion

Our mass measurements for β Pic b and c agree within 1σ compared to previous work by Snellen & Brown (2018), Dupuy et al. (2019a), Nielsen et al. (2020), and Vandal et al. (2020). Our analysis is the first to incorporate the new GRAVITY measurements with uninformative priors while obtaining masses in the expected expected range. Our error bars on the mass of β Pic b are larger than all but Dupuy et al. (2019a) because, like that work, we adopt the inflated errors on the *Hipparcos* proper motions as recommended by Brandt (2018). Our mass posteriors do not change if we exclude the Snellen et al. (2014) relative RV measurement. We were unable to reproduce the $\approx 3\,M_{\rm Jup}$ and $\approx 5\,M_{\rm Jup}$ (when using a uniform prior) mass estimates from Nowak et al. (2020) and Lagrange et al. (2020). Using their slightly different data set, we find $9.5^{+2.0}_{-1.8}\,M_{\rm Jup}$ for β Pic b and $9.2^{+1.0}_{-0.8}\,M_{\rm Jup}$ for β Pic c.

We corroborate the findings by Nielsen et al. (2020) and Nowak et al. (2020) that β Pic c and β Pic b are coplanar. Nowak et al. (2020) found inclinations for β Pic b and c of 88.99 ± 0.01 degrees and 89.17 ± 0.50 degrees, respectively, with a Gaussian prior on the mass of β Pic b. We confirm these inclinations without an informative prior. We find 88.94 ± 0.02 degrees and 89.1 ± 0.7 degrees.

β Pic is surrounded by an extended debris disc and an inner disc that is slightly misaligned with respect to the primary (Smith & Terrile, 1984; Heap et al., 2000). The extended debris disc around β Pic is inclined at 90.0 ± 0.1 degrees (Ahmic et al., 2009; Kraus et al., 2020). β Pic b is thus misaligned by 1.06 ± 0.11 degrees with respect to the debris disc. Our inferred inclination for β Pic c slightly favors misalignment but does not exclude alignment.

Nowak et al. (2020) found that the β Pic system exhibited an oscillating eccentricity for both bodies over a timescale of $\approx 5 \times 10^4$ years using their orbital parameter posteriors. We find variations in eccentricity over a similar timescale and confirmed numerically with `REBOUND` that the system is stable for at least the next 10 Myr. We find that it is moderately likely to observe the current eccentricity of the system amidst all the possible eccentricities over a 10 million year timespan.

The first observational evidence that β Pic b has a significant, nonzero eccentricity was presented by Dupuy et al. (2019a). They discussed the implications of an eccentricity as high as $\approx$0.2 in the context of both single- and multi-planet scenarios; at the time β Pic c was not known. The scenario in which β Pic b formed on a circular orbit but gained eccentricity from interactions with the disk and migrated inward to its current location, with no influence from β Pic c, is still plausible. Such a pathway is available to any sufficiently massive planet. Given that we find that β Pic c is also massive ($8.3 \pm 1\, M_{\mathrm{Jup}}$), it may have also opened a gap in the disk, migrating inward and acquiring eccentricity from gravitational interactions with β Pic b and the disk. Indeed, with two such massive planets in close proximity it is natural to expect that both should have significantly nonzero eccentricities by a system age of $\approx$20 Myr.

Our masses follow from uniform priors, allowing us to independently assess the agreement of the dynamical masses with model predictions. To simplify our model comparisons, we assume an age of 20 Myr for the system, compatible with all available age determinations for the β Pic moving group (Binks & Jeffries, 2014; Mamajek & Bell, 2014; Miret-Roig et al., 2020). We examine the hot-start Cond (Baraffe et al., 2003) models, the Saumon & Marley (2008) models with a hybrid cloud treatment (which we denote as SM08), and the warm-start Spiegel & Burrows (2012) models (SB12) with hybrid clouds and solar metallicity but a range of initial entropies. We perform our comparisons in the K band, as this is the only measurement available for β Pic c (Nowak et al., 2020). We convert luminosities to K-band magnitudes for the SM08 models using Cond colors at the SM08 effective temperatures.

Figure 2.10 shows our results. We find that our dynamical mass measurement for β Pic c is consistent with all models except those with low initial entropies ($\lesssim 10\,k_{\rm B}$/baryon). Our dynamical mass for β Pic b is roughly $1\,\sigma$ below the predictions of hot-start models, and rules out cold starts. Similarly to previous work (Dupuy et al., 2019a; Vandal et al., 2020), none of the disagreements with models are significant beyond $\sim$$1\sigma$, and the precisions of the dynamical masses are insufficient to distinguish between most of the models shown. Stronger tests of models will require significantly better precision, especially for β Pic b.

As Figure 2.10 shows, reaching 0.1–0.5 $M_{\rm Jup}$ levels of precision on the mass of β Pic b is crucial to accurately discern between evolutionary models. The best prospect for improving the mass of β Pic b is long term RV monitoring over the next decade. Even drastically improved absolute astrometry (e.g., *Gaia* DR4) will only provide a modest improvement to the mass measurement of β Pic b. If we assume optimistically that *Gaia* at the end of its mission will achieve the same precision on the $G = 3.7$ mag β Pic A as it has on $G \approx 6$ mag stars (the brightest for which the mission was originally designed), then it would achieve a factor of ~100 improvement on the proper motion of β Pic A.[5] The uncertainty on the mass of β Pic b would shrink by 35%, to $\pm 1.7 M_{\rm Jup}$, if the proper motion precision is improved by a factor of 100–using the same MCMC analysis as presented here with otherwise the same data. Assuming more conservatively that *Gaia* reaches only a factor of 10 better precision on the proper motion of β Pic, the uncertainty on the mass of β Pic b improves by 25%.

2.5 Conclusions

In this paper we have derived masses and orbits of both planets in the β Pictoris system with uninformative priors. We validated our approach against synthetic data from a full N-body integration. Our masses and orbital parameters are derived from two decades of observational data. The GRAVITY data show clear evidence of the gravitational perturbations of β Pic c ($P = 3.346^{+0.050}_{-0.045}$ yr) on the orbit of β Pic b relative to A

---

[5] https://www.cosmos.esa.int/web/gaia/science-performance

($P = 24.27 \pm 0.32\,\mathrm{yr}$). The resulting model-independent masses allow us to compare the observed properties of β Pic b and c with predictions from models of the formation and evolution of giant planets. We summarize our main results below.

1. We find a mass of $9.3^{+2.6}_{-2.5}\ M_{\mathrm{Jup}}$ for β Pic b and $8.3 \pm 1.0\,M_{\mathrm{Jup}}$ for β Pic c with uninformative priors all orbital parameters. The mass constraint on β Pic c is superior due to the RVs covering many orbital periods and due to the impact of β Pic c on the relative astrometry of β Pic b.
2. β Pic b and β Pic c are both consistent with Spiegel & Burrows (2012) warm-start models with initial entropies of at least $10\,k_{\mathrm{B}}$/baryon. They are also both consistent with a 20 Myr age under the hot-start COND evolutionary tracks (Baraffe et al., 2003) and the Saumon & Marley (2008) models using a hybrid cloud model. In all cases, consistency with models would favor a mass for β Pic b that is $\sim 1\sigma$ higher than our dynamical measurement.
3. We find an eccentricity of 0.119 ± 0.008 for β Pic b and $0.21^{+0.16}_{-0.09}$ for c. These modest eccentricities could have been generated by interactions with the disk, or via the mutual interactions between b and c. The eccentricity and mean longitude of β Pic c are poorly constrained because there are only three relative astrometric observations, and these are closely spaced in time. There is a mild covariance between the eccentricity of β Pic b and c owing to the three-body dynamics in the system. Additional GRAVITY relative astrometry on β Pic c will help constrain the eccentricity of β Pic b and especially β Pic c.

4. The mass constraint on β Pic b needs to be improved by a factor of $\sim$3–5 in order to more reliably constrain its age or formation conditions. Long-term RV monitoring over the coming years or decade is needed for better mass constraints on β Pic b. An improved proper motion from a future *Gaia* data release will offer up to a 35% better constraint on the mass of β Pic b (assuming *Gaia* reaches a better precision on the brightest stars).

The new GRAVITY relative astrometry (Nowak et al., 2020; Lagrange et al., 2020) appeared to create tension between dynamical and spectral mass constraints on β Pic b. Our analysis dissolves this tension and results in masses for β Pic b and β Pic c that are consistent with warm and hot start evolutionary models. Additionally, the system is dynamically interesting – with eccentricities of both planets varying by $\sim$50% over 10^4–10^5 year timescales. The precision on the masses and eccentricities of β Pic b and c will improve with continued astrometric and RV monitoring. The planets around β Pic A will continue to provide some of the best tests of super-Jovian planet formation and evolution.

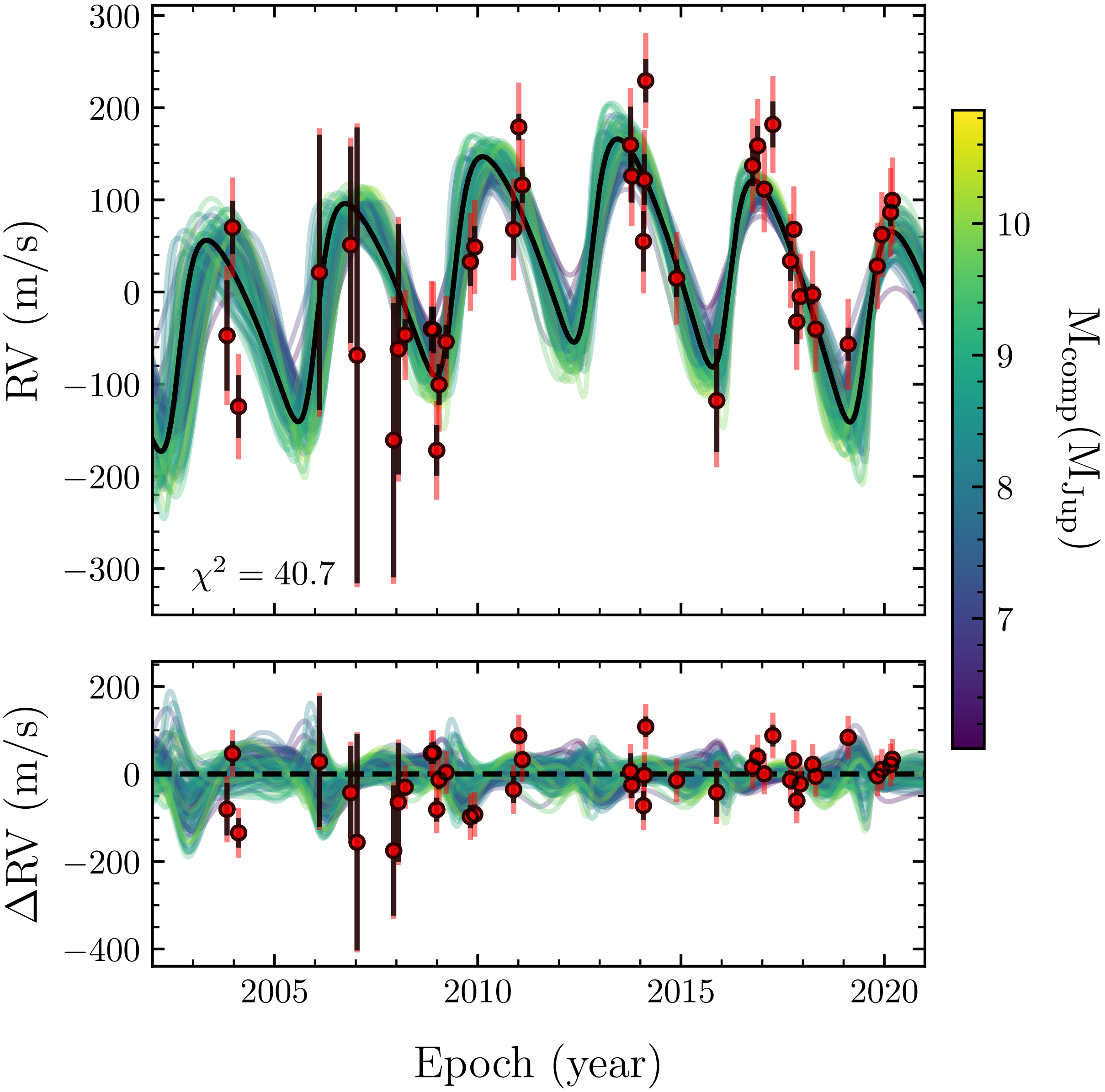

Figure 2.7 The best fit orbit (black) agrees well with the observed β Pic pulsation-corrected RVs. β Pic c has an eccentricity of $e = 0.30$ in the best fit orbit while b has $e = 0.120$. Top panel: The observed RVs overplot with the best fit orbit and a random sampling of other orbits from the MCMC chain. Bottom panel: The RV residuals with respect to the best fit orbit. Both panels: The random sampling of other orbits from the MCMC chain are color coded by the mass of β Pic c. RVs are from Vandal et al. (2020) with the most recent 5 points from Lagrange et al. (2020). The black error bars give the observed errors reported by Vandal et al. (2020) and Lagrange et al. (2020). The red error bars include the best fit jitter of ∼50m/s added in quadrature to the observed errors.

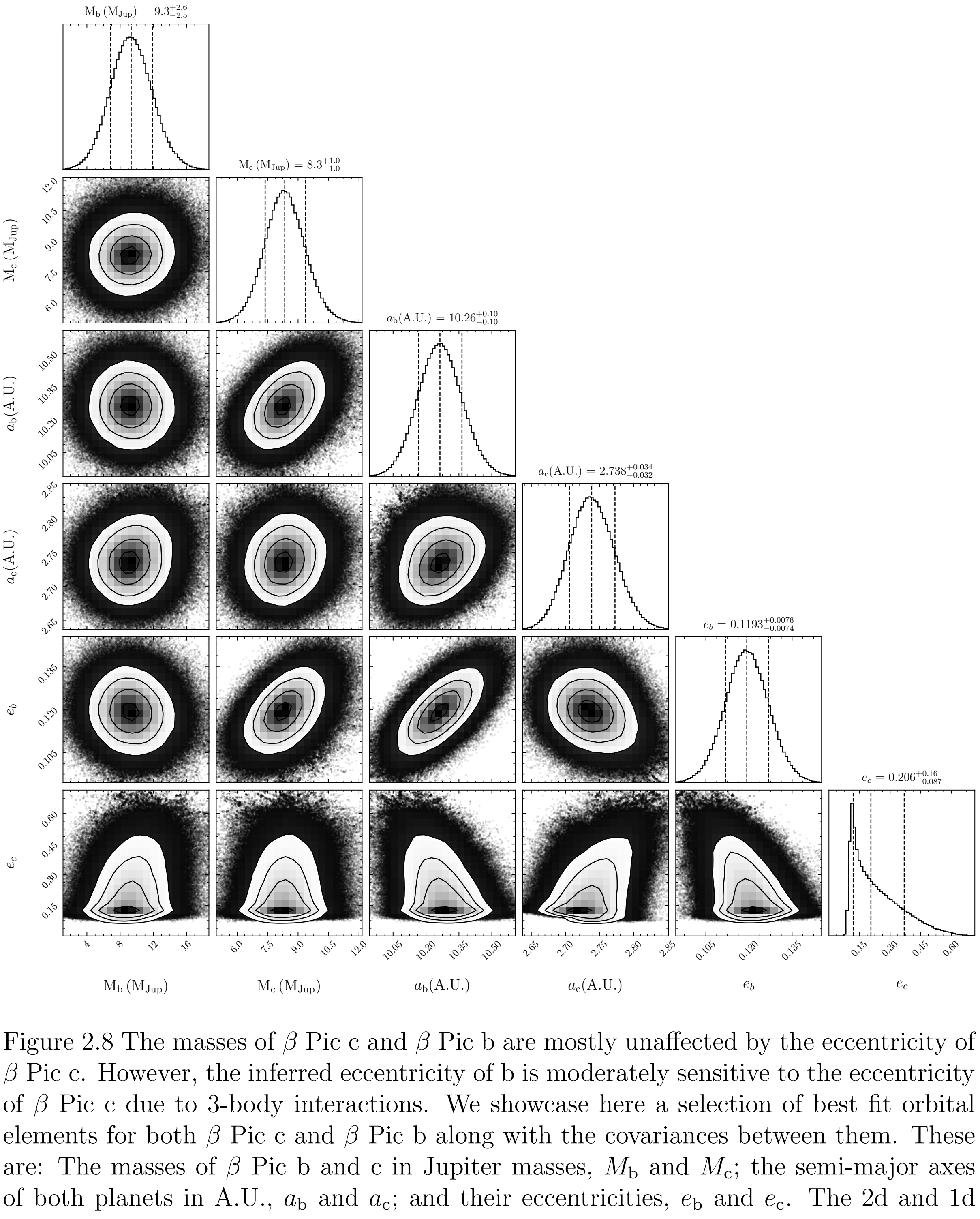

Figure 2.8 The masses of β Pic c and β Pic b are mostly unaffected by the eccentricity of β Pic c. However, the inferred eccentricity of b is moderately sensitive to the eccentricity of β Pic c due to 3-body interactions. We showcase here a selection of best fit orbital elements for both β Pic c and β Pic b along with the covariances between them. These are: The masses of β Pic b and c in Jupiter masses, M_b and M_c; the semi-major axes of both planets in A.U., a_b and a_c; and their eccentricities, e_b and e_c. The 2d and 1d contours are described in Figure 2.2.

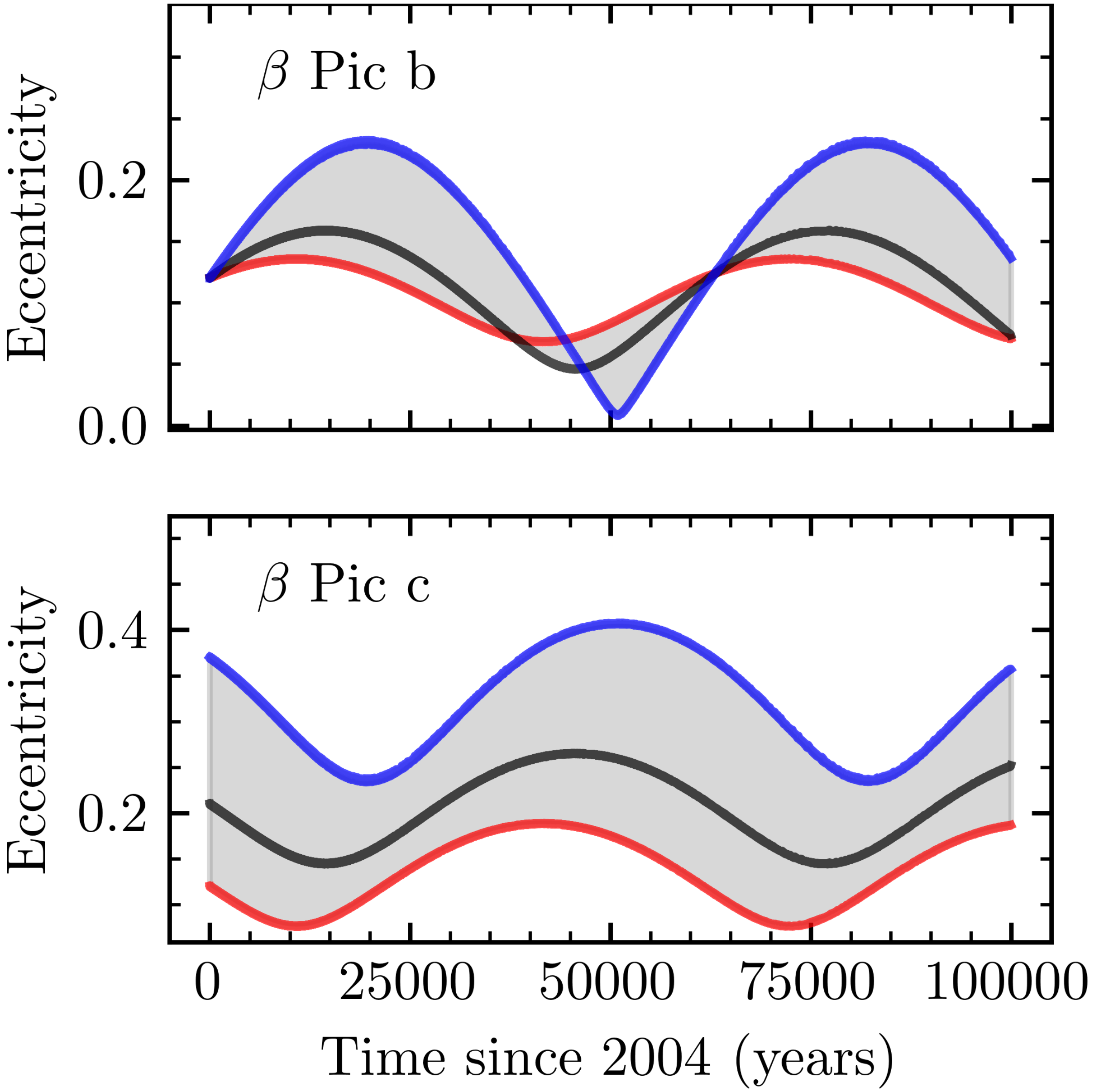

Figure 2.9 Top panel: the eccentricity evolution of β Pic b computed using `REBOUND`'s ias15 integrator (Rein & Spiegel, 2015) for the median (black, 0.21) and 68.3% confident bounds on the eccentricity of β Pic c from Table 2.1 (0.12 is red and 0.37 is blue). Bottom panel: the eccentricity evolution of β Pic c for its median (black) and 68.3% confident eccentricities. The parameter space spanned by the 68.3% confident range of eccentricities is shaded grey. The two planets exchange eccentricity over a ~50,000 year cycle.

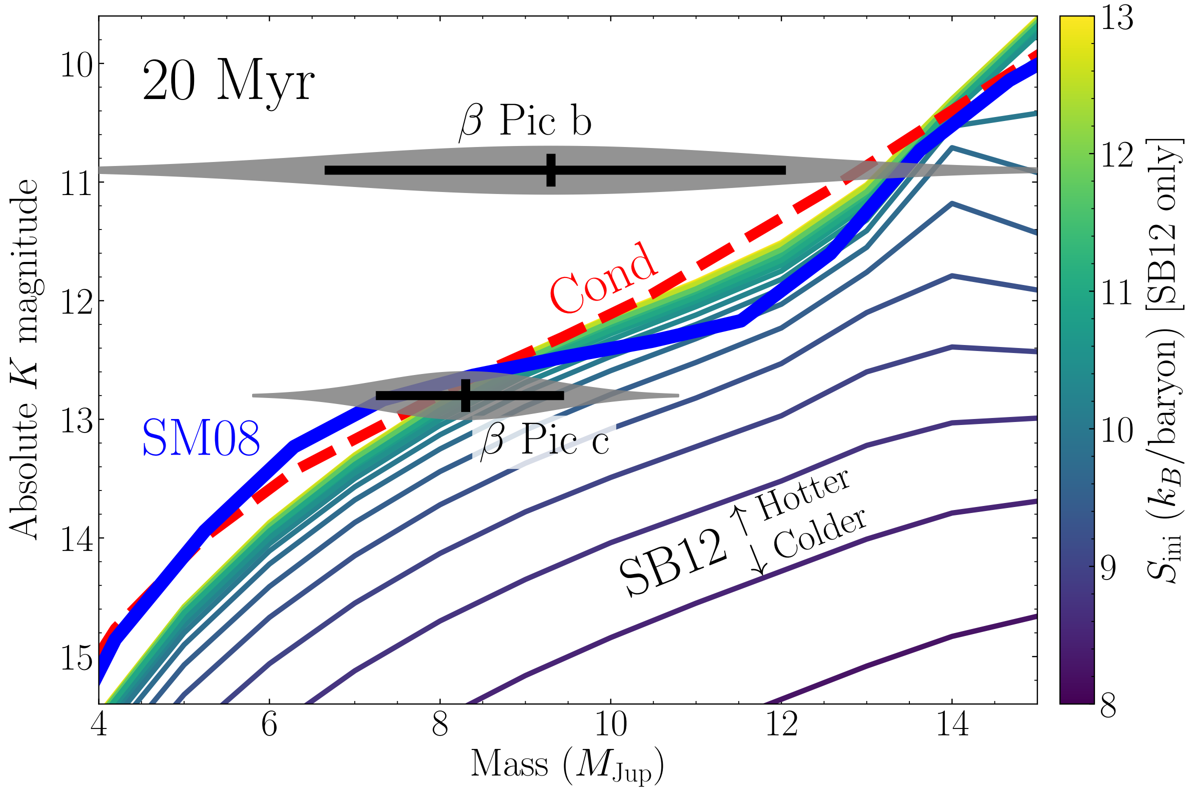

Figure 2.10 Comparison of dynamical mass measurements (gray shaded regions) and observed K-band magnitudes (Nowak et al., 2020) with Cond (Baraffe et al., 2003), SM08 (Saumon & Marley, 2008), and SB12 (Spiegel & Burrows, 2012) models, all at an age of 20 Myr (Binks & Jeffries, 2014; Mamajek & Bell, 2014; Miret-Roig et al., 2020). The SM08 and SB12 models both use hybrid cloud prescriptions and adopt Solar metallicity. The SB12 models also vary (and are color-coded by) their initial entropy. The black lines show $1\,\sigma$ values, while gray shaded regions show the probability density. Our dynamical mass for β Pic c is consistent with all three models assuming a hot start, and rules out a very cold start. Our dynamical mass for β Pic b is $\sim 1\,\sigma$ below the prediction of the hot start models.

Chapter 3

The First Dynamical Mass Measurement in the HR 8799 System

3.1 The Multi-Planet HR 8799 System

HR 8799 hosts four directly imaged giant planets, but until this work, none had a mass measured from first principles. In this chapter, we present the first dynamical mass measurement in this planetary system, finding that the innermost planet HR 8799 e has a mass of $9.6^{+1.9}_{-1.8}\,M_{\rm Jup}$. This mass results from combining the well-characterized orbits of all four planets with a new astrometric acceleration detection (5σ) from the *Gaia* EDR3 version of the *Hipparcos*-*Gaia* Catalog of Accelerations. We find with 95% confidence that HR 8799 e is below 13 $M_{\rm Jup}$, the deuterium-fusing mass limit. We derive a hot-start cooling age of 42^{+24}_{-16} Myr for HR 8799 e that agrees well with its hypothesized membership in the Columba association but is also consistent with an alternative suggested membership in the β Pictoris moving group. We exclude the presence of any additional $\gtrsim$5-$M_{\rm Jup}$ planets interior to HR 8799 e with semi-major axes between ≈3–16 au. We provide proper motion anomalies and a matrix equation to solve for the mass of any of the planets of HR 8799 using only mass ratios between the planets.

The contents of this chapter have been published in the Astrophysical Journal Letters, under the title "The First Dynamical Mass Measurement in the HR 8799 System" (Brandt et al., 2021b).

3.2 Methodology

The acceleration that HR 8799 A experiences is the sum of the acceleration due to each of its four planetary companions. Because the planet masses are small compared to the mass of HR 8799 A ($1.47^{+0.11}_{-0.08}\,M_\odot$; Wang et al., 2018), the star's motion is approximately given by a linear combination of the orbits of the planets, weighted by their masses. We can then optimize the planet masses until the modelled host star acceleration matches the observed value.

We use all 1000 samples from the orbital posteriors from Wang et al. (2018), published on the github repository for the resource `whereistheplanet.com` (Wang et al., 2021). The orbits correspond to the coplanar, dynamically stable case of the HR 8799 system (see Table 4 of Wang et al., 2018). Each set of orbital parameters, together with the masses of the four planets and the star, predicts the motion of HR 8799 A.

Wang et al. (2018) used the *Gaia* DR1 parallax of 24.76 ± 0.64 mas (Gaia Collaboration et al., 2016)[1]. *Gaia* EDR3 measures a much more precise value of 24.462 ± 0.046 mas (Lindegren et al., 2020). We scale all 1000 MCMC samples to the *Gaia* EDR3 parallax, removing its contribution to the uncertainty of the orbital fits. Defining r to be the ratio of a given chain's parallax to the *Gaia* EDR3 value, we multiply the semimajor axes by r (to preserve the relative astrometry) and multiply the system mass by r^3 (to preserve the orbital periods). We keep all other orbital parameters unchanged. Updating the parallax to its *Gaia* EDR3 value ultimately improves the fractional error on the final mass estimate from 20% to 19.5%.

We use the open-source tool `htof` (Brandt et al., 2021d; Brandt & Michalik, 2020; Brandt et al., 2021f) to model *Hipparcos* and *Gaia* observations. In brief, `htof` uses the *Hipparcos* intermediate astrometric data (both ESA (1997) and van Leeuwen (2007a) reductions) and predicted scan angles and observational epochs of *Gaia* (via GOST[2]), to generate synthetic *Hipparcos* and *Gaia* astrometry for any orbit.

The astrometric measurement that we use is the proper motion anomaly. This is the difference between a nearly instantaneous proper motion from *Gaia* EDR3 and a long-term proper motion (the difference in position between the *Hipparcos* and *Gaia* astrometry missions divided by the time between the measurements). We denote these

---

[1] In all cases where we quote a posterior by listing m_{-l}^{+u} or $m \pm \sigma$: m denotes the median with l and u (or singularly, σ) denoting the 16% and 84% confidence intervals, respectively.

[2] https://gaia.esac.esa.int/gost/

proper motion anomalies as, e.g.,

$$\Delta\mu_{\alpha*} = \mu_{Gaia} - \frac{\alpha*_{Gaia} - \alpha*_{Hip}}{t_{Gaia} - t_{Hip}} \tag{3.1}$$

where $\alpha* = \alpha\cos\delta$. Parameters with subscript *Hip* refers to the average of those parameters from *Hipparcos* 2007 and 1997, weighted 60/40 as they are in the HGCA (Brandt, 2018). A set of orbital parameters then gives predicted values for $\Delta\mu_{\alpha*}$ and $\Delta\mu_\delta$ as functions of the masses of the planets, m_b, m_c, m_d, and m_e. Because HR 8799 A's motion closely follows a linear combination of Keplerian orbits, we can represent its predicted proper motion anomaly as the Jacobian

$$\begin{bmatrix} \Delta\mu_{\alpha*} \\ \Delta\mu_\delta \end{bmatrix}_{\text{model}} = \begin{bmatrix} \frac{\partial\Delta\mu_{\alpha*}}{\partial m_b} & \frac{\partial\Delta\mu_{\alpha*}}{\partial m_c} & \frac{\partial\Delta\mu_{\alpha*}}{\partial m_d} & \frac{\partial\Delta\mu_{\alpha*}}{\partial m_e} \\ \frac{\partial\Delta\mu_\delta}{\partial m_b} & \frac{\partial\Delta\mu_\delta}{\partial m_c} & \frac{\partial\Delta\mu_\delta}{\partial m_d} & \frac{\partial\Delta\mu_\delta}{\partial m_e} \end{bmatrix} \begin{bmatrix} m_b \\ m_c \\ m_d \\ m_e \end{bmatrix}. \tag{3.2}$$

We compute the partial derivatives by using only the orbit of a given planet and assigning that planet unit mass. We do this for all 1000 orbital draws. We use `REBOUND` and the `ias15` scheme (Rein & Liu, 2012; Rein & Spiegel, 2015) to integrate the orbits in time.

When we compute the model partial derivatives, we mix the ESA (1997) and van Leeuwen (2007a) positions according to the same 60/40 ratio adopted by the HGCA. Brandt (2018) show in their Section 7 and Figure 2 that a 60/40 mix of the two *Hipparcos* reductions' proper motion measurements better matches the long-term proper motions between *Hipparcos* and *Gaia* than either reduction on its own. The EDR3 version of the HGCA (Brandt, 2021) confirms this finding and also shows that a 60/40 mix of the two *Hipparcos* reductions' position measurements best matches the *Gaia* positions extrapolated back to the *Hipparcos* observational epoch. We compute our positions at the same central epochs as given in the HGCA[3]. This forward modeling allows us to directly compare our proper motion anomalies to the values given in the HGCA. The HGCA is calibrated so that the measured proper motion anomalies have Gaussian uncertainties. We can therefore identify χ^2 with $-2\ln\mathcal{L}$ and find the masses by maximizing the likelihood $\mathcal{L}$, or minimizing

$$\chi^2 = -2\ln\mathcal{L} = \mathbf{d}^T\left(\mathbf{C}_{HG} + \mathbf{C}_{Gaia}\right)^{-1}\mathbf{d} \tag{3.3}$$

where $\mathbf{C}_{HG}$ is the HGCA covariance matrix for the two *Hipparcos*-*Gaia* long-term proper motions, $\mathbf{C}_{Gaia}$ is the HGCA covariance matrix for the two *Gaia* EDR3 proper motions, and

$$\mathbf{d} = \begin{bmatrix} \Delta\mu_{\alpha*\mathrm{model}} - \Delta\mu_{\alpha*\mathrm{HGCA}} \\ \Delta\mu_{\delta\mathrm{model}} - \Delta\mu_{\delta\mathrm{HGCA}} \end{bmatrix}. \tag{3.4}$$

In Equation (3.4), $\Delta\mu_{\alpha*\mathrm{model}}$ and $\Delta\mu_{\alpha*\mathrm{model}}$ are calculated from the right-hand side of Equation 3.2. $\mathbf{C}_{HG}$ and $\mathbf{C}_{Gaia}$ are given in the HGCA. We republish their sum here,

---

[3] These central epochs are, in years for dec. and right-ascension: 2015.85 2015.76 for *Gaia* EDR3 and 1991.35 1991.34 for *Hipparcos*.

along with the anomalies, for ease of reproducibility:

$$\mathbf{C}_{HG} + \mathbf{C}_{Gaia} = \begin{bmatrix} 6.5934 & 0.7473 \\ 0.7473 & 6.9888 \end{bmatrix} 10^{-3} \left(\mathrm{mas\,yr^{-1}}\right)^2 \tag{3.5}$$

$$\begin{bmatrix} \Delta\mu_{\alpha*\mathrm{HGCA}} \\ \Delta\mu_{\delta\mathrm{HGCA}} \end{bmatrix} = \begin{bmatrix} -0.268 \\ -0.348 \end{bmatrix} \mathrm{mas\,yr^{-1}}. \tag{3.6}$$

We compute the partial derivatives of Equation (3.3) against $\Delta\mu_{\alpha*}$, and $\Delta\mu_{\delta}$. Setting these partials to zero gives an under-constrained system of equations. The two components of the proper motion anomaly measure a combination of the masses of the four planets, but not the four masses individually. However, if we assume a relationship between the individual masses, then Equation (3.3) produces an over-constrained system for the masses. If we take a uniform prior on one planet mass and assume mass ratios for the remaining three planets, Equation (3.3) represents a Gaussian mass posterior for each set of orbital parameters.

3.3 Results

In this section we use Equations (3.2)–(3.6), together with varying assumptions about the mass ratios of the four planets, to derive constraints on the mass of HR 8799 e.

3.3.1 Fixed mass ratios

We initially assume the fixed mass ratios derived by Wang et al. (2018) from the observed luminosities assuming hot-start models,

$$m_e = m_d = m_c = 1.25 m_b. \tag{3.7}$$

With the assumed ratios of the masses of the four planets, Equation (3.3) fits one free parameter to two covariant data points. We begin with the calculation of the Jacobians, Equation (3.2) (one for each of the 1000 orbital draws). We fit 1000 Jacobians and sum the posteriors. The resulting posterior is very nearly Gaussian with mean $9.59 M_{\rm Jup}$ and standard deviation $1.84 M_{\rm Jup}$. The residual from the best-fit mass should be χ^2-distributed with one degree of freedom. The best-fit χ^2 is only 0.05: the observed astrometric acceleration from the HGCA agrees almost perfectly with the model prediction using the Wang et al. (2018) orbits and mass ratios.

Because the orbital elements of the HR 8799 planets are so well characterized, the derivatives of the anomalies with respect to each planet's mass vary little between the MCMC draws. The element-wise median Jacobian matrix is an approximation to all 1000 sets of orbital parameters; it is given by

$$\frac{\mathbf{J}_{\rm median}}{\mu{\rm as\,yr^{-1}\,M_{Jup}^{-1}}} = \begin{bmatrix} 2.00 & -3.94 & -3.01 & -24.20 \\ 1.14 & 3.80 & -13.74 & -26.05 \end{bmatrix}. \tag{3.8}$$

The width of our posterior is dominated by the observational uncertainty in absolute astrometry from the HGCA, which corresponds to about $\pm 1.8\, M_{\rm Jup}$ (20%). The uncertainty in the mass of HR 8799 e from the orbital draws alone is just $0.35\, M_{\rm Jup}$ (enlarging the final error bars by only 2%). Using the median Jacobian (Equation (3.8)) together with the mass ratios in Equation (3.7), and broadening with the orbital motion uncertainty ($0.35\,M_{\rm Jup}$), gives a posterior that is indistinguishable from the full posterior using 1000 sets of orbital parameters.

We use the median Jacobian matrix for the remainder of this work and convolve our posterior mass distributions for HR 8799 e with a Gaussian of standard deviation $0.35\,M_{\rm Jup}$ (which accounts for the negligible contribution from orbital motion uncertainty).

3.3.2 Varying mass ratios

The mass ratios of the four planets are known only to roughly ±15% from hot-start models (Wang et al., 2018). We now show that the mass of HR 8799 e is robust to larger variations in the mass ratios. We quantify changes in the mass ratios by three coefficients γ_i, which we define by

$$m_e = \gamma_d m_d = \gamma_c m_c = \gamma_b 1.25 m_b. \tag{3.9}$$

Setting $\gamma_d = \gamma_c = \gamma_b = 1$ yields the fiducial mass ratios.

Figure 3.1 shows the mass posterior of HR 8799 e assuming nine different mass ratios. We vary each of the γ_i independently between 0.7 to 1.4 (corresponding to varying the masses of b,c, and d by roughly $\pm 4\,M_{\rm Jup}$). Varying the mass ratio of planet b or c to planet e by this amount changes the inferred mass of HR 8799 e by $\lesssim 0.1\,M_{\rm Jup}$, or $\lesssim 0.05\sigma$. The top curves in Figure 3.1 show that the mass of HR 8799 e is covariant with HR 8799 d but that the variation is $\approx 1\,M_{\rm Jup}$, or $\approx 0.5\sigma$, even with the extreme range of mass ratios presented. Taylor expanding the maximum likelihood mass, $\langle m_e \rangle$, about our base case mass ratios $\gamma_d = \gamma_c = \gamma_b = 1$ yields

$$\begin{aligned} \frac{\langle m_e \rangle}{M_{\rm Jup}} &\approx 9.56 - 2.50\,(\gamma_d - 1) \\ &\quad - 0.001\,(\gamma_c - 1) + 0.37\,(\gamma_b - 1)\,. \end{aligned} \tag{3.10}$$

The mass of HR 8799 d is moderately covariant with that of HR 8799 e, while HR 8799 b and HR 8799 c have little effect: b is too far away, while c induces a proper motion anomaly nearly perpendicular to that induced by e.

We now generalize Figure 3.1 by marginalizing over the possible range of mass ratios (i.e., effectively summing the posteriors of Figure 3.1). We use independent, log-normal priors (base-e lognormal) centered on unity for each of γ_d, γ_c, and γ_b. For our fiducial case, we take the γ_i priors to have standard deviation 0.15 of the natural logarithm (0.065 dex). This corresponds to ±15%, reflecting hot-start uncertainties (Wang et al., 2018). We also include a worst-case where we use a logarithmic prior with a standard deviation of 0.45 (0.2 dex). This allows for deviations in the mass ratios of roughly -35% and $+55\%$, slightly more than that shown in Figure 3.1.

Figure 3.2 shows the posteriors on planet e's mass under the two prior choices. Our preferred result is a nearly Gaussian posterior of $9.6^{+1.9}_{-1.8}\,M_{\rm Jup}$. Adopting our worst-case prior, allowing for three times the range of mass ratios, yields $9.4^{+2.2}_{-2.1}\,M_{\rm Jup}$.

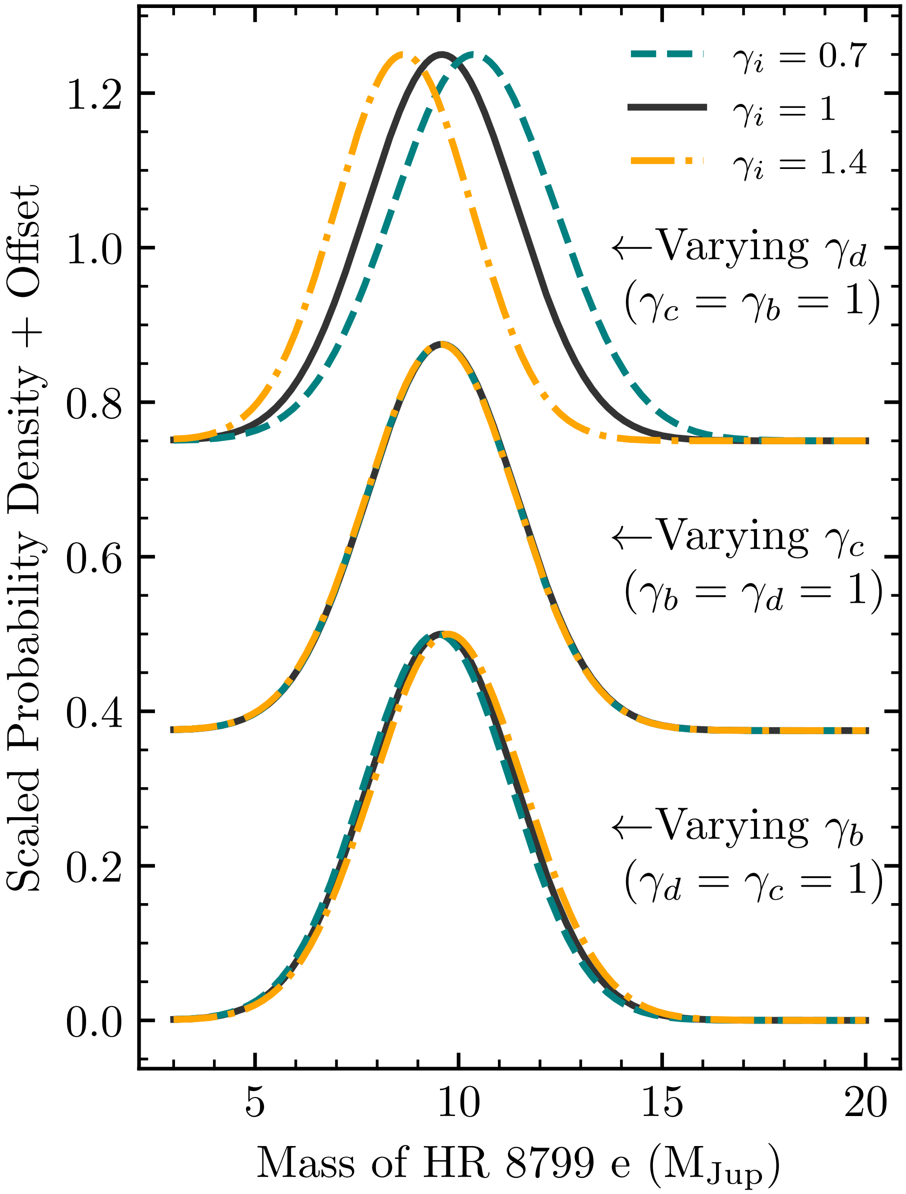

Figure 3.1 Posteriors for the mass of HR 8799 e, varying the assumed mass ratio of each other planet relative to planet e from 70% ($\gamma = 0.7$, teal dashed lines) to 140% ($\gamma = 1.4$, orange dot-dashed lines) of its fiducial value (see Equation (3.9)). The mass of HR 8799 e is moderately covariant with that of planet d but insensitive to the masses of planets b and c.

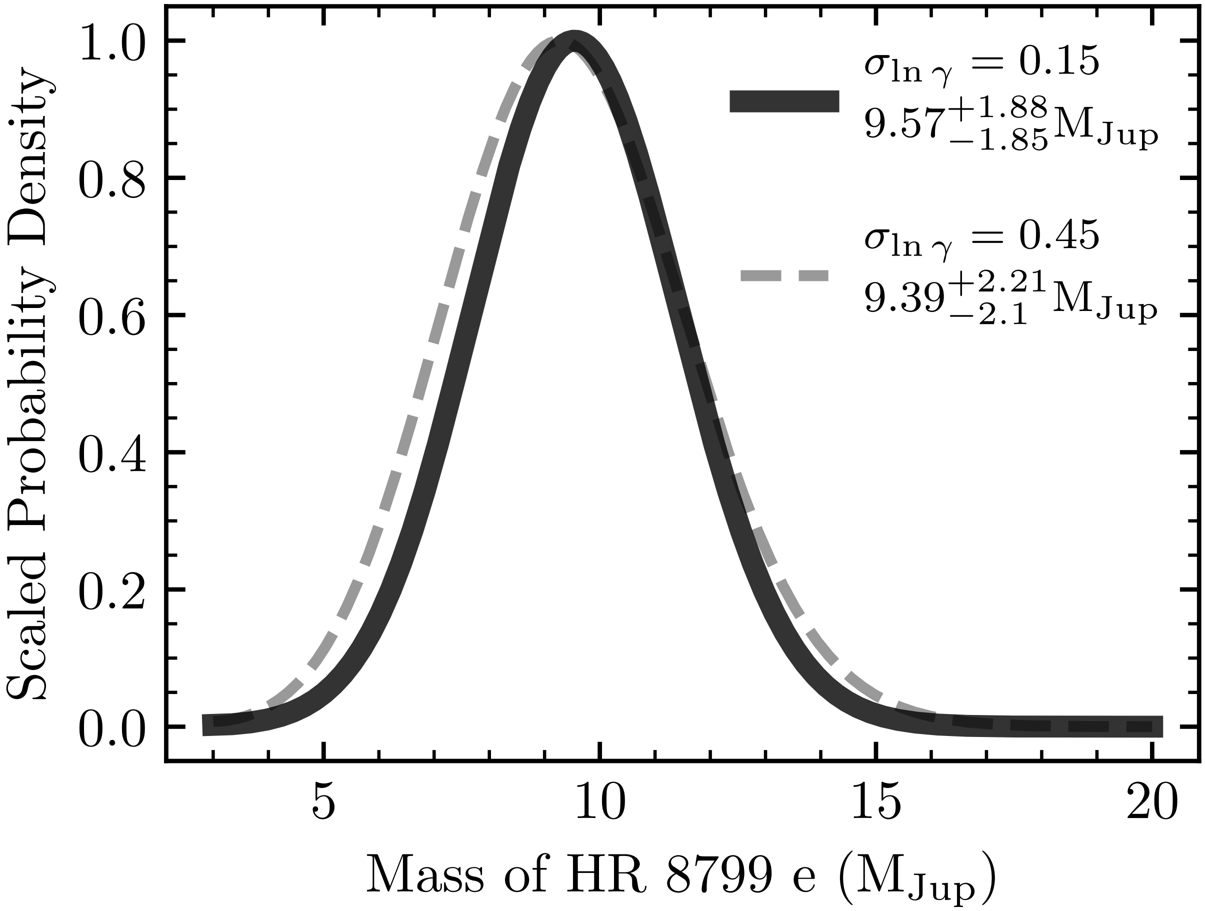

Figure 3.2 The mass posterior of HR 8799 e after marginalizing over bcd-to-e mass ratios (quantified by three γ_i coefficients), each with a lognormal prior. Each posterior is labelled according to both its mass-ratio prior and the resulting posterior (non-rounded for convenience). $\sigma_{\ln\gamma}$ is the natural logarithmic standard deviation of the log-normal prior on the γ_i. Each prior has a natural logarithmic mean of 0.

3.3.3 Additional companions

An additional companion is detectable if it causes a significant astrometric perturbation on the host star. Figure 3.3 shows the semi-major axes and masses that an additional, unseen massive companion would need to cause perturbation large enough to be detected in *Gaia* EDR3. Planets in the blue region above the grey band would have yielded a significant (3σ) astrometric acceleration on the system that we would have seen in our analysis. Additional planets in the parameter space below the grey band (white region) are not excluded. We conclude that additional, unseen massive planets orbiting between 3 and 8 au with masses exceeding $6\,M_{\rm Jup}$ are unlikely, as well as $\gtrsim 7\,M_{\rm Jup}$ companions between 8 au and the orbit of HR 8799 e (≈16 au). We therefore exclude the presence of any $\gtrsim 7\,M_{\rm Jup}$ companions orbiting amidst the inner debris belt, which spans 6–15 au (Frantseva et al., 2020). Our detection limits complement the findings by Wahhaj et al. (2021), who excluded the presence of hot-start planets more massive than $\gtrsim 3\,M_{\rm Jup}$ at two separations: 7.5 and 9.7 au.

3.4 The Age of the HR 8799 System

With our dynamical mass for HR 8799 e, we infer the first cooling age for the planet and thereby the system. Prior to Zuckerman et al. (2011) identifying it as a member of the Columba association (42^{+6}_{-4} Myr; Bell et al. 2015), its age was only loosely constrained (Marois et al., 2008). Lee & Song (2019) recently suggested that it is actually a member of the younger β Pictoris moving group (BPMG, 24 ± 3 Myr; Bell et al., 2015). Moreover,

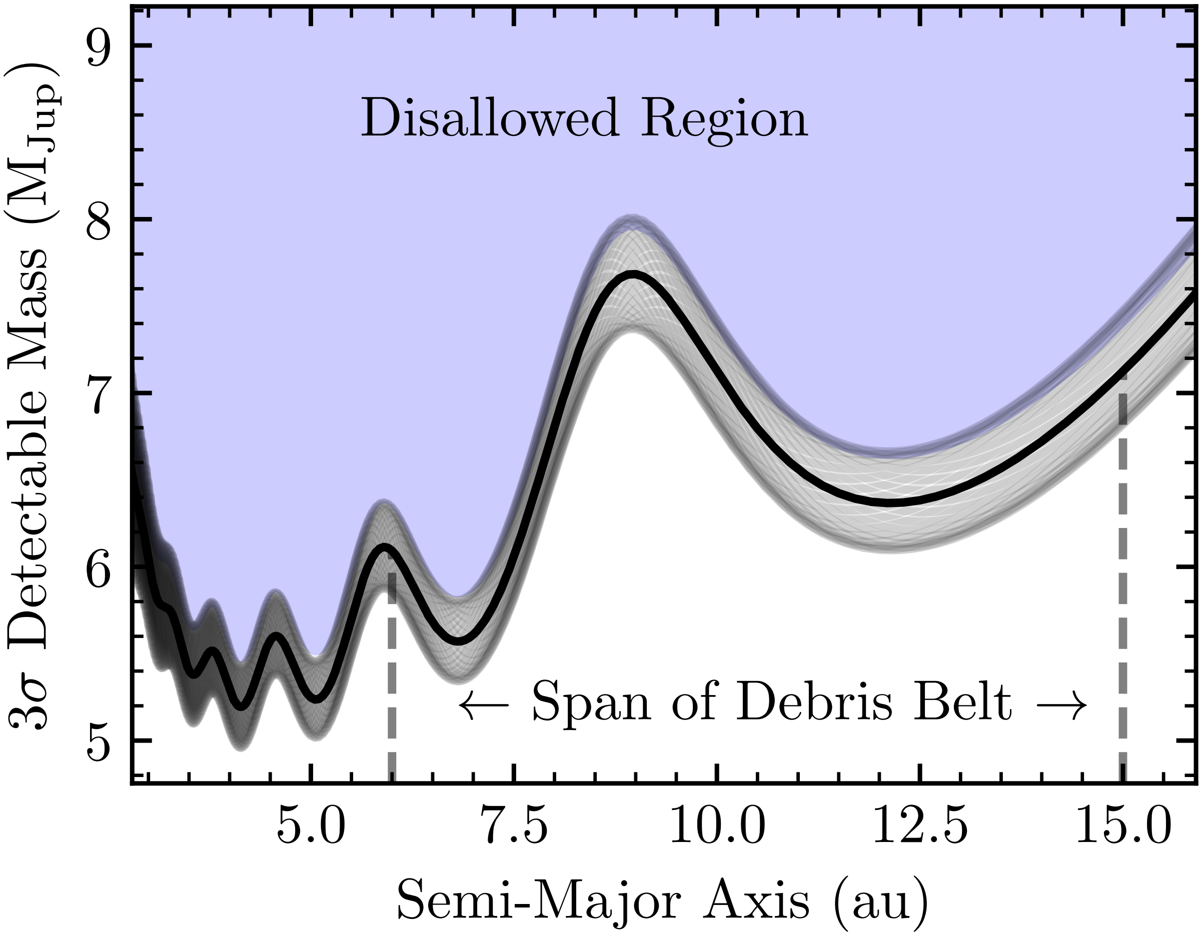

Figure 3.3 The minimum masses and semi-major axes of additional, unseen HR 8799 companions that would have been detected at 3σ using our HGCA *Gaia* EDR3-*Hipparcos* proper motion anomalies. The individual grey lines (forming a grey band together) show the 3σ limits assuming a range of argument of periastron (ω) from 0 to 2π. The black line is the 3σ limit averaged over the all possible ω. Planets lying in the blue "disallowed region" are excluded with at least 99.7% confidence, regardless of their orbital phase. The approximate range of the inner debris belt is indicated by vertical dashed lines (6 to 15 au; Frantseva et al., 2020).

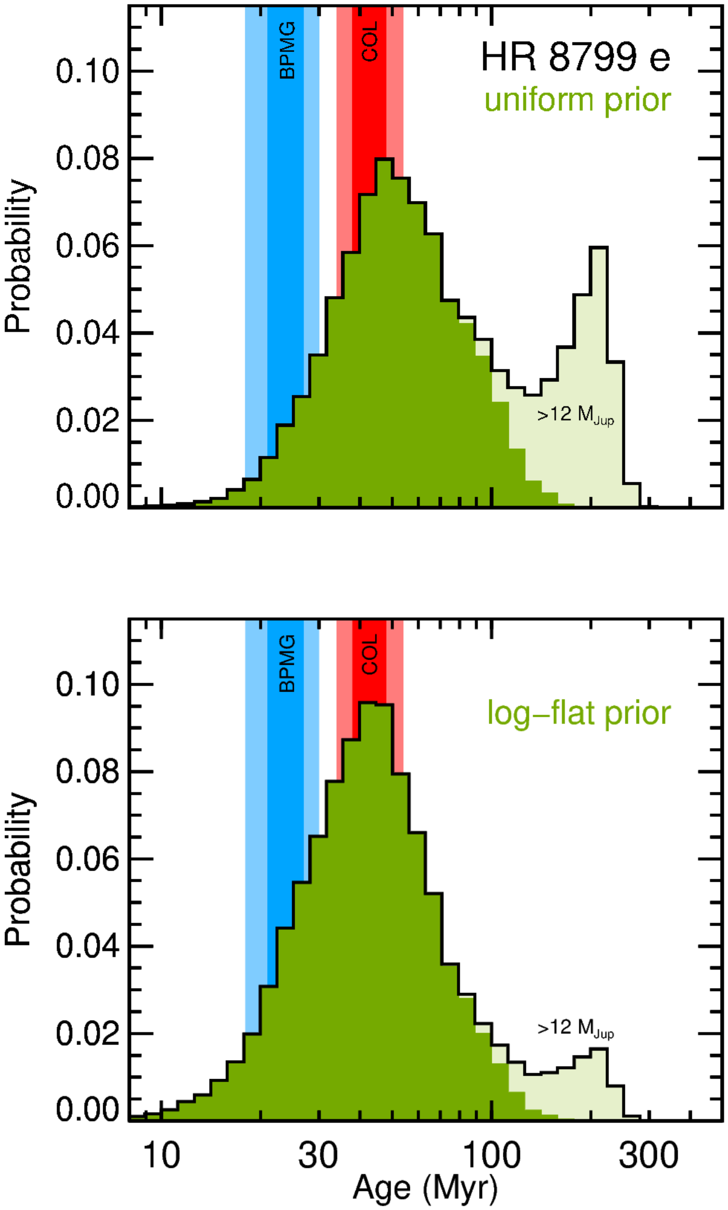

Figure 3.4 Substellar cooling age for HR 8799 e derived from hot-start Saumon & Marley (2008) hybrid models using its observed luminosity and dynamical mass, with either a uniform (top panel) or log-flat (bottom panel) prior on age. The 1σ and 2σ age ranges for the Columba association (red) and the BPMG (blue) are displayed for comparison, and our cooling age is consistent with both. The peak at ≈200 Myr corresponds to the 10% of our mass posterior above 12 $M_{\rm Jup}$ (lighter shading). The maximum likelihood age is 40–50 Myr regardless of the age prior.

HR 8799 was one of four (out of 23) Columba members that Gagné et al. (2018) chose not to use in their BANYAN Σ model due to being outliers, despite it still being considered a bona fide member.

We perform a rejection-sampling analysis using mass and $L_{\rm bol}$, in a similar fashion as Dupuy & Liu (2017) and Brandt et al. (2021, submitted), to derive a hot-start cooling age for HR 8799 e. The HR 8799 planets are too luminous to be consistent with the very low initial entropies predicted by the Marley et al. (2007a) cold-start models (Marleau & Cumming, 2014). Warm- and hot-start scenarios are allowed by the data, and simulations from Berardo et al. (2017), Berardo & Cumming (2017), and Marleau et al. (2019) tend to favor the hot-start scenario in general.

We randomly draw masses from our posterior distribution and ages distributed uniformly or log-flat, then bi-linearly interpolate the evolutionary model grid and compute a test $L_{\rm bol}$. We accept or reject trials in a Monte Carlo fashion depending on how well the trials agree with the observed $L_{\rm bol}$.

We derive a new $L_{\rm bol}$ for HR 8799 e using the absolute magnitude–$L_{\rm bol}$ relations of Dupuy & Liu (2017), the SPHERE photometry from Wahhaj et al. (2021), and the K-band spectrum from Gravity Collaboration et al. (2019). Although the relations of Dupuy & Liu (2017) are derived from field dwarfs, Figure 12 of Filippazzo et al. (2015) demonstrates that young and field objects share the same K-band bolometric corrections within the 0.25 mag scatter of their relations; we adopt this scatter as our uncertainty. We find $K_{\rm MKO} = K_{\rm S,2MASS} = 16.00 \pm 0.02\,{\rm mag}$ and $\log(L_{\rm bol}/L_{\odot}) = -4.52 \pm 0.10\,{\rm dex}$, which is consistent with but twice as precise as the measurement by Marois et al. (2010b).

Figure 3.4 shows posterior distributions of the system age for two different age prior choices (uniform and log-flat). Both posteriors peak at 40–50 Myr but differ at the young and old extremes. The choice of prior significantly affects the old extreme of the posterior. However, under either prior, Columba's age agrees well and BPMG's is consistent (1.2–1.7σ).

Substellar cooling alone does not preclude older ages ($\gtrsim$100 Myr), which have been shown to yield unstable orbits at the correspondingly higher masses ($> 12\,M_{\rm Jup}$). The high end of our mass posterior yields a smaller age peak at $\approx$200 Myr that corresponds to a resurgence in luminosity at older ages due to deuterium fusion. Even though 9.8% of our dynamical mass posterior for HR 8799 e lies above $12\,M_{\rm Jup}$ (where deuterium burning is possible), the planets of the HR8799 system have not been considered deuterium-fusing objects, with masses below $\approx 13\,M_{\rm Jup}$, based on their luminosity and hypothesized youth (Marois et al., 2008; Marois et al., 2010b). As well, they are unlikely to have masses in excess of $13\,M_{\rm Jup}$ on the basis of stability (Pueyo et al., 2015; Wang et al., 2018).

Excluding masses above $12\,M_{\rm Jup}$ yields an age distribution that is approximately Gaussian in $\log t$ for both priors (see dark-shaded posteriors in Figure 3.4). The resulting age posterior (under the log-flat prior) is $\log(t/{\rm yr}) = 7.62 \pm 0.20$ dex (42^{+24}_{-16} Myr).

If the HR 8799 system is indeed a BPMG member, it would be coeval and perhaps co-compositional with the giant planets β Pic b and c that have dynamical masses of $9.3^{+2.6}_{-2.5}\,M_{\rm Jup}$ and $8.3 \pm 1.0\,M_{\rm Jup}$, respectively (Brandt et al., 2021a). This cannot be ruled out by the dynamical masses. HR 8799 e and β Pic c have the same K-band absolute magnitude within the errors, 12.94 ± 0.02 mag and 12.9 ± 0.1 mag, respectively, and their masses are also consistent at 0.6σ.

The above discussion assumes that hot-start models are appropriate for deriving a substellar cooling age. If instead there was significant entropy loss in the formation of HR 8799 e, then it would be younger. Perhaps the initial entropy could even be tuned to match the age of the BPMG in a warm-start scenario. A younger age could also compensate for higher masses when considering the system's long-term stability.

3.5 Conclusions

In this letter, we determine a dynamical mass for HR 8799 e of $9.6^{+1.9}_{-1.8}\,M_{\rm Jup}$ by assuming that planets c, d and e share the same mass to within ≈20%. Marginalizing over a larger range of mass ratios for all four planets yields a dynamical mass of $9.4^{+2.2}_{-2.1}\,M_{\rm Jup}$ for HR 8799 e. We favor the more precise mass for HR 8799 e given that the planets' similar spectra and luminosities strongly suggest similar masses.

Our dynamical mass for HR 8799 e is $2\,M_{\rm Jup}$ (1.2σ) higher than previous estimates based on hot-start models (e.g., $7.2^{+0.6}_{-0.7}\,M_{\rm Jup}$; Wang et al., 2018). We rule out, with 99.7% confidence, any planets with masses greater than $\approx 6 M_{\rm Jup}$ and semi-major axes between ≈3 au and ≈ 8 au, as well as any additional $7\,M_{\rm Jup}$ or larger planets between 8 and 16 au.

We compute an updated bolometric luminosity for HR 8799 e and use hot-start evolutionary models to derive a substellar cooling age. We find 42^{+24}_{-16} Myr if we exclude the high-mass ($>12\,M_{\rm Jup}$) portion of our mass posterior, based on the low luminosity of HR 8799 e and the stability analysis of Wang et al. (2018). This is consistent with both the Columba association and β Pictoris moving group. Notably, the masses and absolute magnitudes of HR 8799 e and β Pic c are consistent within $<1\sigma$.

HR 8799 e, as the innermost planet on a ≈50-year period, induces about 75% of the proper motion anomaly over the ≈25-year *Hipparcos*-*Gaia* baseline. The uncertainty in our dynamical mass is dominated by the *Gaia* proper motion precision of HR 8799. Improved astrometric precision in future *Gaia* data releases will translate directly to improved mass measurements for the HR 8799 planets, especially for HR 8799 e and d.

Chapter 4

Improved Dynamical Masses for Six Brown Dwarf Companions

4.1 Outline

In this chapter, we present comprehensive orbital analyses and dynamical masses for the substellar companions Gl 229 B, Gl 758 B, HD 13724 B, HD 19467 B, HD 33632 Ab, and HD 72946 B. Our dynamical fits incorporate radial velocities, relative astrometry, and most importantly calibrated *Hipparcos-Gaia* EDR3 accelerations. For HD 33632 A and HD 72946 we perform three-body fits that account for their outer stellar companions. We present new relative astrometry of Gl 229 B with Keck/NIRC2, extending its observed baseline to 25 years. We obtain a <1% mass measurement of $71.4 \pm 0.6\,M_{\rm Jup}$ for the first T dwarf Gl 229 B and a 1.2% mass measurement of its host star $(0.579 \pm 0.007\,M_{\odot})$ that agrees with the high-mass-end of the M dwarf mass-luminosity relation. We perform a homogeneous analysis of the host stars' ages and use them, along with the companions' measured masses and luminosities, to test substellar evolutionary models. Gl 229 B is the most discrepant, as models predict that an object this massive cannot cool to such a low luminosity within a Hubble time, implying that it may be an unresolved binary. The other companions are generally consistent with models, except for HD 13724 B that has a host-star activity age 3.8σ older than its substellar cooling age. Examining our results in context with other mass–age–luminosity benchmarks, we find no trend with spectral type but instead note that younger or lower-mass brown dwarfs are over-luminous compared to models, while older or higher-mass brown dwarfs are under-luminous. The presented mass measurements for some companions are so precise that the stellar host ages, not the masses, limit the analysis.

The six systems presented all have directly imaged BD companions on ≈15 to 500 year orbital periods, long-term precision RVs, and significant astrometric accelerations (Brandt, 2018, 2021). The contents of this chapter have been published in the Astrophysical Journal Letters, under the title "Improved Dynamical Masses for Six Brown Dwarf Companions Using Hipparcos and Gaia EDR3" (Brandt et al., 2021e).

4.2 Stellar Ages

We denote posteriors by m^{+u}_{-l}, where u and l give the 68.3% confidence interval about the median value m. We report $m \pm 1\sigma$ if u and l are approximately equal within the quoted precision. HGCA v.EDR3 refers to the *Gaia* EDR3 version of the catalog, and HGCA v.DR2 to the original *Gaia* DR2 version (Brandt, 2018).

Five of our six targets are main sequence, approximately solar-mass stars: Gl 758 A, HD 13724 A, HD 19467 A, HD 33632 A, and HD 72946 A. In this Section we present uniform analyses of their ages based on activity and rotation. Stellar ages, and therefore companion ages, will enable us to compare BD observables with predictions from evolutionary models at our measured dynamical masses.

A star's age can be constrained with gyrochronology; G and K dwarfs lose angular momentum through their magnetized winds as they age (Soderblom, 2010; Ahuir et al., 2020). Activity indices tied to stellar rotation constrain the Rossby number and thereby the age. We convert the Rossby number to a rotation period using the convective overturn time given in Noyes et al. (1984). Finally, we convert the rotation period to an age according to the calibration of Mamajek & Hillenbrand (2008).

We adopt the Bayesian activity-age method that is described in detail in Brandt et al. (2014), and further explained in Li et al. (2021). Our method is identical to the latter work, but we summarize it here and the data involved. We use both the chromospheric activity index $R'_{\rm HK}$ and the X-ray activity index R_X to infer a Rossby number. The R_X measurements come from the *ROSAT* all-sky survey catalogs (Voges et al., 1999, 2000). Some stars have only upper limits on X-ray fluxes; we compute these as 5σ values assuming the uncertainty from the nearest detection in the ROSAT faint source catalog (Voges et al., 2000). The Ca II S-indices are from Pace (2013) and references therein

Table 4.1. Input stellar parameters for the Bayesian age analyses.

Identifier	B_{T} (mag)	$\sigma_{B_{\mathrm{T}}}$ (mag)	V_{T} (mag)	$\sigma_{V_{\mathrm{T}}}$ (mag)	$\log R_X$ (dex)	$\log R'_{\mathrm{HK}}$ (dex)	P_{rot} (days)
Gl 758	7.374	0.015	6.447	0.010	< -5.04	-5.05	$\cdots$
HD 19467	7.788	0.015	7.043	0.010	< -4.75	-4.97	$\cdots$
HD 13724	8.712	0.017	7.948	0.012	< -5.13	-4.78	21, 25.76
HD 33632A	7.102	0.015	6.530	0.010	-5.55	-4.83	$\cdots$
HD 72946	7.933	0.017	7.159	0.011	-4.80	-4.68	$\cdots$

Note. — $<$ denotes a 5σ upper bound on R_X.

(most sources have multiple measured S-indices). The method of Brandt et al. (2014) uses the average of the maximum and minimum S-indices found in the literature (in the Mt. Wilson system). We convert these indices to Mt. Wilson R'_{HK} with the relations from Noyes et al. (1984). The R'_{HK} and R_X values are tabulated in Table 4.1.

In Table 4.1, each source's B_{T}-band and V_{T}-band magnitude comes from the Tycho-2 catalogue (Høg et al., 2000). We denote the errors on the magnitudes with σ, e.g., $\sigma_{V_{\mathrm{T}}}$. We convert the B and V Tycho filters to Johnson B and V, using the transformations from Volume 1 of ESA, 1997. We then use B-V and the activity indices to deduce a stellar age as in Brandt et al. (2014), providing a stellar rotation period when available. Only HD 13724 A has measured periodic, photometric variability, with rotation periods ranging from 21 days (Arriagada, 2011) to 25.76 days (Oelkers et al., 2018). Directly measured rotation periods do not require estimates of the Rossby number or convective overturn time and enable tighter constraints on the stellar age (e.g. Mamajek & Hillenbrand, 2008;

Table 4.2. Posterior values from the Bayesian stellar age estimates.

Identifier	Age Posterior (Gyr)	Notes
Gl 758 A	$8.3^{+2.7}_{-2.1}$	⋯
HD 13724 A	$2.8^{+0.5}_{-0.4}$	*a*
HD 13724 A	$3.6^{+0.6}_{-0.5}$	*b*
HD 13724 A	$3.1^{+0.9}_{-0.7}$	*c*
HD 72946 A	$1.9^{+0.6}_{-0.5}$	⋯
HD 33632 A	1.7 ± 0.4	⋯
HD 19467 A	$5.4^{+1.9}_{-1.3}$	⋯

[a] This is our fiducial case using the Arriagada (2011) 21-day rotation period.

[b] This uses the 25.76 day rotation period from Oelkers et al. (2018).

[c] This estimate does not involve a rotation period as an input parameter.

Brandt et al., 2018). We incorporate these rotation periods as described by Brandt et al. (2014).

The resulting stellar age posteriors are shown in Figure 4.1. We tabulate the median and 68.3% confidence intervals in Table 4.2. In the following subsection, we compare our age estimates with other results in the literature.

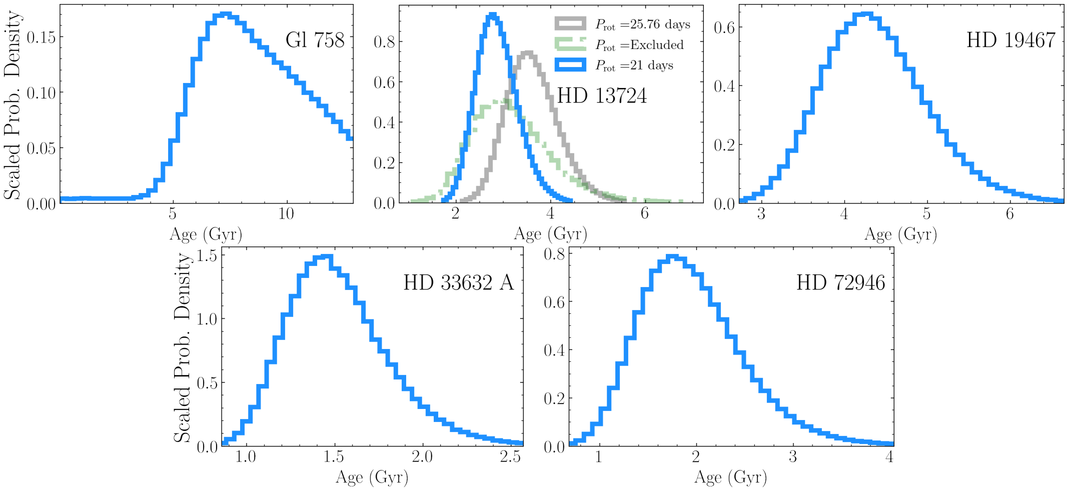

Figure 4.1 Posteriors from the Bayesian stellar age analyses using the method of Brandt et al. (2014). The median values and confidence intervals are listed in Table 4.2. The parameters used in the analyses are listed in Table 4.1. HD 13724 (middle panel of the top row) had three stellar rotation periods considered: 21 days from Arriagada (2011), 25.76 days from Oelkers et al. (2018), and a case where period information was neglected (labelled "Excluded").

4.2.1 Discussion on the ages of individual stars

The ages and masses of our six BD host stars have been extensively studied (see for instance Casagrande et al. 2011, Gomes da Silva et al. 2021 and references therein). Here, we place our results within the context of previous age estimates. We begin with a discussion of the age of Gl 229. As an early-M dwarf, we excluded it from our re-analysis.

Gl 229 The ages of M dwarfs, like Gl 229, are hard to determine because of their extremely long main-sequence lifetimes (see, e.g., West et al. 2008). Brandt et al. (2019) suggested an age of 2.6 ± 0.5 Gyr based on stellar activity but noted that the activity-age relation is poorly calibrated for M dwarfs. They ultimately adopted a pair of wide uniform priors on the age (considering ages between 1 and 10 Gyr). In Section 4.8, we reconsider the age in light of our new dynamical mass and adopt a prior uniform between 1 and 10 Gyr. This is to highlight the significant disagreement between modern models and Gl 229 B's high mass, at all reasonable ages.

Gl 758 This G8V star (Maldonado et al., 2012) is favored to be old. Brandt et al. (2018) inferred an activity-based age (using the same Brandt et al. 2014 method as we do now) that favored old ages $\gtrsim$6 Gyr with a long tail to $\approx$13 Gyr. We infer here a nearly identical posterior of $8.3^{+2.7}_{-2.1}$ Gyr and between 0.7 Gyr and 11.5 Gyr with 99.7% confidence. This age is broadly consistent with all values in the literature. Casagrande et al. (2011) found the age between 4.53 and 12.06 Gyr (16% and 84% confidence intervals) with Padova isochrones (Bertelli et al., 2008, 2009) and a significantly older age between 8.5 and 13.4 Gyr using BASTI isochrones (Pietrinferni et al., 2004, 2006, 2009). Pace (2013) adopted an age of 11.16 ± 2.46 Gyr, albeit based on Casagrande et al. (2011). More recently, Luck (2017) re-examined the age of Gl 758 and explored the best-fit age using a wide variety of isochrones. They inferred 6.42 Gyr using the earlier, Bertelli et al. (1994) isochrones; 5.31 Gyr using the Demarque et al. (2004) isochrones that implemented (at the time) an improved prescription for convective core-overshoot; and 7.72 Gyr with the Dartmouth stellar evolution database (Dotter et al., 2008). Most recently, Bowler et al. (2020) argued for minimum and maximum ages of 6 and 10 Gyr, respectively, from various age determinations in the literature. We adopt the age prior shown in Figure 4.1 for the BD model analysis in Section 4.8.

HD 13724 HD 13724 A is a G3/G5V dwarf (Kharchenko, 2001). HD 13724 A's measured $R'_{\rm HK}$ and R_X, combined with the 21-day period from Arriagada (2011) (consistent with the most-recent 20.2±1.2 day period derived by Rickman et al., 2019), yield a precise age of $2.8^{+0.5}_{-0.4}$ Gyr. We use this age in our comparisons to BD models in Section 4.8.

We infer a slightly older age of $3.6^{+0.6}_{-0.5}$ Gyr if we instead adopt the 25.76 day rotation period from Oelkers et al. (2018). The measured rotation periods favors old ages, however, the star's activity (neglecting any rotation period information) gives a similarly old age of $3.1^{+0.9}_{-0.7}$ Gyr. All three of our age estimates are much older than the 1.04 ± 0.88 Gyr found by Rickman et al. (2020), which was inferred from grids of Geneva stellar models (Ekström et al., 2012; Georgy et al., 2013).

There is a lack of consensus within the literature on the age of HD 13724, albeit younger ages seem to be favored. Most recently, Gomes da Silva et al. (2021) report a posterior of 0.47 ± 0.36 Gyr, and Delgado Mena et al. (2019) report an age posterior of 1.11 ± 0.98 Gyr. Results from Casagrande et al. (2011) are more consistent with our analyses that include the rotation period. Casagrande et al. (2011) use the Padova isochrones to constrain the age to between 0.94 and 5.51 Gyr (16% and 84% confidence intervals)— a wide range that encompasses all of the aforementioned age posteriors. Their 5% and 95% confidence intervals on the age are 0.27 and 7.18 Gyr. Casagrande et al. (2011) found similarly wide posteriors using the BASTI isochrones. Stanford-Moore et al. (2020) infer an old age, similar to our own, based on stellar activity: centered on 5 Gyr, and between 1.4 and 12 Gyr with 95% confidence. If HD 13724 is young, it is an unusually inactive star and a slow rotator for its age.

HD 13724 could have an anomalously high surface metallicity that skews stellar-evolution inferred ages to young values (gravitational settling depletes surface metallicity as Solar-type stars age; Thoul et al., 1994). However, we show in Section 4.8 that the BD age constraints favor a ≈1 Gyr age for HD 13724 A (close to that assumed by Rickman et al. 2020), and that our rotation period-informed age of $2.8^{+0.5}_{-0.4}$ Gyr is 3σ inconsistent with the inferred BD age. This would make its rotation period of 20-30 days much slower than that expected from gyrochronology, and render HD 13724 A an interesting test case for gyrochronology in G dwarfs.

HD 19467 This G3V star (Gomes da Silva et al., 2021) has only an upper bound on its X-ray activity index and a chromospheric activity slightly less than Solar (Isaacson & Fischer, 2010; Gomes da Silva et al., 2021). The activity of this solar-type star (e.g., Mints & Hekker 2017) points to HD 19467 A being nearly a solar twin. We infer an activity-age of $5.4^{+1.9}_{-1.3}$ Gyr, with a 95% confidence intervals of 3.4 to 9.2 Gyr. This agrees well with the gyrochronology estimate of 5.6 ± 0.8 Gyr derived by Maire et al. (2020a) from ASAS photometry.

Our activity-based age is slightly younger than most isochronal estimates in the literature, but generally consistent within 1–2σ. For example, Casagrande et al. (2011) report 8.7 ± 3.4 Gyr, and Lorenzo-Oliveira et al. (2018) give an activity age of 8.8 ± 0.3 Gyr. However, some estimates prefer even older ages that would be modestly inconsistent with our analysis, e.g., 10.5 ± 1.9 Gyr by Aguilera-Gómez et al. (2018). Our age agrees with the best-fit activity age of 6.18 Gyr by Isaacson & Fischer (2010). For contrast, the recent dynamical analysis by Maire et al. (2020a) adopted an age of 8^{+2}_{-1} Gyr, favoring isochronal estimates. The median of that estimate is older than what we adopt but within our 95% confidence interval.

HD 33632 A HD 33632 A is an F8V star (Anderson & Francis, 2012) that is similarly as active as the Sun (Pace, 2013; Egeland et al., 2017). HD 33632 A may be slightly more massive than the Sun (e.g., $1.01 \pm 0.05\, M_\odot$ from Mints & Hekker 2017; $1.03 \pm 0.04\, M_\odot$ from Ramírez et al. 2012; $1.10\, M_\odot$ from Casagrande et al. 2011). The activity of HD 33632 A implies a young age; magnetic braking has not yet slowed the star significantly. Isaacson & Fischer (2010) estimated, from activity, a fast rotation period of 9 days. Combining the R_X activity index and the chromospheric activity, we infer an age of 1.7 ± 0.4 Gyr for HD 33632 A.

Other activity age estimates range from 2–5 Gyr, e.g., 3.5 Gyr from Isaacson & Fischer (2010) and $3.9^{+3.3}_{-1.8}$ Gyr from Stanford-Moore et al. (2020). Isochronal ages favor $\gtrsim 2$ Gyr; with posteriors that are consistent with 1.5 Gyr. For instance, Ramírez et al. (2012) determined a best-fit age of 4.8 ± 1.4 Gyr. The analyses by Casagrande et al. (2011) found maximum likelihood ages for HD 33632 A of 2.2 Gyr and 3.2 Gyr using Padova and BASTI ischrones, respectively. The 16% and 84% confidence interval ages were 1–4.15 Gyr with Padova and 1.5–4.5 Gyr using BASTI — both fully consistent with our 1.7 ± 0.4 Gyr age estimate. The abundance of neutron capture elements provides age estimates of HD 33632 A near ≈1.5–2.5 Gyr (Spina et al., 2017; Currie et al., 2020). The recent analysis of HD 33632 Ab by Currie et al. (2020) adopted an age prior of $1.5^{+3.0}_{-0.7}$ Gyr, fully consistent with our 1.7 ± 0.4 Gyr age.

HD 72946 This G5V star (Kharchenko et al., 2007) has a variety of age constraints, including from isochrones and lithium abundances (Ramírez et al., 2012; Luck, 2017; Aguilera-Gómez et al., 2018), that place it anywhere from 0.5 Gyr to as old as ≈9 Gyr. There is no measured photometric rotation period for the star, but there are measurements of the X-ray emission (Voges et al., 1999) and of the chromospheric activity (Pace, 2013; Bouchy et al., 2016). The latter indicates a star marginally more active, and therefore younger, than the Sun. Combining the X-ray and chromospheric activity indices, we infer an age of $1.9^{+0.6}_{-0.5}$ Gyr from our Bayesian analysis. This younger age is consistent with estimates in the literature, albeit literature estimates span a wide range.

Maire et al. (2020b) is the most similar (to our method) and the most recent age analysis. Maire et al. (2020b) used the Mamajek & Hillenbrand (2008) relations with an average chromospheric activity ($\log R'_{\rm HK} = -4.60$ dex, which is slightly more active and thereby younger than our adopted Gray et al. (2003); Pace (2013) index of -4.68 dex) to infer a 15-day rotation period, implying an age near 1 Gyr. However, using an average projected rotational activity, they placed a more stringent upper bound on the rotation period of 12 days, excluding ages older than 1 Gyr, or ≈1.5 Gyr given liberal uncertainties. They ultimately chose to adopt 0.8–3 Gyr as the range of probable ages, which is in excellent agreement with our inferred $1.9^{+0.6}_{-0.5}$ Gyr age.

For further comparison, Ramírez et al. (2012) derived an isochronal age of $4.9^{+4.6}_{-2.1}$ Gyr.

Casagrande et al. (2011) inferred an age between 1.09 and 9.27 Gyr (16% and 84% confidence intervals) with the Padova isochrones. They found similar results, 1.20 and 9.64 Gyr, with BASTI isochrones. Aguilera-Gómez et al. (2018) report an isochronal age of $8.7^{+2.6}_{-4.2}$ Gyr — favoring an age much older than our estimate and that by Maire et al. (2020b) but still marginally consistent with both estimates.

4.3 Radial Velocities and Relative Astrometry

All six systems have both direct imaging of the BD companions and radial velocity (RV) measurements of the host star. In this Section, we summarize the direct imaging and RV data for each of the sources and present new Keck/NIRC2 imaging of Gl 229 B. Table 4.3 lists the sources of the relative astrometry we use to fit each system.

We retrieve the RV data for every source from Vizier.[1] For many of the sources, a large fraction of the RVs come from the HIRES instrument on Keck (Vogt et al., 1994), originally published by Butler et al. (2017). We use the recently recalibrated HIRES data from Tal-Or et al. (2019). Many other RV measurements come from the HARPS instrument at the European Southern Observatory (ESO) La Silla 3.6-m telescope (Mayor et al., 2003). We use the recently recalibrated HARPS data produced by Trifonov et al. (2020). For every source presented in this work except for HD 72946, the RVs do not cover a full orbital period. *Hipparcos-Gaia* absolute astrometry thus plays a crucial role in constraining the companion's mass and orbit.

---

[1] https://vizier.u-strasbg.fr/viz-bin/VizieR

Table 4.3. Summary of the relative astrometry that we use in our orbital analyses.

Identifiers	Reference	Measurements[a]
Gl 229 A/B	TB20	7
Gl 229 A/B	Table 4.5	2
Gl 758 A/B	BB18	4
HD 13724 A/B	R20	9
HD 19467 A/B	TRENDSV	5
HD 19467 A/B	C15	1
HD 19467 A/B	BB20	1
HD 72946 A/B	M20	2
HD 72946/ HD 72945	*Gaia* EDR3	1
HD 33632 A/B	*Gaia* EDR3	1
HD 33632 A/Ab	Table 4.6	1

Note. — (A/B) refers to relative astrometry between A and B. For example, HD 72946 A/B refers to relative astrometry of HD 72946 B about HD 72946 A. The data reference points to the publication where the data are retrievable either in print or through a data source (e.g., Vizier) clearly linked to that publication. We do not reproduce the data here so that data remain consolidated within their original published source.

[a] The number of pairs of position angle/separation measurements.

References. — *Gaia* EDR3 – Lindegren et al. (2020), M20 – Maire et al. (2020b), TB20 – Brandt et al. (2019), BB18 – Bowler et al. (2018), R20 – Rickman et al. (2020), TRENDSV – Crepp et al. (2014), C15 – Crepp et al. (2015a), BB20 – Bowler et al. (2020)

Table 4.4. Summary of the additional Gl 229 RVs from HIRES.

Epoch BJD	RV m/s	RV error m/s
2458116.862	8.72	1.23
2458117.852	8.71	1.24
2458396.142	6.98	1.18
2458777.037	16.16	1.08
2458794.994	4.30	1.14
2458880.798	15.76	1.05
2458907.830	12.75	0.95
2459101.122	6.79	1.15
2459267.794	19.94	1.03

References. — Rosenthal et al. (2021), A. Howard, priv. commun.

4.3.1 Gl 229

We adopt HIRES and HARPS radial velocities using the recently calibrated data sets by Tal-Or et al. (2019) and Trifonov et al. (2020), respectively. The combined RV data set consists of 248 observations spanning twenty years. We add nine new HIRES observations (Rosenthal et al., 2021) of Gl 229 A, spanning 2018 through early 2021. These additional HIRES RV data are summarized in Table 4.4. The additional three years of RVs show slight curvature in the RV time series of Gl 229 A; this curvature is consistent with that expected from the previous best fit orbits of Gl 229 B.

We use the relative astrometry from Brandt et al. (2019) that consists of six observations between 1995 and 2000 using the Wide Field and Planetary Camera 2 (WFPC2) aboard the *Hubble Space Telescope* (*HST*) and one 2012 observation from the Subaru telescope with HiCIAO (Suzuki et al., 2010). As suggested in Brandt et al. (2019), we double the formal PA errors on the 1995 November and 1996 November *HST* observations (two epochs). These two epochs used different guide stars than the other four *HST* epochs.

We also present new relative astrometry of Gl 229 B, extending the direct imaging baseline to twenty-five years. We observed Gl 229 on 2020 October 24 UT and 2021 January 5 UT with NIRC2 in narrow-camera mode and the natural guide star adaptive optics system at the Keck II telescope (Wizinowich et al., 2000; Johansson et al., 2008). In order to obtain high-S/N, unsaturated images of both the host star and companion, we alternated taking shallow and deep exposures. All data were taken using an 864×120-pixel subarray to reduce the minimum allowable exposure time. The images of Gl 229 B were obtained with an exposure time per coadd of 0.5 s, 100 coadds, and the CH_4s filter ($\lambda_C = 1.592\,\mu$m and $\Delta\lambda = 0.126\,\mu$m). For unsaturated images of Gl 229 A, we used different filters and exposure times at the two epochs. At the first epoch, the image quality was poorer, so we used the CH_4s filter, exposure time per coadd of 0.01 s, and 100 coadds. At the second epoch, we used the narrower $H_{\rm cont}$ filter ($\lambda_C = 1.580\,\mu$m and $\Delta\lambda = 0.023\,\mu$m), exposure time per coadd of 0.5 s, and 100 coadds.

In shallow images, Gl 229 A is unsaturated while the companion is undetected. In deep images, the companion is clearly resolved while the primary is saturated (though the wings of the star's point-spread-function, PSF, are usable). This poses a challenge in measuring the separations of the system. Furthermore, the adaptive optics (AO) corrections for the observations are imperfect and time-varying, especially for the images observed in October 2020.

To obtain relative astrometry, we implement a least-squares PSF-fitting algorithm that uses the unsaturated PSFs of Gl 229 A as templates to fit for the positions of both Gl 229 A and Gl 229 B in deep images. Gl 229 A is saturated in the deep images. We mask hot and saturated pixels and fit the outer wings and speckles of Gl 229 A's diffraction pattern. Figure 4.2 shows three representative PSFs, the masked pixels, and the residuals from this procedure. The outer speckles are sufficiently well-measured in the shallow images that they can centroid well the saturated PSF.

For every deep image, we use all shallow images from the same night to fit for a relative separation and position angle (PA) of the system. Thus for every image where the companion is detected, we obtain $\approx$15–20 templates that allow us to estimate the mean and uncertainty of the result. This gives very precise relative offsets in detector coordinates x and y ($\sim$0.01 pix).

We convert from detector coordinates into sky coordinates using the same method as Dupuy & Liu (2017) and Bowler et al. (2018). This accounts for differential atmospheric refraction and aberration. We correct for differential chromatic refraction, although this effect is negligible compared to the uncertainties. We use the calibration of Service et al. (2016) to correct for distortion; we subtract $0\overset{\circ}{.}262 \pm 0\overset{\circ}{.}002$ from the PA of the y-axis of NIRC2 given in the header; and account for the pixel scale and its uncertainty

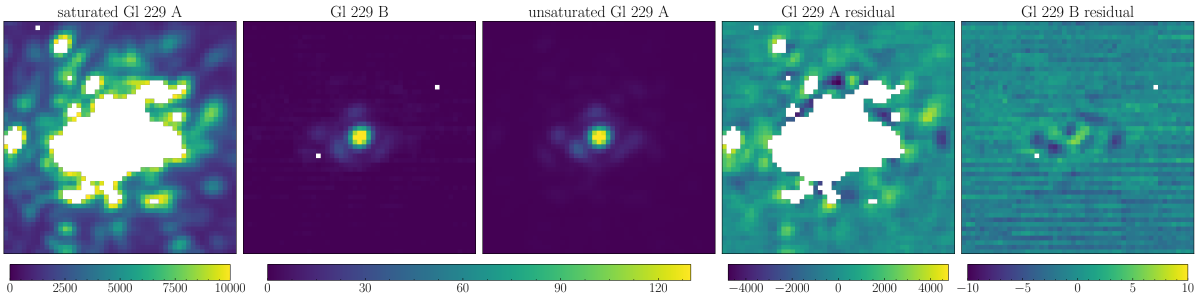

Figure 4.2 PSF fitting for Gl 229 A/B. We use the unsaturated PSF of Gl 229 A (third panel from left) as a template to fit Gl 229 B (second from left) and the saturated PSF of Gl 229 A (left). We mask saturated and hot pixels; these are shown in white. The right two panels show the corresponding residuals.

Table 4.5. New Keck/NIRC2 relative astrometry of Gl 229 A/B.

Date (UT)	Sep (mas)	$\sigma_{\rm Sep}$ (mas)	PA (deg)	$\sigma_{\rm PA}$ (deg)	Filter
2020 Oct 24	4922.1	2.3	179.564	0.024	CH_4s
2021 Jan 5	4890.5	2.4	179.735	0.024	$H_{\rm cont}$ + CH_4s

(9.971 ± 0.004 mas pix^{-1}).[2] The 1.1 mas uncertainty in the Service et al. (2016) distortion solution implies a relative astrometry noise floor of 1.1 mas in separation and 0°.013 in PA for Gl 229 AB, which we add in quadrature to the other errors. Table 4.5 lists our relative astrometry for Gl 229 B. These measurements are only 2–3σ different than what the Brandt et al. (2019) orbital fit predicts.

4.3.2 Gl 758

We use the four epochs of Keck/NIRC2 direct imaging from Bowler et al. (2018). Like Bowler et al. (2018), we use RVs from the Automated Planet Finder (APF) at Lick

---

[2] https://github.com/jluastro/nirc2_distortion/wiki

Observatory, HIRES, and RVs from the Tull Coudé spectrograph (Tull et al., 1995). The only difference in this RV data set between our analysis and that of Bowler et al. (2018), is that we are able to use the new calibrated HIRES RVs from Tal-Or et al. (2019). The entire RV data set from the three instruments consists of 526 measurements spanning nearly twenty years.

4.3.3 HD 13724

The companion to HD 13724 was first discovered with RVs by Rickman et al. (2019), using CORALIE (Queloz et al., 2000). Rickman et al. (2020) followed up with high-contrast imaging and measured the first dynamical mass for the companion. We adopt the same relative astrometry as Rickman et al. (2020). Like Rickman et al. (2020), we include HARPS and CORALIE radial velocities. The only difference in the HARPS dataset is that we use the newly calibrated Trifonov et al. (2020) data. The overall RV baseline of the combined HARPS and CORALIE dataset is roughly twenty years and comprises 170 measurements. We are unable to include ≈5 unpublished CORALIE RVs from ≈2020 that were shown in Rickman et al. (2020).

Any two RV instruments will almost never agree on a measure of the RV offset (also known as the RV zero point) of a star due to unique systematics in the data processing pipeline or instrument. Instrument upgrades and small changes in a data reduction pipeline can also perturb the RV zero point. The CORALIE instrument was upgraded in June 2007 (Ségransan et al., 2010) and again in November 2014. We follow Rickman et al. (2020) and Cheetham et al. (2018) and treat the CORALIE pre- and post-upgraded instruments as independent RV instruments, thereby splitting the CORALIE dataset in three: CORALIE-98 (before the 2007 upgrade), CORALIE-07 (between the 2007 and 2014 upgrades), and CORALIE-14 (after the 2014 upgrade). Accordingly, our fits to HD 13724 include four RV offsets: one for HARPS and one for each CORALIE state.

4.3.4 HD 19467

We adopt the relative astrometry used in the recent work by Maire et al. (2020a). This consists of 7 measurements total, spanning ≈8 years between 2011 and 2018, from Crepp et al. (2014), Crepp et al. (2015b), and Bowler et al. (2020).

The radial velocities of HD 19467 A consist of HARPS measurements calibrated by Trifonov et al. (2020) and HIRES measurements calibrated by Tal-Or et al. (2019). The RVs span a baseline of more than twenty years, although this is less than one-tenth of the orbital period of HD 19467 B.

4.3.5 HD 33632 A & B

HD 33632 Ab was discovered by Currie et al. (2020) with direct imaging from Subaru/CHARIS (Groff et al., 2013, 2015) and Keck/NIRC2. We adopt the same relative astrometry here. These data were presented in cartesian coordinates in Currie et al. (2020); we present them in polar coordinates in Table 4.6.

We use RVs from the Lick planet search with the Hamilton spectrograph (Fischer

Table 4.6. Relative astrometry of HD 33632 A/Ab.

Date (UT)	Sep (mas)	$\sigma_{\rm Sep}$ (mas)	PA (deg)	$\sigma_{\rm PA}$ (deg)	Instrument	Filter
2018-10-18	781	5	257.0	0.4	CHARIS	JHK
2018-11-01	774	5	256.7	0.4	NIRC2	L'
2020-08-31	746	5	262.8	0.4	CHARIS	JHK
2020-09-01	746	5	262.7	0.4	CHARIS	JHK

Note. — These are the same data first used in Currie et al. (2020).

et al., 2014). The RVs for HD 33632 A span roughly eleven years, a small fraction of the nearly 100-year period of HD 33632 Ab.

HD 33632 has a co-moving M dwarf companion HD 33632 B that is resolved in *Gaia* EDR3 (Scholz, 2016; Gaia Collaboration et al., 2021). The companion has a projected separation of $33\farcs99086 \pm 0\farcs00003$. We convert the correlated *Gaia* EDR3 positions into correlated relative astrometry (separation and PA). The resulting separations and PAs for HD 33632 B about HD 33632 A are in Table 4.7.

4.3.6 HD 72946

We use the two epochs of relative astrometry taken with VLT/SPHERE (Beuzit et al., 2019) presented in Maire et al. (2020a). HD 72946 has a co-moving stellar companion, HD 72945, at a separation of ≈10″ (≈250 au) (Gaia Collaboration et al., 2021). We convert the *Gaia* EDR3 absolute astrometry of HD 72946 and HD 72945 into relative

Table 4.7. Relative astrometry derived from *Gaia* EDR3 for wide stellar companions.

Date (UT)	Identifiers	Sep (mas)	σ_{Sep} (mas)	PA (deg)	σ_{PA} (deg)	Sep–PA corr.
2016 Jan 1	HD 33632 A/B	33990.86	0.03	20.34068	0.00011	−0.28
2016 Jan 1	HD 72946/ 72945	10044.18	0.09	204.76056	0.00061	0.74

Note. — All data are derived from Gaia Collaboration et al. (2021). The raw positions and correlations were fetched from the *Gaia* archive (`https://gea.esac.esa.int/archive/`).

astrometry following the same procedure as for HD 33632 A/B. The resulting separation and PA for HD 72945 about HD 72946 are in Table 4.7.

The RVs for HD 72946 come from the ELODIE (Baranne et al., 1996) and SOPHIE (Bouchy & Sophie Team, 2006) instruments as published by Bouchy et al. (2016). These RVs span roughly sixteen years — a full orbital period of HD 72946 B.

4.4 Host Star Astrometry

Absolute astrometry from *Hipparcos* and *Gaia* give powerful constraints on the masses of giant long-period companions. We use absolute astrometry of the host stars to measure the dynamical properties of the six systems with high precision. We follow the procedures described in Brandt (2018); Dupuy et al. (2019a); Brandt et al. (2021f,a), which are similar to those adopted by, e.g., Feng et al. (2019); Lagrange et al. (2019a, 2020). In brief, we use the proper motion anomalies between *Hipparcos*, *Gaia* EDR3 and the *Hipparcos*-*Gaia* long-term proper motion to measure the acceleration vector of the host star in the plane of the sky. The acceleration offers additional constraints on the dynamical properties of the companion (and particularly its mass).

We use calibrated *Gaia* EDR3 and *Hipparcos* astrometry from the *Hipparcos*-*Gaia* v.EDR3 catalog of accelerations, originally produced for *Gaia* DR2 by Brandt (2018). We use the HGCA because it rotates the *Hipparcos*, *Gaia*, and *Hipparcos*-*Gaia* proper motions into the same reference frame in order to make them suitable for orbit fitting. The HGCA also calibrates all uncertainties to produce Gaussian residuals with the expected variance.

HD 33632 A and HD 72456 have outer third bodies: HD 33632 B and HD 72945, respectively. We analyse both two-body (ignoring the outer stellar companion) and three-body orbital fits. `orvara` uses their proper motions and proper motion correlations to help constrain their orbit, both of which are available in the *Gaia* archive.[3] However, neither HD 33632 B nor HD 72945 are in the HGCA. We apply a proper motion error inflation of a factor of 2 for HD 33632 B (Cantat-Gaudin & Brandt, 2021) and 1.37 for HD 72945 (Brandt, 2021) to account for any low-level systematics. We correct for projection effects in the proper motion and apply the Cantat-Gaudin & Brandt (2021) magnitude-dependent correction, which aligns the proper motion of the Gaia EDR3 sources brighter than G=13 with the International Celestial Reference Frame. Both corrections are negligible compared to the inflated proper motion errors, but we include them for completeness. The final proper motions and errors for HD 33632 B are $-144.58 \pm 0.15\,\mathrm{mas\,yr^{-1}}$ in right-ascension and $-139.53 \pm 0.11\,\mathrm{mas\,yr^{-1}}$ in declination For

---

[3] `https://gea.esac.esa.int/archive/`

HD 72495, we use $-130.31 \pm 0.19\,\mathrm{mas\,yr^{-1}}$ in right-ascension and $-133.12 \pm 0.15\,\mathrm{mas\,yr^{-1}}$ in declination

4.5 Orbit Fitting

We use `orvara` to fit for the orbital parameters of each system. The code employs MCMC with `ptemcee` (Foreman-Mackey et al., 2013; Vousden et al., 2016). Absolute astrometry is processed and fit for the five astrometric parameters by `htof` (Brandt et al., 2021d; Brandt & Michalik, 2020) at each MCMC step. We use a parallel-tempered MCMC with 20 temperatures; for each temperature we use 100 walkers with at least 400,000 steps per walker, thinned at the end by at least a factor of 50. Our MCMC chains converge typically between 20,000 and 80,000 steps; we conservatively discard the first 75% of each chain as burn in and use the remainder for inference. The chains for Gl 229 and the three-body fits to HD 33632 and HD 72946 were run for two million steps to ensure convergence and that the full parameter space was explored. We use the same criteria presented in Brandt et al. (2021a) to verify the convergence of our chains. The convergence criteria include the Gelman-Rubin Diagnostic (Gelman & Rubin, 1992; Roy, 2019).

These aforementioned methods and analysis tools are nearly identical to those presented in Brandt et al. (2019), Currie et al. (2020), Brandt et al. (2021a), Brandt et al. (2021f), and Li et al. (2021). We fit either 9 or 16 parameters for either 2 or 3 bodies total, respectively. These are the six Keplerian orbital elements[4] for each companion plus its mass and an RV jitter to be added to the RV uncertainties. We use a single RV jitter per star rather than per instrument, attributing the jitter to stellar activity. Our results are consistent if we adopt a different jitter for each instrument. `orvara` marginalizes out each instrument's RV zero-point, parallax, and barycenter proper motion. We perform fits to HD 72946 and HD 33632 that include and exclude their widely-separated, co-moving stellar companions (i.e., we do both two and three-body fits to these two systems). `orvara`'s three-body approach was shown to be accurate in Brandt et al. (2021a) via a set of `REBOUND` validation tests. We refer the reader to Section 2.3 of Brandt et al. (2021a) for the discussion of the three-body approach.

4.5.1 Priors on Orbital Elements

We assume uninformative priors for all the orbital elements: uniform except for inclination i, where we assume the standard geometric prior, and with semi-major axis and companion mass where we assume log-flat priors. We adopt a log-flat prior on each RV jitter.

4.5.2 Priors on Stellar Masses

We assume uniform priors on the masses of HD 13724 A, Gl 758 A, and Gl 229 A. We adopt stellar evolution masses as priors on the primary mass of the other three systems

---

[4] `orvara` fits for $\sqrt{e}\cos\omega$, $\sqrt{e}\sin\omega$ instead of the eccentricity e and the argument of periastron ω directly.

(HD 33632, HD 19467, and HD 72946), for which the RV baseline is short. For these three systems, the constraints on the mass of the primary from any orbital fit are many factors worse than those known (even loosely) from stellar evolution. For instance, a completely uninformative fit to HD 19467 yields a posterior on the primary mass of $1.6 \pm 2 M_{\odot}$. This is a G3 dwarf, and so we know that the mass is near $1 M_{\odot}$ with much higher confidence than $\pm 2 M_{\odot}$. Adopting a prior informed by stellar evolution theory is appropriate. For similar reasons, we could adopt a prior on the primary mass of Gl 758, however, adopting a tight prior on the primary mass adds a negligible improvement to the inferred secondary mass.

We adopt the same Gaussian priors on the primary masses as were used in the most recent dynamical analyses of the systems. These are $1.1 \pm 0.1 M_{\odot}$ for HD 33632 A (Currie et al., 2020), $0.953 \pm 0.022 M_{\odot}$ for HD 19467 A (Maire et al., 2020a), and $0.986 \pm 0.027 M_{\odot}$ for HD 72946 A (Maire et al., 2020b). These choices enable direct comparisons of our results to the preceding orbital analyses. Moreover, they are consistent with isochronal mass estimates in the literature (compare and see Casagrande et al., 2011; Ramírez et al., 2012; Mints & Hekker, 2017). Our use of an informative prior ultimately has a negligible effect on our mass constraints for HD 19467 B and HD 33632 Ab. But it improves the precision of our inferred mass for HD 72946 B by a factor of ≈3.

We also adopt priors on the distant stellar companions in the three-body fits to the HD 72946 and HD 33632 systems in Section 4.6. As we show in Section 4.6, adding the third body (HD 72945 or HD 33632 B) does not change significantly the inferred parameters of the BD companion. However, in both cases adding the tertiary stellar body without placing a prior on its mass degrades the convergence of the chains because the stellar companion's mass is unconstrained by the data. We adopt stellar-evolution based priors on the masses of the stellar companions for these two systems. For the M dwarf (Currie et al., 2020) companion HD 33632 B, we adopt a $0.22 \pm 0.03 M_{\odot}$ prior consistent with the mass-magnitude relation from Mann et al. (2019), and with stars of similar spectral type (roughly M4; Scholz 2016), e.g., V1352 Ori; $0.23 M_{\odot}$; GJ 3709 B, $\approx 0.27 M_{\odot}$ or HD 239960, $\approx 0.21 M_{\odot}$; Gaidos & Mann 2014. For the F8V-type companion HD 72945 (Anderson & Francis, 2012), we use a mass prior of $1.1 \pm 0.1 M_{\odot}$ prior (consistent with e.g., $1.15 - 1.25\, M_{\odot}$ from Luck 2017; or $1.21^{+0.02}_{-0.03}\, M_{\odot}$ found by Ramírez et al. 2012).

4.6 Orbit & Dynamical Mass Results

In this Section we discuss the inferred orbital elements and masses for each star from our MCMC orbit fits. We improve the secondary mass constraints for all systems, and obtain a large (factor of 7) improvement on the mass precision of Gl 229 B.

We use the reduced chi-squared statistic to assess the goodness of fit: $\chi^2/n_{\rm dof} = \sum({\rm data} - {\rm model})^2/n_{\rm dof}\sigma^2$, where $n_{\rm dof}$ is the number of degrees of freedom. Our use of an RV jitter term enforces a reduced chi-squared near unity for the RV data, but there is no such condition for the relative or absolute astrometry. Table 4.8 gives the chi-squared statistics from the best-fit orbit (for every source) for relative and absolute astrometry.

Figure 4.3 shows the relative orbits of the six systems studied here. Figure 4.4 summa-

Table 4.8. The goodness-of-fit to the relative and absolute astrometry of each orbital fit.

System	$\chi^2_{\mathrm{PA}} + \chi^2_{\mathrm{Sep}}$	$N_{\mathrm{PA}} + N_{\mathrm{Sep}}$	$\chi^2_{Hipparcos\,\mu}$	$\chi^2_{Gaia\,\mu}$	$\chi^2_{\mathrm{HGCAlong-baseline}\,\mu}$
Gl 229	15.82	18	5.77	0.45	0.49
Gl 758	5.06	8	2.64	0.61	0.36
HD 13724	6.22	18	2.65	0.07	0.08
HD 19467	11.73	14	0.26	2.91	0.69
HD 33632	1.41	8	5.87	0.01	0.01
HD 72946	0.01	4	4.59	0.27	1.84

Note. — The χ^2 quoted here are for the maximum likelihood orbits. $\chi^2_{\mathrm{PA}} + \chi^2_{\mathrm{Sep}}$ is the total χ^2 of the fit to the relative astrometry. The reduced chi-squared of the RVs are near one by construction, and so are unlisted. The χ^2 for each proper motion (μ) includes both μ_δ and μ_α, and so it is composed of 2 data points. $N_{\mathrm{PA}} + N_{\mathrm{Sep}}$ is the combined number of PA and separation measurements (twice the number of relative astrometry measurements).

rizes our improvements to the BD masses, displaying the marginalized mass posteriors using the HGCA v.EDR3 in comparison to an otherwise identical analysis using the HGCA v.DR2. We improve the mass precision of four of the BDs by factors of two to five after adopting the *Gaia* EDR3 astrometry. Predicted positions (separation, PA etc.) are available at any epoch via `http://www.whereistheplanet.com/` (Wang et al., 2021). The chains used for the predicted positions on `http://www.whereistheplanet.com/` are included in the supplemental data.

Corner plots for every fit, which show the orbital parameter covariances, are contained within Figure set 4.5. Figure sets 4.6, 4.7, and 4.8 show the fits to the proper motions, relative astrometry, and RVs, respectively.

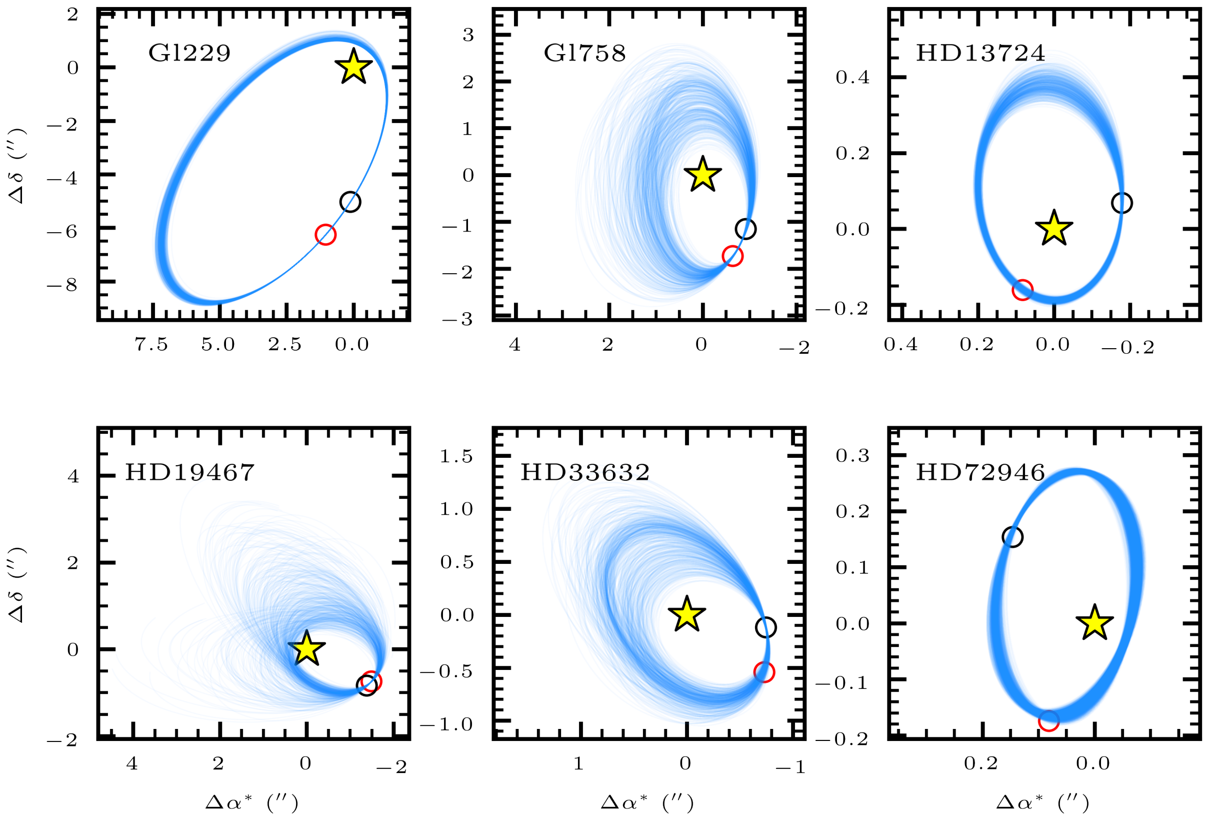

Figure 4.3 Relative orbits (in arc seconds) from fitting RVs, relative astrometry, and absolute astrometry from the v.EDR3 HGCA. Five hundred random orbital draws are shown. Positions at 2010 (red) and 2020 (black) are marked by circles. The host-star of each system is marked with the star symbol at the origin.

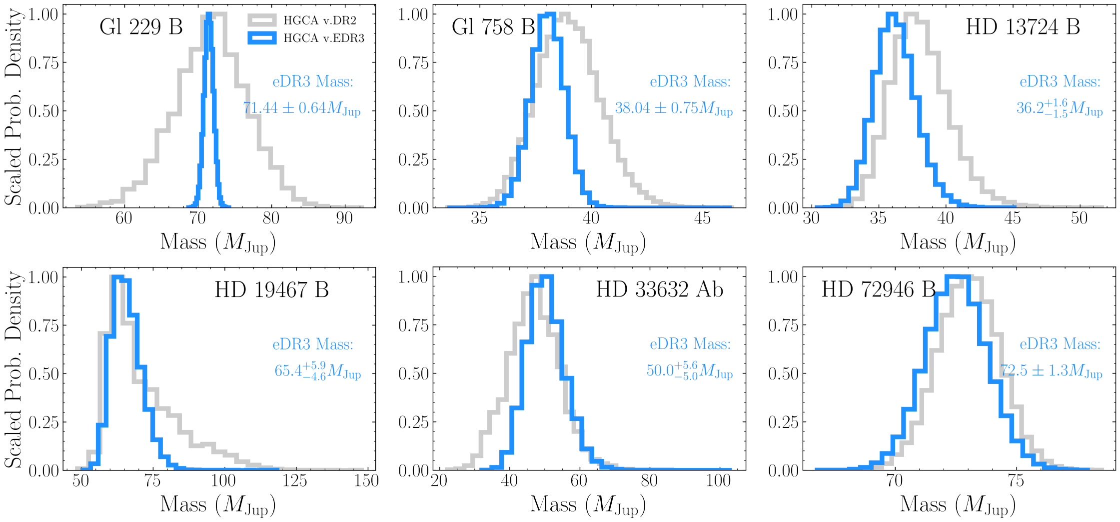

Figure 4.4 Marginalized mass posteriors for the six companions using the HGCA v.EDR3 (blue; Brandt 2021) and HGCA v.DR2 (gray; Brandt, 2018). Each posterior has been scaled to a peak value of 1. The median and 1σ confidence intervals of our new *Gaia* EDR3 mass posteriors are listed in each sub panel. The HGCA v.EDR3 yields masses that are more precise by factors of 2–5 over HGCA v.DR2 for four systems. The other input data (RVs and relative astrometry) are identical between the fits shown here.

4.6.1 Gl 229

Our orbital posteriors are summarized in Table 4.11, which is presented in the appendix. The corner plot and covariances of select orbital parameters are shown in Figure 4.5. We infer a mass of $71.4 \pm 0.6\,M_{\rm Jup}$ for Gl 229 B, and an eccentricity of $0.851^{+0.002}_{-0.008}$, the highest precision of both to-date. Our mass agrees with the previously published value of $70.4 \pm 4.8\,M_{\rm Jup}$ (Brandt et al., 2019) yet is a factor of seven more precise. The χ^2 on the *Hipparcos*-*Gaia* long term proper motion is just 0.5 (Table 4.8); the observed proper motion anomaly of Gl 229 A is in almost exactly the same direction predicted by the best-fit orbit. The fit to the RVs is summarized in Figure 4.8. Because of the long period and significant RV jitter, the *Hipparcos*-*Gaia* absolute astrometry plays a crucial role in constraining the mass of the secondary. Figure 4.7 showcases the fit to the relative astrometry, and Figure 4.6 shows the fit to the *Gaia* and *Hipparcos* proper motions. The goodness-of-fit statistics are good for the relative astrometry ($\chi^2 = 16$ for 18 data points), but the proper motion in declination from *Hipparcos* is discrepant and leads to a poor χ^2 of 5.5 (with 2 data points). Excluding both *Hipparcos* proper motions from the fit changes our best-fit mass and errors by $\leq 0.1\,M_{\rm Jup}$ ($\leq 0.2\sigma$). Likewise, omitting the new relative astrometry and/or the new HIRES RVs changes the mass by $\lesssim 0.2\,M_{\rm Jup}$ and has a negligible effect on the precision of our mass measurement. Nearly all of the improvement in our mass constraint comes from the *Gaia* EDR3 proper motions.

The new NIRC2 relative astrometry improves the mass constraint on the primary star Gl 229 A by a factor of ≈5, removes almost all of the covariance between primary and secondary masses (upper left panel of Figure 4.5), and reduces the semi-major axis uncertainty by a factor of 3. The other parameters are not affected by the new relative astrometry.

For Gl 229, neither the RVs nor relative astrometry cover a significant fraction of the ≈240-year period (Table 4.11), as both data sets have baselines of ≈25 years. These facts, in addition to the nearly face-on orbit, result in a degeneracy between ω and Ω. There are four local maxima for ω and Ω in the posterior distribution, with one mode significantly higher than the others (Figure 4.5). Revealing this multi-modality in an MCMC analysis requires exhaustively exploring the parameter space. An analysis that begins its MCMC chains from previously published orbital parameters could miss such

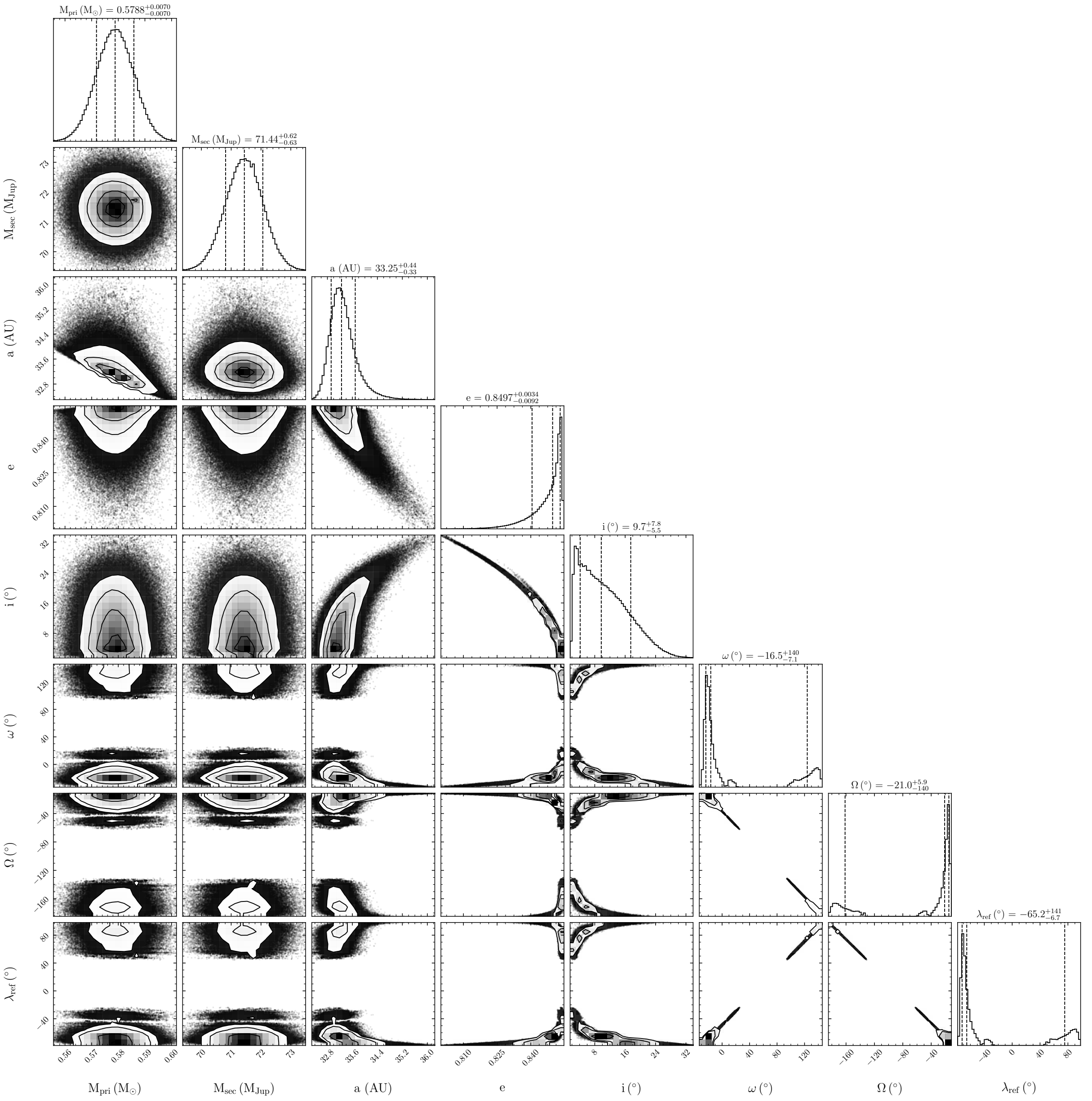

Figure 4.5 Orbital elements, with respect to the star, for Gl 229 B. In the 1D histograms, the vertical-dashed lines about the center dashed lines give the 16% and 84% quantiles around the median. In the 2d histograms, the contours give the 1-σ, 2-σ, and 3-σ levels. The Figure set contains the corner plots for every fit. The complete Figure set (10 images) is permanently hosted with the published version of this article at DOI:10.3847/1538-3881/ac273e.

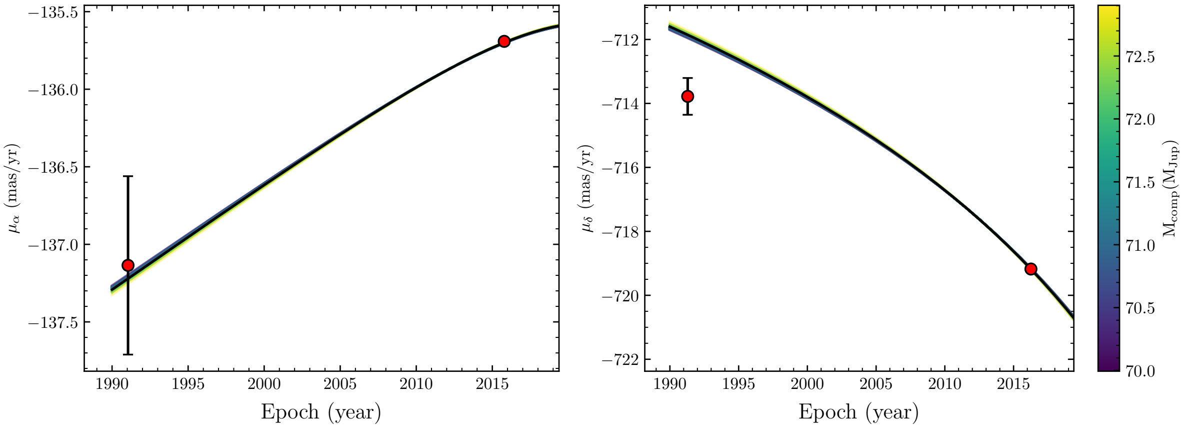

Figure 4.6 Model proper motions compared to the calibrated *Hipparcos* (dot at 1991.25) and *Gaia* EDR3 proper motions (dot near 2016) from the HGCA. The best fit orbit is shown in black. A random sampling of other orbits from the MCMC chain are shown and are color coded by the mass of Gl 229 B. The formal χ^2 of the fit to each proper motion are listed in Table 4.8. The complete Figure set (6 images) is permanently hosted with the published version of this article at DOI:10.3847/1538-3881/ac273e.

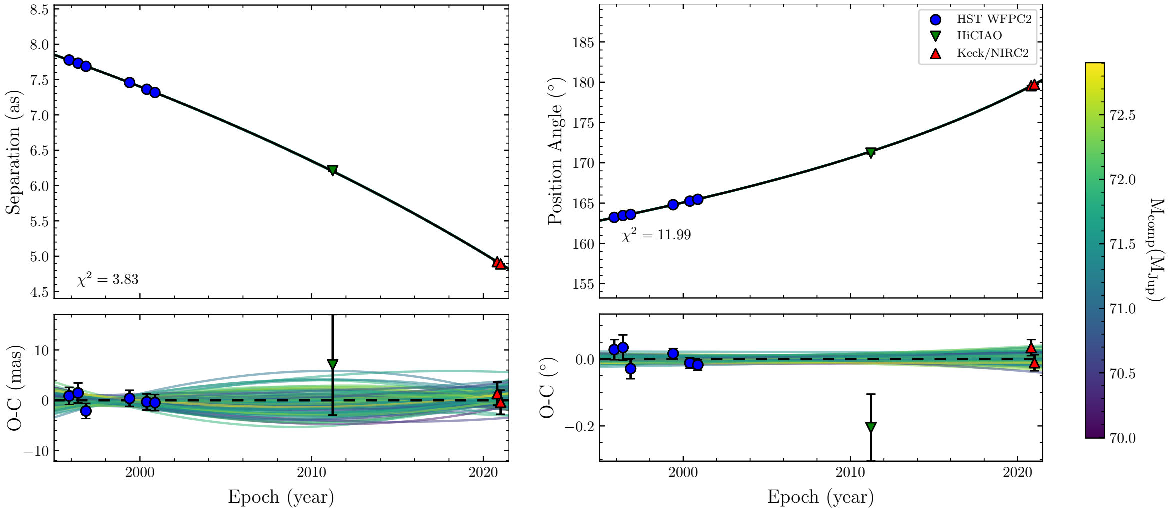

Figure 4.7 Left: relative separation of Gl 229 B. Right: PA of Gl 229 B. A random sampling of orbits from other MCMC steps are shown and are color coded by the mass of Gl 229 B. The best fit orbit is shown in black. The formal χ^2 of the fit to the data are inset and listed in Table 4.8. The complete Figure set (12 images) is permanently hosted with the published version of this article at DOI:10.3847/1538-3881/ac273e.

additional modes in the posteriors.

The true value of Ω and ω could be identified using high-precision relative astrometry.

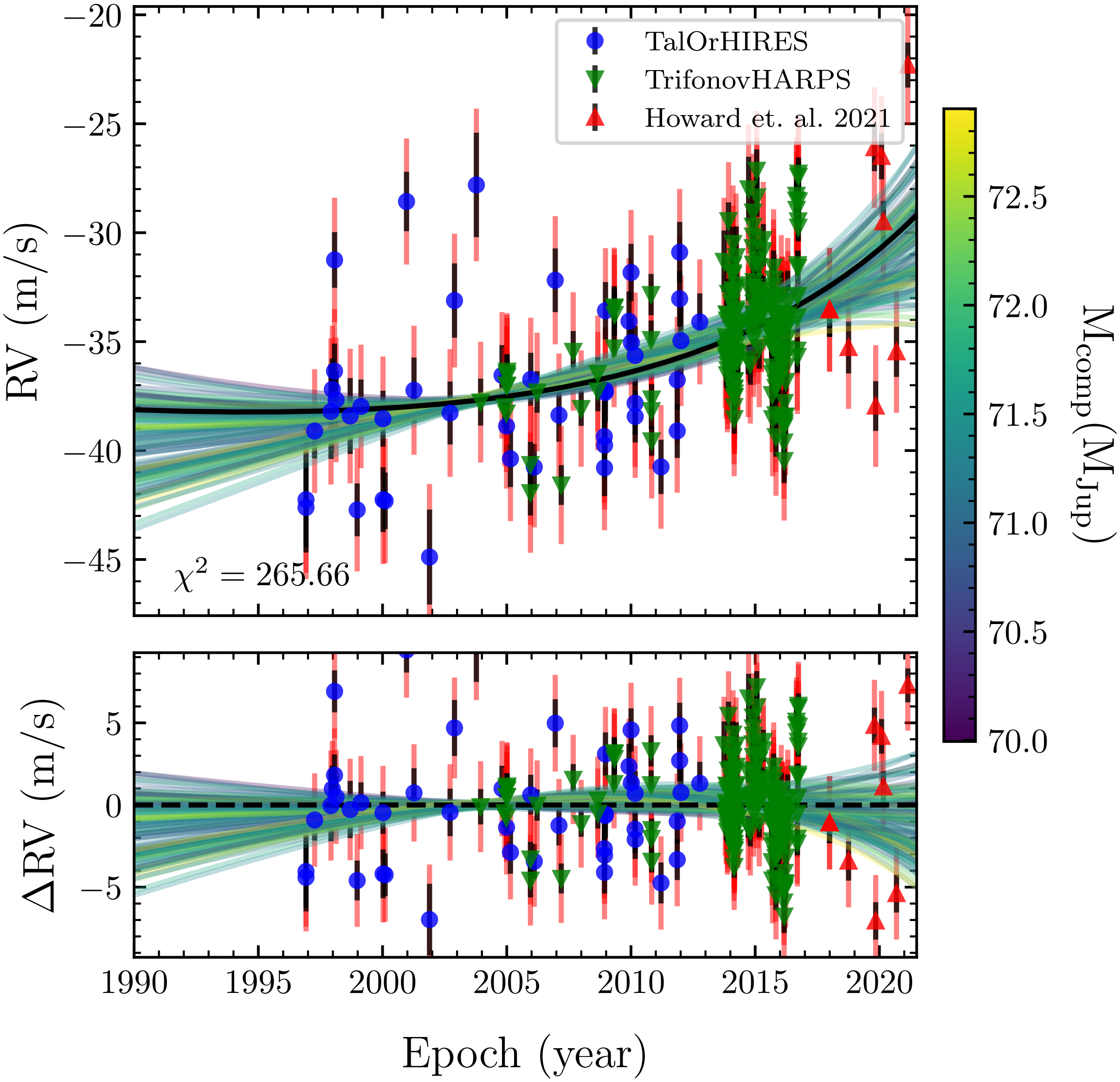

Figure 4.8 Top panel: The observed RVs of Gl 229 A overplot with the best fit orbit (in black) and a random sampling of other orbits from the MCMC chain. Bottom panel: The RV residuals with respect to the best fit orbit. Both panels: The random sampling of other orbits from the MCMC chain are color coded by the mass of Gl 229 B. The black error bars give the observed errors. The red error bars include the best fit jitter added in quadrature to the observed errors. The formal χ^2 of the fit to the data is inset in the top panel. The complete Figure set (6 images) is permanently hosted with the published version of this article at DOI:10.3847/1538-3881/ac273e.

A few VLT/GRAVITY observations, taken now, would serve the same purpose (from the perspective of constraining power) as several $\sigma \approx 3\,\mathrm{mas}$ measurements over the next decade from more 'classic' direct-imaging instruments such as NIRC2 or HiCIAO. A single, 100 μas precise GRAVITY observation (Nowak et al., 2020; Lagrange et al., 2020) would improve the orbital period, inclination, eccentricity, Ω, and ω constraints by 20% to 80%. As we discuss in Section 4.8, Gl 229 B is surprisingly massive, and could be a BD-BD binary. Ultra-precise GRAVITY astrometry might detect the astrometric signature of such an unseen companion.

Tuomi et al. (2014) found evidence for a new super-Earth-sized planet in the Gl 229 system, Gl 229 b. Feng et al. (2020) found that Gl 229 b was still yet to be confirmed but reported the discovery of an additional planet, Gl 229 c. Both are at least super-Earth-sized; their minimum masses are $8.5 \pm 2.0\,M_{\oplus}$ (Gl 229 b) and $7.3 \pm 1.3\,M_{\oplus}$ (Gl 229 c) with RV semi-amplitudes between 1 and 2 m/s (Feng et al., 2020). We perform an identical analysis including these two planets in our orbital fit by subtracting off their RV signals. We use the *maximum a posteriori* (MAP) orbital elements provided in Table 2 of Feng et al. (2020). We infer a mass for Gl 229 B (the BD) that is just 0.24 M_{Jup} (0.4σ) higher than that from the case that ignored the candidate inner planets. The other orbital elements are nearly identical as well. These two inner planets combined contribute $<3\,\mathrm{m/s}$ of RV perturbations, with short orbital periods compared to the total RV baseline. This is only a factor of $\approx$2 larger than the median RV error (including jitter). The reported planets have too little mass and are too close to the star (both have semi-major axes < 1 au; Feng et al. 2020) to impact the inferred mass of Gl 229 B.

We infer a mass of $0.579 \pm 0.007\,M_\odot$ for the primary, Gl 229A. This agrees with the v.DR2 fit by Brandt et al. (2019), who found $0.54^{+0.04}_{-0.03}\,M_\odot$. Our improved primary mass precision is entirely due to the two additional epochs of relative astrometry from NIRC2. In a fit that ignores that new relative astrometry, we find a primary mass with a factor of $\approx$5 worse precision: $0.545^{+0.033}_{-0.030}\,M_\odot$. This is expected because Gl 229 B has a long orbital period ($237.9^{+5.1}_{-4.6}$ years) and so the *Gaia* proper motion is quasi-contemporaneous (i.e., all scans occurred approximately at the same orbital phase of Gl 229 B). In the single-epoch approximation (Brandt et al., 2018), the astrometric acceleration of Gl 229 A on the sky, combined with the parallax and angular separation, constrains only the mass of the companion; it yields no constraint on the mass of the primary. These two facts are partly why we obtain such a better secondary mass constraint after adopting *Gaia* EDR3 astrometry. The additional relative astrometry drives most of the precision increase on the orbital elements (including both period and primary mass, related by Kepler's third law), while the improved *Gaia* EDR3 absolute astrometry drives the improved precision on the secondary mass.

Potential mass systematics below the 1% level

We achieve our highest mass precision for Gl 229 B (0.9%), so we consider here potential systematics below the $\sim$1% level. Unknown systematics within *Gaia* or *Hipparcos* proper motions are unlikely to be a concern; the HGCA dealt with these systematics and corrected them far below this level (Figure 6 of Brandt 2021). The RV star reference set in that work (all non-accelerators according to RV trends) is nicely calibrated into a Gaussian core, with minimal evidence for outliers.

However, a potential source of systematics is the fact that we do not have the *Gaia* EDR3 intermediate astrometric data (the individual positions and uncertainties per transit). This systematic is rooted in the fitting, per MCMC step, of the five-parameter astrometric model to *Gaia* transits. We use the resulting positions and proper motions to compute a likelihood given the measured HGCA proper motions. A *Gaia* transit consists of four components: the transit time, the scan angle of the transit, the along-scan formal error, and whether this particular transit was used in the final solution. `htof` uses scan angles and epochs from the *Gaia* GOST[5] tool. `htof` automatically rejects GOST observations that fall into the documented satellite dead times. `htof` assumes uniform along-scan errors for all observations of one source. Deviations from these assumptions, whether from varying precision or additional rejected observations, will change the relative weighting of different transits in the astrometric fit. As a result, the time of minimal positional uncertainty—the central epochs in the HGCA—may differ between the forward-modeled and catalog values. Using the incorrect central epochs would lead to inferring an incorrect astrometric acceleration.

For Gl 229, we find that `htof`'s computed central epochs from the *Gaia* EDR3 GOST scanning law are only 0.036 yr and 0.017 yr different from the true *Gaia* EDR3 values in right ascension and declination, respectively. The acceleration that we measure is primarily between the midpoint of *Hipparcos* and *Gaia*, around 2004, and *Gaia* in 2016.

---

[5] https://gaia.esac.esa.int/gost/

A discrepancy of 0.036 years is about 0.3% of this baseline, and would lead to a ≈0.3% error in the astrometric acceleration. This is a factor of ≈3 smaller than the ≈1% precision of the HGCA acceleration measurement, though the acceleration of Gl 229 A is increasing as its companion approaches periastron.

The following test shows quantitatively the impact of the GOST approximation on Gl 229 B. By disabling `htof` in `orvara`, `orvara` employs a different approximation (Brandt et al., 2018) that forces the central epoch to be equal exactly to the catalog values. In this case, we find the best fit companion mass grows by $0.5\,M_{\rm Jup}$ — slightly less than 1σ. We do not need access to the full *Gaia* intermediate astrometric data for Gl 229, but it will become essential in the future to push mass precisions well below 1%.

4.6.2 Gl 758

Gl 758 B (Thalmann et al., 2009), a late-T dwarf, has a rich history of dynamical mass measurements. Bowler et al. (2018) measured $42^{+19}_{-7}\,M_{\rm Jup}$ using RVs and relative astrometry. Calissendorff & Janson (2018) and Brandt et al. (2018) improved this estimate with *Hipparcos-Gaia* DR2 accelerations; Brandt et al. (2018) inferred $38.1^{+1.7}_{-1.5}\,M_{\rm Jup}$.

We add newly calibrated RVs from Tal-Or et al. (2019) and update the absolute astrometry using the HGCA v.EDR3. Table 4.12 (presented in the appendix) summarize our posteriors and priors. We infer a mass for Gl 758 B of $38.0 \pm 0.8\,M_{\rm Jup}$, twice as precise as the previous estimate. We infer an eccentricity of 0.24 ± 0.11; a circular orbit remains allowed at $\approx 2\sigma$. The secondary mass posterior is nearly Gaussian. Our priors are all uninformative, but adopting a stellar-evolution informed prior on the mass of Gl 758 A has a negligible effect on Gl 758 B's mass measurement. Using a primary mass prior of $0.96 \pm 0.03\,M_{\odot}$ (consistent with Takeda 2007; Luck 2017) yields a secondary mass that is shifted by only $0.1 M_{\rm Jup}$.

Our inferred eccentricity is more modest than that found by Bowler et al. (2018) ($0.58^{+0.07}_{-0.11}$), but consistent with the most recent estimate of 0.26 ± 0.11 by Brandt et al. (2018). The latter work included absolute astrometric accelerations. Astrometric accelerations favor lower eccentricities than what one would infer from RVs alone. The RV baseline is short compared to the orbital period and so the RV constraint on the eccentricity is relatively weak. The mild eccentricity that we confirm for Gl 758 B cements its place in like company with HD 33632 Ab.

4.6.3 HD 13724

We derive a dynamical mass of $36.2^{+1.6}_{-1.5}\,M_{\rm Jup}$ for HD 13724 B with uninformative priors on all system parameters. The posteriors are summarized in Table 4.13. All posteriors (except for the argument of periastron, ω) are nearly Gaussian.

The mass of this primary star is weakly constrained with RVs, relative astrometry, and *Gaia* DR2 astrometric accelerations, and poorly constrained if one excludes astrometric accelerations. Thus, previous studies assumed an informative prior on the mass of HD 13724 A, allowing better constraints on the secondary mass. Rickman et al. (2020) placed a Gaussian prior of $1.14 \pm 0.06\,M_{\odot}$ on HD 13724 A, a range inferred from the

Ekström et al. (2012) and Georgy et al. (2013) grids of Geneva stellar models. Adopting this prior on the primary has a small ($\approx$1-2$M_{\rm Jup}$, or roughly 1σ) effect on our inferred mass for the secondary.

Combining the new, higher precision *Gaia* EDR3 accelerations with RVs and relative astrometry yields a useful dynamical constraint on the primary mass. We find a dynamical mass constraint of $0.95^{+0.076}_{-0.067}\,M_{\odot}$ for HD 13724 A, using an uninformative prior on the primary mass. Our dynamical mass precision is comparable to the precision of predictions from stellar evolution: e.g., $1.08^{+0.04}_{-0.03}\,M_{\odot}$ reported by Aguilera-Gómez et al. (2018), and the Rickman et al. (2020) prior. We discuss the discrepancy between our dynamical mass and those from stellar evolution in Section 4.7.2.

Our dynamical mass for HD 13724 B, $36.2^{+1.6}_{-1.5}\,M_{\rm Jup}$, is in tension with the first dynamical mass measurement of $50.5^{+3.3}_{-3.5}\,M_{\rm Jup}$ found by Rickman et al. (2020), but agrees well with the minimum mass ($m\sin(i) = 26.77^{+4.4}_{-2.2}M_{\rm Jup}$, determined from RVs alone in the initial discovery by Rickman et al. (2019). Using our inferred inclination of $45.1^{+2.1}_{-1.8}$ degrees and the minimum mass from Rickman et al. (2019), we calculate $m = 37.8^{+7.2}_{-4.3}\,M_{\rm Jup}$. Our inferred eccentricity of 0.335 ± 0.026 is significantly more modest than $e = 0.64 \pm 0.07$ reported by Rickman et al. (2020). Our parameters are consistent with $e = 0.34^{+0.09}_{-0.05}$ found by Rickman et al. (2019).

The two salient differences between our analysis and that of Rickman et al. (2020) are that we include *Hipparcos*-*Gaia* accelerations and that we do not adopt a prior on the mass of HD 13724 A. If we instead adopt the same stellar mass prior of $1.14\pm0.06\,M_{\odot}$, we find a secondary mass of $38.6^{+1.2}_{-1.1}\,M_{\rm Jup}$ and an eccentricity of 0.346 ± 0.026, consistent with our results using an uninformative prior on the primary star's mass. Entirely excluding the *Hipparcos*-*Gaia* accelerations does not resolve the tension either, although it does weaken it. Such an analysis yields an eccentricity of 0.39 ± 0.15 and mass of $41.1^{+9.9}_{-6.2}\,M_{\rm Jup}$.

4.6.4 HD 19467

Using the same primary mass prior as Maire et al. (2020a), we infer a mass of $65.4^{+5.9}_{-4.6}\,M_{\rm Jup}$ for HD 19467 B. All posteriors are summarized in Table 4.14. Our BD mass is consistent with $74^{+12}_{-9}\,M_{\rm Jup}$ as found by Maire et al. (2020a) using *Gaia* DR2, and is roughly twice as precise. We infer an eccentricity of 0.54 ± 0.11, in good agreement with the 0.56 ± 0.09 found by Maire et al. (2020a), as well as the earlier measurement of $0.39^{+0.26}_{-0.18}$ by Bowler et al. (2020). Our inferred period of 320^{+200}_{-80} years is shorter than (but fully consistent with) both 420^{+170}_{-250} years inferred by Bowler et al. (2020) and 398^{+95}_{-93} years from Maire et al. (2020a).

Removing the primary mass prior results in a secondary mass that is fully consistent with the measurement with a prior. However, as mentioned in Section 4.5.2, this yields a posterior for the mass of the G3 dwarf star that is much broader than constraints from stellar evolution.

HD 19467 B's high eccentricity is unlikely to be due to the interactions between it and an undiscovered, inner and massive companion; the RV curve has no signatures of residual few-year signals with semi amplitudes $\gtrsim 10$ m/s. HD 19467 B's high eccentricity places it in like-company (among HR 7672 B and 1RXS2351+3127 B) with the BD population-level peak of eccentricities near $e = 0.6 - 0.9$ studied by Bowler et al. (2020). Without *Gaia* EDR3, Bowler et al. (2020) could only exclude zero eccentricity at $\approx 2\sigma$.

4.6.5 HD 33632 A, Ab, & B

For the L/T transition object HD 33632 Ab we derive a dynamical mass of $50^{+5.6}_{-5}\,M_{\rm Jup}$, which is consistent with and 1.5 times more precise than the mass of $46.4^{+8.1}_{-7.5}$ derived by Currie et al. (2020). This precision improvement is due to *Gaia* EDR3 astrometry. The other orbital parameters are modestly improved and are summarized in Table 4.15. Our preferred fit uses a prior of $1.1 \pm 0.1\,M_{\odot}$ on the host star (Section 4.5.2). An uninformative prior slightly degrades our inferred secondary mass to $50.5^{+6.2}_{-5.1} M_{\rm Jup}$, with a posterior of $1.05^{+0.26}_{-0.21}\,M_{\odot}$ for the primary.

Currie et al. (2020) found bimodalities in the posteriors for both the eccentricity and semi-major axis for HD 33632 B. Our new analysis (using the informative primary mass prior) and more precise *Gaia* EDR3 astrometry breaks both degeneracies: we find $a = 23.6^{+3.2}_{-4.5}$ au and $e = 0.12^{+0.18}_{-0.09}$. Our inclination constraint is modestly improved, $45.2^{+4.7}_{-11}$ degrees, compared to the Currie et al. (2020) result.

HD 33632 A has a widely-separated co-moving M dwarf companion at a common parallax. The analysis of Currie et al. (2020) noted this companion but did not include it in a dynamical fit. We perform a three-body fit to the system, adopting the priors discussed in Section 4.5.2.

Gaia EDR3 provides a $\approx$20 μas constraint (a fractional separation error of 10^{-6}) on the relative position between HD 33632 A and B. This results in sharp likelihood peaks and ridges across parameter space and slows convergence of our MCMC chain. We reduce the precision of the *Gaia* EDR3 relative astrometry (Table 4.7) by a factor of 100 (resulting in separation error of 3 mas), this results in poorer constraints on HD 33632 B but converged posteriors. Additionally, we run the chains for two million steps.

In Figure set 4.5, we show orbital elements for HD 33632 Ab and HD 33632 B (the stellar companion), respectively, from the three-body fit. We summarize all posteriors and priors in Table 4.16. The RV, relative astrometry, and proper motion fits look identical to the two-body fits and so those are not present in the corresponding Figure sets.

The key conclusion is that including the stellar companion does not appreciably shift the MAP values for the BD. The constraints on most of the orbital parameters of the stellar-companion are weak. However, we find a modest constraint on the inclination of HD 33632 B of $74.6^{+4.1}_{-11}$ degrees. We adopt the results of the two-body fit with HD 33632 A and Ab as our preferreed orbital elements for the BD.

Unlike the other BDs considered in this study, HD 33632 Ab appears to have a definitively low eccentricity. The MAP value is near 0.06 and circular orbits are allowed. But, like the other BD's studied herein, HD 33632 Ab is massive, confidently weighing between 40 and 60 $M_{\rm Jup}$.

4.6.6 HD 72946

We infer a secondary mass of $72.5 \pm 1.3\,M_{\rm Jup}$ using the primary mass prior from Section 4.5.2 (the same prior adopted by Maire et al. 2020b). This agrees well with the mass of $72.4 \pm 1.6\,M_{\odot}$ found by Maire et al. (2020b). We derive an eccentricity, period, and inclination of 0.489 ± 0.007, 15.92 ± 0.10 years, and $59.5^{+1.2}_{-1.1}$ degrees, respectively. These agree with the values reported by Maire et al. (2020b), but we improve the precision in period by a factor of ≈1.5 and inclination by a factor of ≈2. Our eccentricity, period, and $M \sin i$ constraints all agree with the initial RV discovery work by Bouchy et al. (2016).

The posteriors from this two-body fit are summarized in Table 4.17. Our analysis adopts an informative prior on the primary star's mass. With an uninformative primary mass prior, we infer a much less precise BD mass of $76.2^{+4.6}_{-4.2}\,M_{\rm Jup}$ together with a primary mass of $1.11^{+0.14}_{-0.12}\,M_{\odot}$.

As noted by Maire et al. (2020b) and Bouchy et al. (2016), HD 72946 has a wide stellar companion, HD 72945, separated by $\approx 10''$ (≈250 au) (Gaia Collaboration et al., 2021). Neither of the latter authors included this companion in their fit. We perform a three-body fit to the system, including this companion. We adopt priors as discussed in Section 4.5.2.

In Figure set 4.5, we show orbital elements (with covariances) for HD 72946 B and HD 72945 from the three-body fit. The orbital elements are tabulated in Table 4.18. The RV, relative astrometry, and proper motion fits look identical to the two-body fits and so those are not present in the corresponding Figure sets. As in the three-body fit to

HD 33632 A/Ab/B, convergence is slowed due to the $\approx$100 μas precision on the separation of the stellar companion. However, the effect here is much smaller than the case with HD 33632 (where we had to inflate the *Gaia* EDR3 errors). We therefore quote results from fits using the relative astrometry from Table 4.7 without inflating those errors.

The exceptional precision of EDR3 provides a good measurement of 200^{+52}_{-41} A.U. for the semi-major axis of the orbit of HD 72945 about HD 72946 (see the marginalized posterior in the corner plot from Figure set 4.5). We obtain a good measurement because we have four constraints for the six phase space components of the stellar companion: separation, PA, and both proper motions.

Comparing the three-body and two-body fits, The inferred BD mass is shifted by less than 0.5σ, and the eccentricity is nearly identical. Like with HD 33632, including the outer stellar companion, HD 79245, does not appear to influence significantly the inferred properties of the BD companion. As with HD 33632, we adopt the two-body parameters for the BD due to the exceptional quality of those chains.

HD 72946 B, like Gl 229 B, is a BD whose mass is near the hydrogen-burning limit. Continued orbital monitoring and better measurements of the RV trend will establish the orbit of the outer, stellar companion and determine whether the HD 72946 AB / HD 72945 system is unstable to Kozai-Lidov oscillations (Kozai, 1962).

4.7 Primary Masses and Stellar Evolution

We highlight here our model-independent primary masses for the three systems where we have useful constraints. These are HD 13724, Gl 229, and HD 72946. We focus our discussion on the dynamical masses of HD 13724 A and Gl 229 A. Our measurement of the high-mass M-dwarf Gl 229 A is especially precise (1.2%; $\pm 0.007 M_{\odot}$). The high-mass end of M dwarfs is particularly interesting due to the well-known ≈10% tension between models and observations in radius-mass space (see Figure 21 of Choi et al., 2016). A relatively small number of precise individual masses constrain the $\approx$0.4–0.6 $M_{\odot}$ regime of the mass-magnitude relation (Figure 21 of Benedict et al., 2016).

We calculate bolometric luminosities for HD 13724 A, Gl 229 A, and HD 72946 A by combining Tycho (Høg et al., 2000) V_T and B_T magnitudes with a parallax-distance from *Gaia* EDR3 and a bolometric correction. This procedure is described in detail in Li et al. (2021). In brief, we convert the Tycho $B_T - V_T$ index into the Johnson $B - V$ index with the transformations provided by ESA (1997) (Eqn. 1.3.26 therein). We adopt the bolometric corrections from Table 5 of Pecaut & Mamajek (2013). We use the *Gaia* EDR3 parallax to obtain a distance posterior, and thereafter bolometric magnitude and luminosity posteriors. Table 4.9 shows the resulting bolometric luminosities and 1-σ confidence intervals together with our dynamical mass measurements.

Table 4.9. Bolometric luminosities for the three sources with dynamical primary mass constraints.

Identifier	Bol. Luminosity $L_\odot$	Mass $M_\odot$
HD 13724 A	1.199 ± 0.014	$0.95^{+0.08}_{-0.07}$
Gl 229 A	0.0430 ± 0.0005	0.579 ± 0.007
HD 72946 A	0.871 ± 0.009	$1.11^{+0.14}_{-0.12}$[a]

[a] This posterior comes from the orbital fit to HD 72946 A/B that did not include a prior on the mass of A.

4.7.1 Gl 229 A & the Mass-Luminosity-Relation

With our new 1.2% precise mass for Gl 229 A, the star becomes one of the few early-M dwarfs with an individually measured dynamical mass. We now compare it to calibrations of mass-magnitude relations from Benedict et al. (2016) and Mann et al. (2019), sometimes also referred to as the Mass-Luminosity Relation (MLR). Benedict et al. (2016) used mass measurements of individual stars within binaries; Mann et al. (2019) used the larger sample of binaries with measured total system masses.

We use $V = 8.129$ mag from Paunzen (2015), adopting an error of 0.010 mag, and the *Gaia* EDR3 parallax of 173.574 ± 0.017 mas to calculate an absolute magnitude of $M_V = 9.326 \pm 0.010$ mag. That agrees with $M_V = 9.33 \pm 0.01$ mag reported by Holmberg et al. (2009). Gl 229 A is so bright in the infrared that 2MASS gives a poor photometric measurement, so we use $K_s = 4.15 \pm 0.05$ mag from Leggett (1992), assuming that the conversion between 2MASS and CIT photometric systems is negligible within the errors.

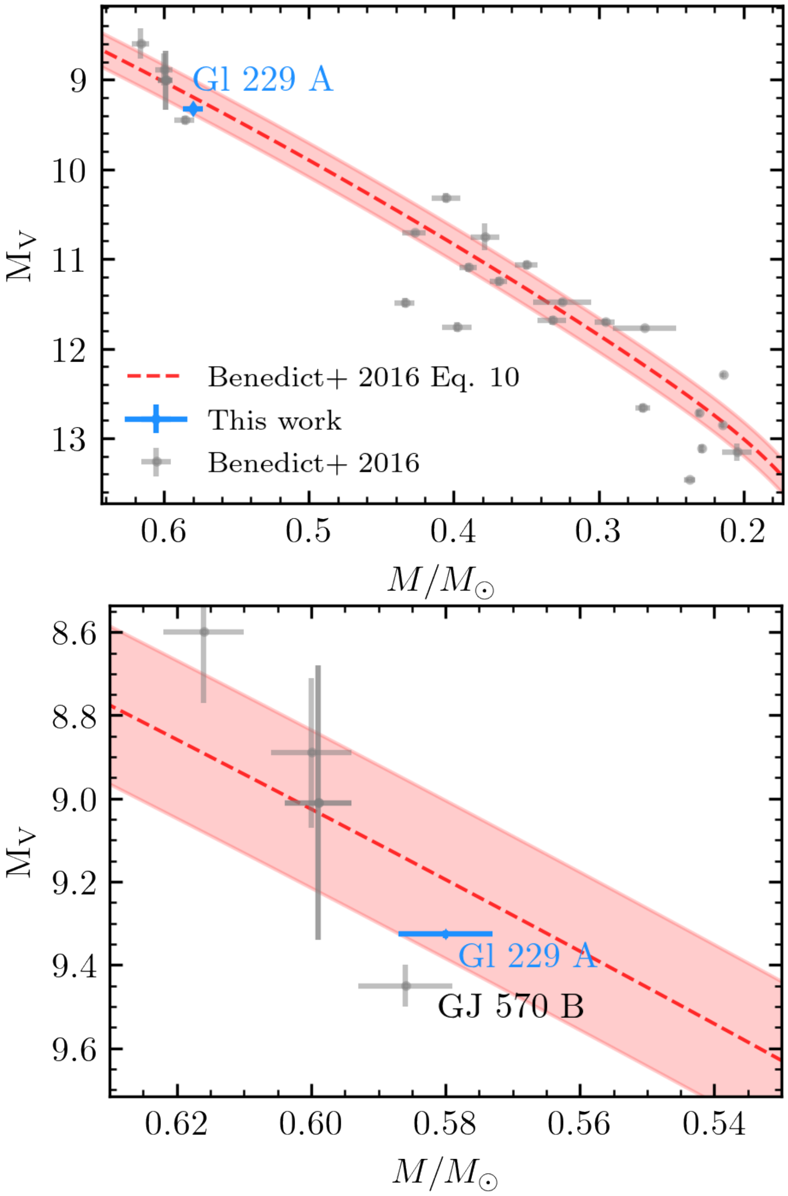

Figure 4.9 Absolute magnitude as a function of mass for Gl 229 A (shown in blue using the mass from this work) and binaries from Benedict et al. (2016) (shown in gray). The best-fit mass-magnitude relation from Benedict et al. (2016) is the dashed red line, and the red region gives the root-mean-square scatter about the best fit relation (0.19 in $M_{\rm V}$; Benedict et al., 2016). The error bar on $M_{\rm V}$ for Gl 229 A is smaller than the diamond plotting symbol in both panels. Gl 229 A borders a gap at $\approx$0.45–0.55 $M_\odot$ in the Benedict et al. (2016) calibration sample. The lower panel shows a zoom-in highlighting Gl 229 A and the five other stars on the high-mass end of the Benedict et al. (2016) sample (GJ 570 B, GJ 278 C/D, GU Boo A/B). Note that GJ 278 C/D overlap.

Figure 4.9 displays Gl 229 A and the binaries from Benedict et al. (2016) in mass-M_V space. Their best-fit double-exponent empirical mass-magnitude relation is overplotted as a dashed red line. Our mass for Gl 229 A is as precise as the other five stars above $0.5\,M_\odot$ from the Benedict et al. (2016) sample. However, binary stellar evolution has played an unknown but potentially important role in some of the other stars that constrain the mass-magnitude relation at this high-mass end. GJ 278 C/D, for example, has a sub-1 day orbital period (Feiden & Chaboyer, 2013), implying that stellar tides may have influenced their evolution. Gl 229 A, with its distant BD companion, is free from such concerns. As the bottom panel of Figure 4.9 shows, our dynamical mass for Gl 229 A agrees within the uncertainties of the mass-M_V relation from Benedict et al. (2016).

We use the code[6] provided by Mann et al. (2019) to estimate a mass of $0.549 \pm 0.015\,M_\odot$ for Gl 229 A from its K-band photometry and parallax. This is lower than our measured mass but consistent within 1.8σ. Thus, Gl 229 A provides a remarkable corroboration of both the empirical mass-magnitude relation of Benedict et al. (2016) and Mann et al. (2019) at their quoted uncertainties.

4.7.2 The dynamical and stellar-evolution masses for HD 13724 A

Our dynamical mass for HD 13724 A is $0.95^{+0.08}_{-0.07}\,M_\odot$, about half as precise as predictions from stellar evolution. However, stellar evolution predictions favor systematically higher

---

[6] `https://github.com/awmann/M_-M_K-`

masses: $1.101 \pm 0.022\,M_\odot$ (Delgado Mena et al., 2019), $1.120 \pm 0.010\,M_\odot$ (Gomes da Silva et al., 2021), and $1.08^{+0.04}_{-0.03}\,M_\odot$ (Aguilera-Gómez et al., 2018). To confirm whether our dynamical mass is discrepant with stellar evolution models, we build solar-like, non-rotating models using Modules for Experiments in Stellar Astrophysics (MESA) (version 15140; Paxton et al., 2011, 2013, 2015, 2018, 2019).

We start by calibrating a solar model using the simplex test suite in MESA, which adopts Grevesse & Sauval (1998) abundances and includes the effects of diffusion (Thoul et al., 1994) and exponential overshoot mixing (Herwig, 2000). The key parameters are the initial mass fractions of helium (Y) and metals (Z), the mixing length parameter (α), and the overshoot parameter ($f_{\rm ov}$). The solar-calibrated values of α and $f_{\rm ov}$ are then used to generate a set of models with masses ranging from $0.95\,M_\odot$ to $1.15\,M_\odot$. We use the tracks in this mass range to extract a range of masses that agree with the observed luminosity and $T_{\rm eff}$.

We adopt a linear enrichment law, $Y = Y_{\rm p} + ({\rm d}Y/{\rm d}Z)Z$ where $Y_{\rm p} = 0.249$ and ${\rm d}Y/{\rm d}Z = 1.4$. We use the effective temperature ($T_{\rm eff}$) of $5824 \pm 19\,{\rm K}$ from Tsantaki et al. (2013). We inflate the errors to $\pm$50K to reflect the spread of the other $T_{\rm eff}$ measurements in the literature from high resolution spectra (Delgado Mena et al., 2014; Datson et al., 2015; Soubiran et al., 2016).

Figure 4.10 plots our MESA models in a Hertzsprung-Russell diagram, along with isochrones at 1 Gyr steps. Models with an initial $[Z/X] = 0.26$–0.27, ages of 1–4 Gyr, and masses $1.10 \pm 0.02\,M_\odot$ simultaneously match the observed luminosity ($1.199 \pm 0.014\,L_\odot$),

$T_{\rm eff}$, and surface $[Z/X]$ (0.241; Nissen et al., 2020). Our 0.95-$M_\odot$ model is a factor of 1.5 too low in luminosity at the measured effective temperature, firmly ruling it out. However, HD 13724's dynamical mass has a sufficiently large uncertainty that the tension with our MESA-derived mass of $1.10 \pm 0.02\,M_\odot$ is $\approx 2\sigma$. Future data will improve the mass precision and determine whether this system is meaningfully discrepant with the predictions of stellar evolutionary models.

Evolutionary models can also constrain the star's age given a luminosity and either an effective temperature or a dynamical mass. Figure 4.10 shows that MESA currently provides only a weak constraint of 1–4 Gyr. Even this constraint depends on modeling details in the convective zone that set the effective temperature. A precise dynamical mass would remove this dependence and enable a better age estimate from stellar models.

4.8 Benchmark Tests of Substellar Evolutionary Models

Substellar cooling models predict an object's luminosity given its age, mass, and composition. Benchmark BDs with known physical parameters provide the strongest tests of these models. Here we combine our masses and ages with measured BD luminosities to assess substellar evolutionary models.

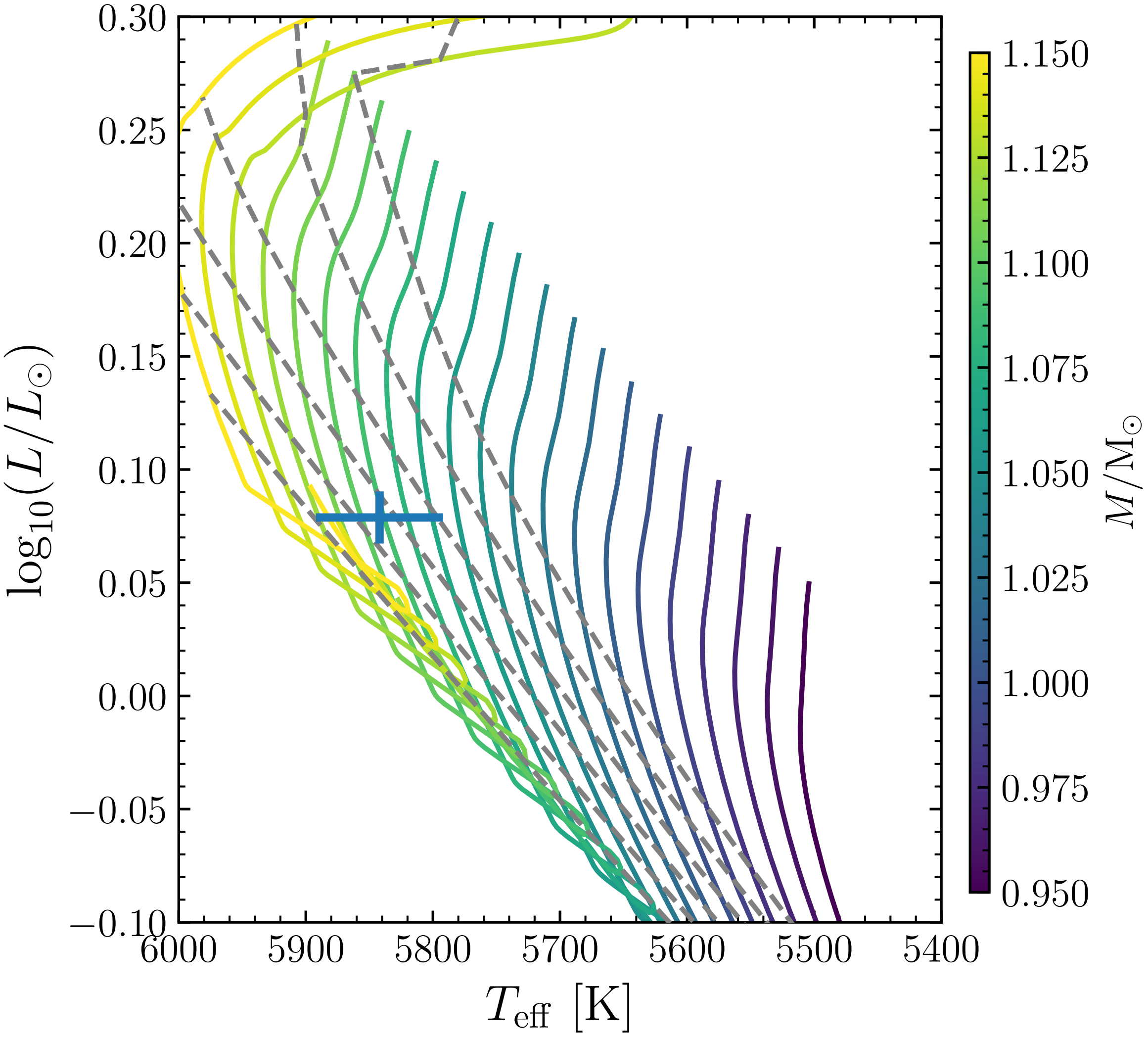

Figure 4.10 Newly computed MESA evolution tracks at a variety of masses, tuned to HD 13724 A's surface composition and calibrated with the simplex test suite. The blue point and error bars give the luminosity from Table 4.9 and the (error inflated) $T_{\rm eff}$ from Tsantaki et al. (2013). Each model is run until core hydrogen is exhausted. The dashed gray lines are isochrones at 1,2,3,4,5 and 6 Gyr; 1 Gyr is the bottom leftmost and 6 Gyr is the top rightmost. Models with masses of $1.10 \pm 0.02 M_\odot$ and ages 1–4 Gyr are consistent with the observed $T_{\rm eff}$ and luminosity.

4.8.1 Overview of Evolutionary Models

We consider three sets of evolutionary models. Each makes different assumptions about the most influential unknown: the atmospheric boundary condition. The earliest-developed models in our set are from Burrows et al. (1997). These make a number of different assumptions than the more recent models we consider, the most important being their lower atmospheric opacities. This is partly due to knowledge of opacity sources improving and expanding as more complete molecular line lists have been developed with time. The Burrows et al. (1997) models also use lower-opacity "gray" atmospheres at higher temperatures, as their main focus was to explore cooler BDs and giant planets ($T_{\rm eff} < 1300\,{\rm K}$).

The second set of substellar models is from Saumon & Marley (2008). The "hybrid" calculations of these evolutionary models assume cloudy atmospheres at warmer effective temperatures ($T_{\rm eff} > 1400\,{\rm K}$), no clouds at $T_{\rm eff} < 1200\,{\rm K}$, and a combination of cloudy and cloud-free atmospheres at intermediate temperatures.

The third set are the ATMO 2020 evolutionary models from Phillips et al. (2020). These are the latest cloud-free evolutionary models in the same lineage as "Cond" (Baraffe et al., 2003) and BHAC15 (Baraffe et al., 2015). Unlike the hybrid models from Saumon & Marley (2008) that are applicable over the whole $T_{\rm eff}$ range of the companions we examine here, ATMO 2020 is only intended to apply to cloud-free, later-type T dwarfs like Gl 229 B, Gl 758 B, and (marginally) HD 19467 B. For completeness, we still compare all companions to all models, even though the ATMO 2020 and Burrows et al. (1997)

models are only intended for cooler BDs.

4.8.2 Description of Benchmark Tests

We perform two types of benchmark tests. For one, we use our determinations of age and $L_{\rm bol}$ to derive a model mass that we compare to our dynamical masses. For the other, we derive BD ages from models given $L_{\rm bol}$ and mass measurements and compare to our inferred host star ages. Not all models are computed beyond an age of 10 Gyr, so in the following analysis we restrict all age distributions to $\leq$10 Gyr. Figure 4.11 shows our sample in comparison to evolutionary model isochrones, as well as other previous mass measurements for ultracool dwarfs.

Our method for computing posterior distributions is based on the Monte Carlo rejection sampling approaches described in Dupuy & Liu (2017) and Dupuy et al. (2018). We begin by drawing random masses and ages. When inferring mass, we use our age posterior and a distribution uniform in $\log M$; when inferring age we use a distribution uniform in time and our dynamical mass posterior. We then bi-linearly interpolate the evolutionary model grid at each mass and age to compute a test $L'_{\rm bol}$ along with effective temperature, radius, and surface gravity. For each luminosity test value, we compute $\chi^2 = (L_{\rm bol} - L'_{\rm bol})^2/\sigma^2_{L_{\rm bol}}$ and then determine the global minimum of all trials ($\chi^2_{\rm min}$). We accept each pair of mass and age into our output posterior with probability $\exp(-(\chi^2 - \chi^2_{\rm min})/2)$. This produces not only an output distribution of the parameter of interest (mass or age) but also any other properties interpolated from the evolutionary

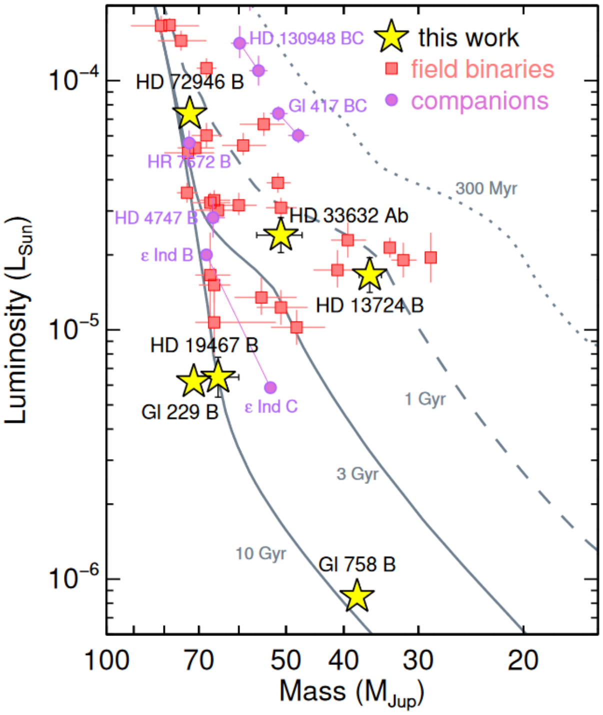

Figure 4.11 Luminosity as a function of mass for our sample (yellow stars) compared to the field dynamical mass sample (red squares), other companions (purple circles), and hybrid Saumon & Marley (2008) evolutionary models (gray lines). The ensemble of 41 objects plotted here generally overlap with evolutionary models at the ages expected for field dwarfs; only Gl 229 B is more massive (less luminous) than the oldest model isochrone at its luminosity (mass). Most of the 35 literature masses here are from Dupuy & Liu (2017), with additional masses from Cardoso (2012), Lazorenko & Sahlmann (2018), Brandt et al. (2018), Dupuy et al. (2019b), and Sahlmann et al. (2020, 2021). There are dynamical mass measurements for ϵ Ind BC from both Cardoso (2012) and Dieterich et al. (2018). We opt to plot the former because they are in better agreement with a new analysis of the orbit by Chen et al. (2022).

models.

To perform quantitative benchmark tests, we then compare the model-derived posterior distributions to the observed ones. We use a one-tailed test of the null hypothesis (that the two distributions are consistent), following Bowler et al. (2018). Given independent draws from the two distributions, we compute the probability that a draw from the model-derived posterior is larger or smaller than the draw from the observed posterior. We convert this probability into a Gaussian sigma.

4.8.3 Summary of Luminosity Measurements

Not all of our benchmark sample have published $L_{\rm bol}$ values, so we derive those that are needed using empirical relations. For Gl 229 B and Gl 758 B we use the same $L_{\rm bol}$ values from Filippazzo et al. (2015) and Bowler et al. (2018), respectively.

For HD 19467 B and HD 33632 Ab, we use the K-band absolute magnitude–$L_{\rm bol}$ relation of Dupuy & Liu (2017) and the photometry reported by Crepp et al. (2014) and Currie et al. (2020) to compute luminosities of $\log(L_{\rm bol}/L_{\odot}) = -5.19 \pm 0.08$ dex and -4.62 ± 0.07 dex, respectively. Our derived luminosity for HD 19467 B is consistent with those by Maire et al. (2020a) who found $\log(L_{\rm bol}/L_{\odot}) = -5.17^{+0.10}_{-0.08}$ dex and -5.31 ± 0.12 dex from J and K bands, respectively. Our derived luminosity for HD 33632 Ab agrees with Currie et al. (2020), who found $\log(L_{\rm bol}/L_{\odot}) = -4.62^{+0.04}_{-0.08}$ dex.

For HD 13724 B, we first convert the Rickman et al. (2020) SPHERE medium-band photometry to standard systems. We compute synthetic photometry from their best-matching template spectrum (2MASS J10595185+3042059; Sheppard & Cushing, 2009). For the MKO system, we find $J_{\mathrm{MKO}} = 17.41 \pm 0.05\,\mathrm{mag}$, $H_{\mathrm{MKO}} = 17.61 \pm 0.14\,\mathrm{mag}$, $K_{\mathrm{MKO}} = 17.16 \pm 0.17\,\mathrm{mag}$. For the 2MASS system we find $J_{\mathrm{2MASS}} = 17.62 \pm 0.05\,\mathrm{mag}$, $H_{\mathrm{2MASS}} = 17.55 \pm 0.14\,\mathrm{mag}$, $K_{\mathrm{S,2MASS}} = 17.03 \pm 0.17\,\mathrm{mag}$. Using the K_{MKO}-band photometry and Dupuy & Liu (2017) L_{bol} relation we find $\log(L_{\mathrm{bol}}/L_{\odot}) = -4.78 \pm 0.07\,\mathrm{dex}$.

For HD 4113 C, we adopt a bolometric correction of $\mathrm{BC}_J = 2.0 \pm 0.5\,\mathrm{mag}$ based on Figure 12 of Filippazzo et al. (2015) that, combined with the photometry reported in Cheetham et al. (2018), gives $\log(L_{\mathrm{bol}}/L_{\odot}) = -6.30 \pm 0.22\,\mathrm{dex}$.

While HD 72946 B has a luminosity of $-4.11 \pm 0.10\,\mathrm{dex}$ from Maire et al. (2020b), for consistency and improved precision, we compute one here using the Dupuy & Liu (2017) relation. As with HD 13724 B, we first convert the SPHERE photometry from Maire et al. (2020b) to the MKO and 2MASS systems using their best-matching template spectrum (2MASS J03552337+1133437; Bardalez Gagliuffi et al., 2014). We find $H_{\mathrm{MKO}} = 14.55 \pm 0.05\,\mathrm{mag}$ and $H_{\mathrm{2MASS}} = 14.50 \pm 0.05\,\mathrm{mag}$, resulting in $\log(L_{\mathrm{bol}}/L_{\odot}) = -4.133 \pm 0.023\,\mathrm{dex}$.

4.8.4 Discussion of Individual Objects

Our benchmark tests fall into two categories: cases where our measured mass is significantly more precise than the model-derived mass (Figure 4.12) or where the precision of the two are comparable (Figure 4.13). For a substellar object of fixed L_{bol}, mass M and age t scale approximately as $M \propto t^{0.49}$ (Burrows et al., 2001). Thus, as Liu et al. (2008) noted while discussing the first T-dwarf mass benchmark, a 5% mass uncertainty

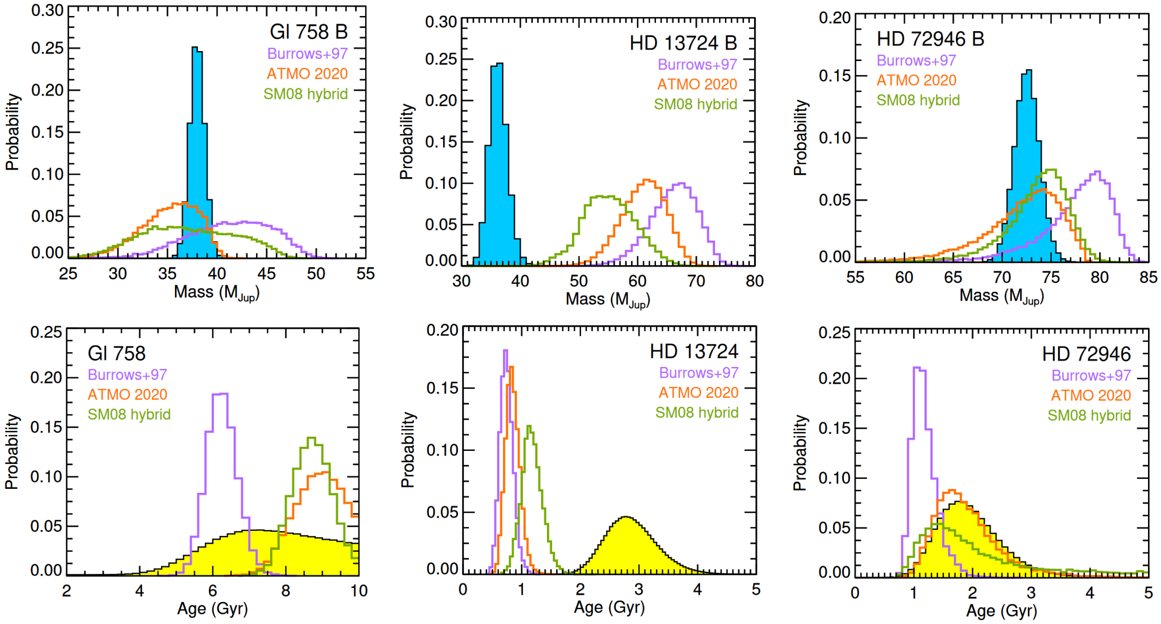

Figure 4.12 Top row: dynamical mass posterior distributions (filled blue histograms) compared to the mass posteriors derived from substellar evolutionary models given $L_{\rm bol}$ and host star ages (unfilled histograms). Bottom row: host star activity-based age probability distributions (filled yellow histograms) compared to the age posteriors derived from evolutionary models given $L_{\rm bol}$ and mass (i.e., substellar cooling ages). Objects shown here have masses measured so precisely that the limiting factor is the uncertainty in the host star's age. Gl 229 B would be shown here, but as discussed in Section 4.8, its luminosity is too low to be consistent with the ATMO 2020 or SM08 models even at 10 Gyr. HD 13724 B is the only object here that appears significantly discrepant with any models. As a mid-T dwarf, clouds are likely to be important in modeling its evolution, thus only the hybrid models (3.8σ disagreement) are relevant for it.

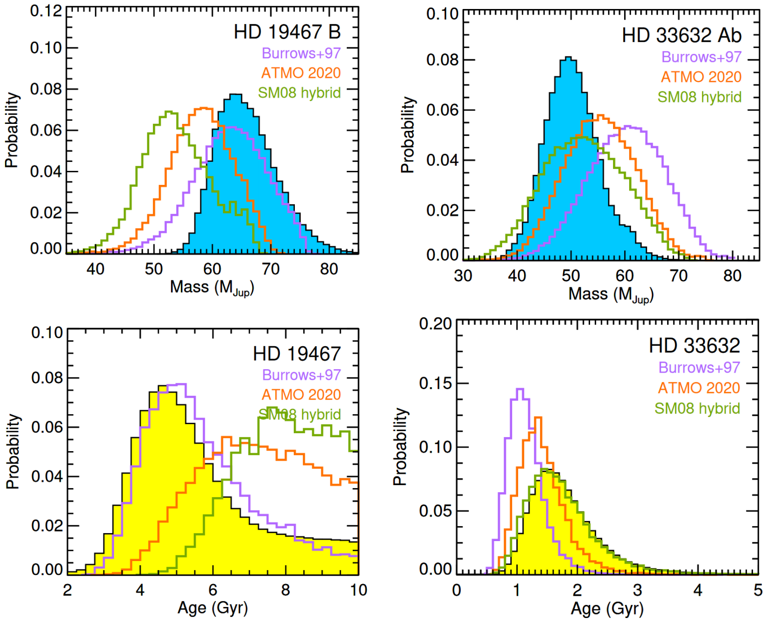

Figure 4.13 Figure 4.12 gives a description of the posteriors. The objects displayed here have comparable uncertainties in mass and age. The only significant discrepancy with models is the mid-T dwarf HD 19467 B where the measured mass is higher than expected from the Saumon & Marley (2008) hybrid models.

propagates to a 10% uncertainty in the model-derived age. The limiting factor in the benchmark test will be the independently determined age, unless the fractional error in the age is no more than twice the fractional error in mass. Our age determinations range in precision from 15%–30%, while our measured masses range in precision from 1%–10%. Our masses for HD 19467 B and HD 33632 Ab are within a factor of two of the precision of their stellar ages, but in most cases our benchmarks are dominated by the precision of the age determination and not the mass precision (Gl 229 B, Gl 758 B, HD 13724 B, and HD 72946 B).

Gl 229 B Brandt et al. (2019) found an unexpectedly high mass of $70 \pm 5\,M_{\rm Jup}$ for Gl 229 B. Our new mass of $71.4 \pm 0.6\,M_{\rm Jup}$ greatly increases the significance of the tension with evolutionary models (Figure 4.14). Gl 229 B's luminosity is 11σ lower than predicted by the hybrid Saumon & Marley (2008) models for an object of this mass even at 10 Gyr.

The Burrows et al. (1997) models overlap with our mass measurement well and give a cooling age of $7.5^{+0.3}_{-0.4}$ Gyr. As discussed in detail by Saumon & Marley (2008), the primary reason that the Burrows et al. (1997) models predict lower luminosities at a given mass and age is their lower global opacity ($L_{\rm bol} \propto \kappa_R^{0.35}$, where κ_R is the Rosseland mean atmospheric opacity; Burrows & Liebert, 1993). The differences in predicted luminosity are 0.3–0.6 dex (Figure 6 of Saumon & Marley, 2008), especially around the H-fusion mass boundary at old ages.

There are two ways Gl 229 B could have such a low global opacity. First, it and its host star could inherit a low metallicity. However, Brandt et al. (2019) concluded that a sub-solar metallicity for Gl 229 A is implausible given an assortment of measurements that are consistent with solar metallicity (Neves et al., 2013; Gaidos & Mann, 2014; Gaidos et al., 2014). Secondly, Gl 229 B could have acquired a sub-solar metallicity during its formation. However, companions formed by disk fragmentation are expected to be at least as metal rich as the host star (e.g., Boley & Durisen, 2010). The only processes that alter the companion metallicity, such as concentration of solids at the site of fragmentation (e.g., Haghighipour & Boss, 2003; Rice et al., 2006) or planetesimal capture (e.g., Helled & Bodenheimer, 2010), only increase metallicity.

There are two possibilities by which Gl 229 B's mass could be reconciled with models without needing an unusually low opacity: either it is a tight binary, or there is another massive companion in orbit around Gl 229 A, which would muddle our interpretation of the astrometric acceleration. The latter scenario was ruled out in Brandt et al. (2019) and is even more unlikely with the more precise *Gaia* EDR3 proper motions and the additional RVs. The HGCA acceleration is significant at $\approx 115\sigma$ (a difference of 13,000 in χ^2), and it points in exactly the direction expected. An additional massive companion would be very unlikely to preserve the low χ^2 value for the proper motion anomalies of just 0.49 ($n_{\rm dof} = 2$).

Gl 229 B itself being an unresolved binary remains a plausible explanation that would not require radical changes to substellar evolutionary models. Brandt et al. (2019) noted that it is not unusually luminous for its spectral type, making a nearly equal-flux companion unlikely. A faint companion would need a sufficiently small orbit to elude detection by astrometric perturbations in the relative astrometry, and we discuss this possibility in more detail below in Section 4.8.5.

Gl 758 B This late-T-type companion has one of the most precise masses in our sample, with a fractional error of 2.0%. Our benchmark test is dominated by the uncertainty in the host star's age, which we discussed in detail in Section 4.2. All three substellar models' cooling ages are consistent with our broad host star age distribution (4.7–10 Gyr at 2σ). The Burrows et al. (1997) models give a much younger age (6.2 ± 0.4 Gyr) than hybrid (8.7 ± 0.6 Gyr) or ATMO ($8.9^{+0.8}_{-0.6}$ Gyr) models. This companion remains the sole test of models at the cold temperatures (ATMO 2020-derived $T_{\rm eff} = 603 \pm 9$ K) of older, lower-mass BDs.

HD 13724 B As discussed in Section 4.2, we find a host star age from gyrochronology that is significantly older than other determinations in the literature (e.g., 1.0 ± 0.9 Gyr; Rickman et al., 2020). Here we conservatively adopt the youngest of our age distributions ($2.8^{+0.5}_{-0.4}$ Gyr) that uses the stellar rotation period of 21 days from Arriagada (2011). A very similar age posterior would result from using the 20.2 ± 1.2 day rotation period from Rickman et al. (2019). As a mid-T dwarf, the hybrid evolutionary models are the most appropriate for HD 13724 B, and indeed they provide the best agreement in our benchmark test. Still, they give a substellar cooling age that is highly discrepant (3.8σ) with our host star age distribution (Figure 4.12).

Given the disagreement in the literature about the age of this host star, it is possible that this is a case where the host star itself is atypical (e.g., rotating slowly at a young age). However, it is also possible that substellar evolutionary models are to blame for the young derived age, as this phenomenon has been observed before for moderately young ($\lesssim$1 Gyr) BDs (Dupuy et al., 2009, 2014). We discuss both of these possibilities in more detail in Section 4.8.5.

HD 19467 B Maire et al. (2020a) concluded that both hybrid and cloud-free models predict a luminosity $\approx 1\sigma$ higher than is observed given their measured mass. Our benchmark test yields qualitatively similar results: the Burrows et al. (1997) models best match the observed luminosity given the measured mass and host-star age, even though those models should not be appropriate for a mid-T dwarf. Quantitatively, there is only a 1.4σ difference between our host star age distribution and the cooling age derived from the (most appropriate) Saumon & Marley (2008) hybrid models. Therefore, with a mass of $65^{+6}_{-5}\,M_{\rm Jup}$ and a spectral type of T5.5 ± 1.0 (Crepp et al., 2015b), HD 19467 B joins the ranks of unexpectedly (but not anomalously) massive T dwarfs near the substellar mass boundary, comparable to WISE J0720−0846 B ($66 \pm 4\,M_{\rm Jup}$ and T5.5 ± 0.5; Dupuy et al., 2019b).

HD 33632 Ab This is the other L/T transition object in our sample, and our mass measurement of $50^{+6}_{-5}\,M_{\rm Jup}$ is about twice as precise as that in Currie et al. (2020). Our host star age of $1.7^{+0.4}_{-0.6}$ Gyr agrees very well with all model-derived cooling ages, especially the Saumon & Marley (2008) hybrid models that are appropriate for an object of this type.

HD 72946 B Our mass agrees well with that found by Maire et al. (2020b) who discussed in detail its host star's somewhat young age (0.8–3 Gyr) and somewhat metal-rich composition ([Fe/H] = 0.11 ± 0.03 dex; Bouchy et al. 2016). All three substellar model-derived ages agree with our adopted host star age distribution ($1.9^{+0.5}_{-0.6}$ Gyr), with the largest difference being 1.4σ for the Burrows et al. (1997) models. Such cloud-free, low-opacity models should not be appropriate for this companion's spectral type (L5.0 ± 1.5), and they should be especially ill-suited given the metallicity implied by the host star. HD 72946 B is therefore a case of a warm companion (hybrid-derived $T_{\rm eff}$ =

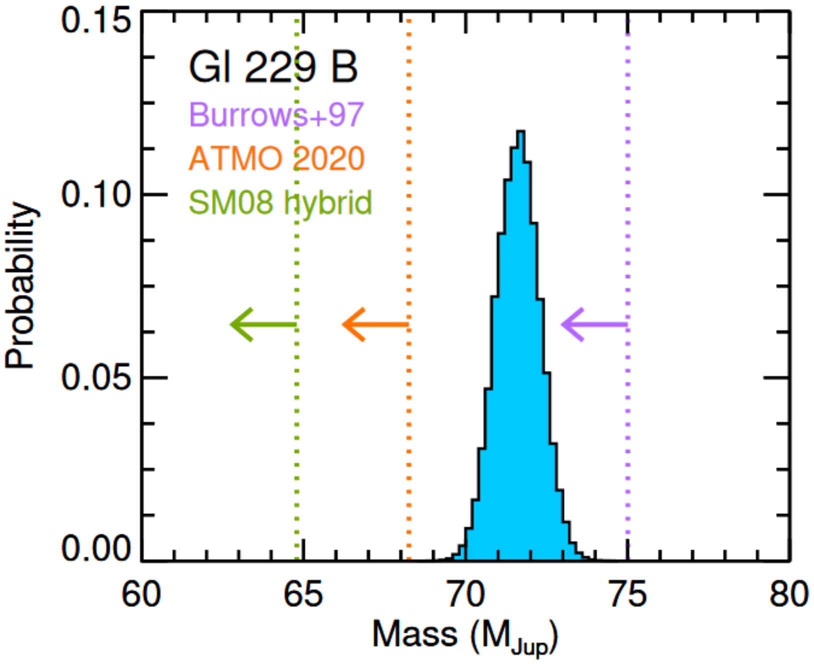

Figure 4.14 Same plot as in Figures 4.12 and 4.13 except showing the 3σ upper limits in mass predicted by models corresponding to an age of <10 Gyr for Gl 229 B. Only the evolutionary models of Burrows et al. (1997) are consistent with our measurement. No other models, to our knowledge, predict that such a massive object can achieve such a low luminosity within a Hubble time.

1700 ± 90 K) with a mass ($72.5 \pm 1.3\, M_{\rm Jup}$) that could be on either side of the hydrogen-fusion boundary, and a host-star age that is consistent with its substellar cooling age.

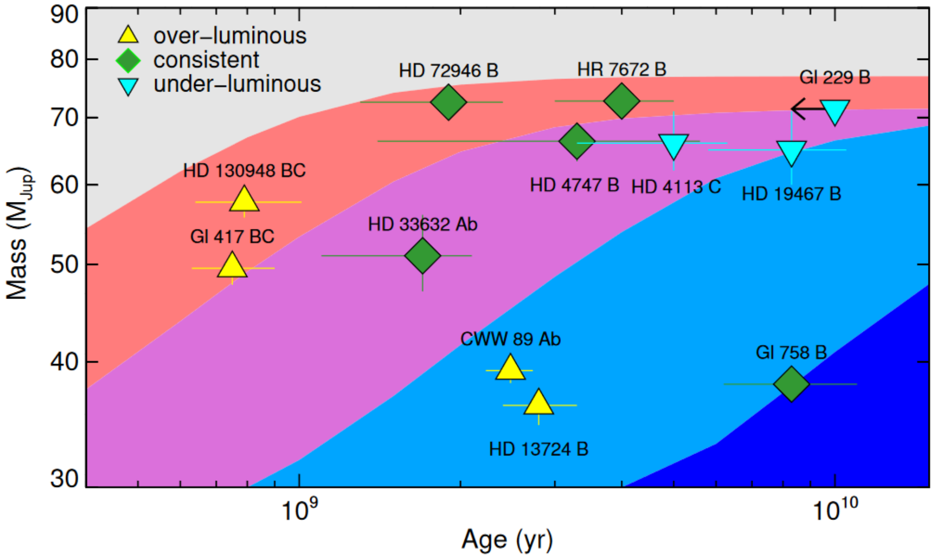

Figure 4.15 All BDs that have dynamically-measured masses, directly-determined luminosities, and independently determined ages. Each system is classified as being either consistent with models ($<1\sigma$ discrepant; green diamonds) or over- or under-luminous as compared to models given the mass and host star age (up-pointing and down-pointing triangles, respectively). In fact, of the discordant objects, all but HD 19467 B (1.1σ) are $\geq 2.5\sigma$ discrepant with models. The background shading indicates effective temperature as predicted by hybrid Saumon & Marley (2008) evolutionary models ranging from >1800 K (gray), 1800–1400 K (red; ≈L4–L9), 1400–1100 K (purple; ≈T0–T4), 1100–600 K (light blue; ≈T5–T7), and <600 K (dark blue). While no trend is apparent with temperature, the ensemble of measurements is consistent with observations favoring higher luminosities than models for young and low-mass BDs, lower luminosities for old and high-mass BDs, and agreement in between. Note that the average masses of HD 130948 BC (Dupuy et al., 2009) and Gl 417 BC (Dupuy et al., 2014) are plotted because those benchmark tests are based on total, not individual, masses.

Table 4.10. Comparison of observed and BD evolutionary model-derived fundamental properties of benchmark systems.

Object	System age			Companion mass			Companion $\log(L_{\rm bol}/L_{\odot})$		
	Host star (Gyr)	BD model (Gyr)	Δ	Dynamical ($M_{\rm Jup}$)	BD model ($M_{\rm Jup}$)	Δ	Observed (dex)	BD model (dex)	Δ
Gl 417 BC[a]	$0.74^{+0.12}_{-0.15}$	0.62 ± 0.03	$+0.9\sigma$	$110.1^{+1.4}_{-1.5}$	118 ± 6	-1.2σ	-4.23 ± 0.03	-4.45 ± 0.09	$+1.6\sigma$
HD 130948 BC[a]	$0.79^{+0.14}_{-0.15}$	$0.44^{+0.04}_{-0.03}$	$+3.0\sigma$	116.0 ± 0.4	141^{+6}_{-5}	-4.5σ	-3.93 ± 0.06	-4.27 ± 0.11	$+3.0\sigma$
HD 33632 Ab	$1.7^{+0.4}_{-0.6}$	$1.7^{+0.4}_{-0.6}$	$+0.1\sigma$	$50^{+5.6}_{-5}$	52 ± 8	-0.2σ	-4.69 ± 0.07	$-4.65^{+0.14}_{-0.11}$	-0.2σ
HR 7672 B[b]	1.9 ± 0.3	4^{+3}_{-2}	-1.3σ	72.7 ± 0.8	$71.6^{+1.8}_{-1.3}$	$+0.7\sigma$	-4.25 ± 0.05	$-4.21^{+0.06}_{-0.07}$	-0.5σ
HD 72946 B	$1.9^{+0.5}_{-0.6}$	$1.8^{+0.5}_{-0.6}$	-0.0σ	72.5 ± 1.3	$73.7^{+1.8}_{-1.5}$	-0.4σ	-4.12 ± 0.02	$-4.21^{+0.13}_{-0.10}$	$+0.7\sigma$
CWW 89 Ab[c]	2.48 ± 0.25	$0.35^{+0.09}_{-0.12}$	$+7.3\sigma$	39.2 ± 1.1	74 ± 3	-3.9σ	-4.19 ± 0.14	-5.22 ± 0.09	$+6.1\sigma$
HD 13724 B	$2.8^{+0.4}_{-0.5}$	1.18 ± 0.17	$+3.8\sigma$	$36.2^{+1.6}_{-1.5}$	55^{+4}_{-5}	-3.6σ	-4.78 ± 0.07	$-5.44^{+0.12}_{-0.13}$	$+3.9\sigma$
HD 4747 B[b]	$2.9^{+0.4}_{-0.5}$	$3.6^{+1.4}_{-2.0}$	-0.4σ	66 ± 3	66^{+5}_{-3}	$+0.1\sigma$	-4.55 ± 0.08	$-4.59^{+0.06}_{-0.08}$	$+0.3\sigma$
HD 4113 C[d]	$5.0^{+0.7}_{-0.8}$	$14.0^{+1.0}_{-0.7}$	-4.4σ	66^{+5}_{-4}	23 ± 5	$+6.1\sigma$	-6.30 ± 0.22	$-4.72^{+0.18}_{-0.17}$	-5.2σ

Table 4.10 (cont'd)

Object	System age			Companion mass			Companion $\log(L_{\rm bol}/L_{\odot})$		
	Host star (Gyr)	BD model (Gyr)	Δ	Dynamical ($M_{\rm Jup}$)	BD model ($M_{\rm Jup}$)	Δ	Observed (dex)	BD model (dex)	Δ
HD 19467 B	$5.1^{+1.3}_{-1.7}$	9^{+2}_{-3}	-1.4σ	$65.4^{+5.9}_{-4.6}$	53^{+6}_{-5}	$+1.8\sigma$	-5.16 ± 0.08	$-4.7^{+0.2}_{-0.3}$	-1.3σ
Gl 758 B	$7.5^{+1.8}_{-1.4}$	8.8 ± 0.6	-0.8σ	38.0 ± 0.75	35 ± 4	$+0.8\sigma$	-6.07 ± 0.03	$-5.98^{+0.09}_{-0.16}$	-0.7σ
Gl 229 B	< 10	$\cdots$	$\cdots$	71.4 ± 0.6	64.78 ± 0.10	$+10\sigma$	-5.208 ± 0.007	$-4.52^{+0.06}_{-0.07}$	-11σ

Note. — The source of the data tabulated here is from this work unless otherwise noted. For Gl 229 B, model results are given at a fixed age of 10 Gyr.

[a] For Gl 417 BC and HD 130948 BC, results in the system age and companion mass columns are based on their total dynamical mass rather than individual masses. In the companion $\log(L_{\rm bol}/L_{\odot})$ column, we quote results for the fainter component assuming the model-derived mass ratio for the system. We use the most recent orbital fits from Dupuy & Liu (2017), replacing the *Hipparcos* parallaxes used in that work with *Gaia* EDR3 parallaxes.

[b] The mass and luminosity measurements of HD 4747 B and HR 7672 B are from Brandt et al. (2018), and we have updated the stellar activity ages here using the same methods as in Brandt et al. (2018).

[c] The mass, cluster age, and luminosity measurements of CWW 89 Ab are from Beatty et al. (2018).

[d] The dynamical mass of HD 4113 C is from Cheetham et al. (2018), the luminosity is from this work, and we have computed a stellar activity age posterior for consistency with other systems presented here.

4.8.5 Emerging Trends in Benchmark Tests

We have derived masses and ages for six benchmark BDs. We now place these six systems in context, combining them with other benchmark BD systems with measured masses, ages, and luminosities.

Two of these are HD 130948 BC (Dupuy et al., 2009) and Gl 417 BC (Dupuy et al., 2014), BD+BD binaries orbiting young, solar-type host stars. In both cases their total masses have been measured dynamically. Previous results were based on *Hipparcos* parallactic distances, so we have checked whether newer *Gaia* EDR3 parallaxes would significantly change the results. For HD 130948, the *Gaia* parallax is consistent within 0.1%, implying negligible changes. But for Gl 417, the *Gaia* EDR3 parallax is 3.4% (3.5σ) smaller than the *Hipparcos* parallax used by Dupuy et al. (2014). By Kepler's Third Law, $M_{\rm sys} \propto a^3$, the larger semimajor axis implied by the larger *Gaia* distance results in an 11% higher system mass of $110.1^{+1.4}_{-1.5}\,M_{\rm Jup}$ as compared to previous work. This reduces the tension with models, corresponding to a 1.6σ discrepancy in luminosity (models too faint) for the companion Gl 417 C.

Three more objects, HR 7672 B and HD 4747 B (Crepp et al., 2012b, 2016; Brandt et al., 2018) and HD 4113 C (Cheetham et al., 2018; Rickman et. al., in prep), are directly imaged companions that have measured masses using *Gaia* DR2, and where the RV phase coverage is sufficient that *Gaia* EDR3 measurements do not significantly change published results.

The only BD mass benchmark with an age determined by association to a stellar cluster, rather than gyrochronology, is CWW 89 Ab, a transiting BD in Ruprecht 147 (2.48 ± 0.25 Gyr; Torres et al., 2018). We use the secondary-eclipse, irradiation-corrected luminosity from Beatty et al. (2018) and dynamical mass from Carmichael et al. (2019). We limit our discussion to the evolutionary parameter space circumscribed by these older ($\gtrsim$300 Myr), massive ($>$30 $M_{\rm Jup}$) BDs, and do not consider the much younger and/or lower mass objects such as β Pic bc (Nowak et al., 2020; Lagrange et al., 2020; Brandt et al., 2021a) and 2MASS J15104786−2818174 Aab (Triaud et al., 2020).

There are other examples of BD benchmarks in the literature that we do not include in the following due to large uncertainties in their ages or masses. Gl 802 B is a likely old (probable thick-disk member) and over-luminous BD with a mass of $66 \pm 5\,M_{\rm Jup}$ (Ireland et al., 2008). HD 47127 B is a massive ($> 68\,M_{\rm Jup}$) and old (7–10 Gyr) BD companion to a white dwarf (Bowler et al., 2021). And finally, the BD binary ϵ Ind BC has discrepant dynamical mass measurements (Cardoso, 2012; Dieterich et al., 2018) and a poorly-constrained host-star age (King et al., 2010).

We place benchmark results into three broad categories: consistent with evolutionary models, over-luminous, and under-luminous. The over-luminous cases are BDs that are brighter than models predict given their independently measured mass and age. Equivalently, their measured masses are surprisingly low, or their host stars' ages are older than the substellar cooling ages. Such systems include HD 13724 B (cooling ages are 3.8σ younger than the star's activity age) as well as literature systems HD 130948 BC (3.0σ), Gl 417 BC (0.9σ), and CWW 89 Ab (7.3σ). The under-luminous cases are observed to be fainter than model predictions given their mass and age; their dynamical masses are higher than expected. These include Gl 229 B ($\approx 10\sigma$) and HD 19467 B (though at just 1.4σ), as well as HD 4113 C ($\approx 5\sigma$).

Here we are considering only the benchmark results where models are appropriate to each object's corresponding spectral type. We use hybrid Saumon & Marley (2008) models for all objects and ATMO 2020 for late-T objects. While the Burrows et al. (1997) models provide the best match the observed mass, age, and luminosity for some BDs, they lack state-of-the-art cloud and opacity treatments, and have also been shown to overpredict the lithium-depletion mass boundary (Dupuy & Liu, 2017).

Figure 4.15 displays the benchmark test results as a function of mass and age, with model-predicted T_{eff} ranges highlighted with background shading. Table 4.10 lists the quantitative discrepancies from each benchmark test. Each temperature band contains a mixture of benchmark results, suggesting that any significant problems with substellar cooling models are not restricted to a particular surface temperature. Instead, there is a trend of objects being over-luminous at low masses and young ages and under-luminous at high masses and old ages. This more closely resembles a correlation with surface gravity, which increases towards higher masses and older ages (as the radius decreases). This analogy is not precise, as $\log g$ does not actually map one-to-one on a mass-age diagram.

The over-luminosity problem at moderately young ages has been known since the earliest measurement of a substellar mass-age-L_{bol} benchmark (HD 130948 BC; Dupuy et al., 2009). Our addition of HD 13724 B stakes out the lowest mass at which this has now been observed, comparable to the unique mass-radius-age-$L_{\rm bol}$ benchmark CWW 89 Ab. CWW 89 Ab's radius is consistent with evolutionary models for its mass and age, implying that its interior structure agrees with fundamental degeneracy physics. This rules out deposition of excess energy in the deep interior (e.g., from tides) as the cause of the over-luminosity. Beatty et al. (2018) suggest that strong alkali absorption of flux from CWW 89 A could induce a thermal inversion if the BD's cooling is inhibited. This could occur if a high C/O ratio in the BD makes CH_4 and H_2O chemically unfavorable relative to CO. Such a mechanism would not be available to wide companions like HD 13724 B, HD 130948 BC, and Gl 417 BC and is regardless ruled out by the strong H_2O features in their spectra.

Magnetic fields have been unsuccessful at reproducing the over-luminosity problem (Mullan & MacDonald, 2010), and other potential solutions remain elusive. One promising avenue may be to examine the temperature-pressure profile directly through retrieval methods. This has shown evidence for upper atmosphere heating in L dwarfs (e.g., Burningham et al., 2017), and perhaps such surface processes could make some BDs appear over-luminous.

In contrast, the under-luminosity problem (a.k.a., the "over-massive BD problem") is fundamentally different; it can always be explained by unresolved multiplicity. While we place HD 19467 B in this category, its dynamical mass is only 1.8σ different from the model-derived mass. Gl 229 B and HD 4113 C are the true touchstones for the BDs that are much too faint for their mass. Both Gl 229 B and HD 4113 C could be unresolved binaries, although the multi-decade collection of high-quality observations of Gl 229 B greatly restrict what binary scenarios are plausible with a total mass of $71.4\,M_{\rm Jup}$. AO imaging like in Figure 4.2 rules out wide, massive companions ($\gtrsim$0.3 au), such as a $41{+}30\,M_{\rm Jup}$ system. The lack of any perturbations $\gtrsim$2 mas in relative astrometry implies the photocenter orbit of a putative binary must be $\lesssim$0.012 au. For instance, a 10-$M_{\rm Jup}$ companion (contributing negligible flux) could only be orbiting B with $a \lesssim 0.04$ au. A 10-$M_{\rm Jup}$ companion on a 0.04 au circular orbit would cause Gl 229 B to display significant RV variations with a semi-amplitude of $5.4\,{\rm km\,s^{-1}} \times (M_2 \sin i / 10\,M_{\rm Jup})(a/0.04\,{\rm au})^{-1/2}$. A brighter, more massive secondary would have a higher upper limit on a and impart a larger RV semi-amplitude on Gl 229 B (as long as the binary orbit is not near face-on). Such RV monitoring has not been carried out, but if future observations rule out a massive companion, then something more fundamental must be amiss with evolutionary models. This would likely require extreme changes to how models treat properties like the equation of state, heat transport in the interior, or the influence of rotation or magnetism.

In between the two extremes outlined above, BDs with masses of 40–70 $M_{\rm Jup}$ at intermediate ages of 1–4 Gyr seem to agree very well with evolutionary models. Unfortunately, age constraints for these stars are relatively weak: magnetic fields weaken and magnetic braking becomes less efficient, limiting the precision of gyrochronology (e.g., van Saders et al., 2016). BDs with more precise age measurements at a few to several Gyr, whether from BDs in more distant open clusters or from asteroseismic ages of their host stars, would provide stronger tests of BD cooling models in this region of parameter space.

4.9 Conclusions

We have derived masses and orbits for six systems containing BD companions: Gl 229, Gl 758, HD 13724, HD 19467, HD 33632, and HD 72946. Our analysis utilizes long-term RV monitoring, relative astrometry (including new Keck/NIRC2 measurements we report for Gl 229 B), and *Hipparcos*-*Gaia* proper motion accelerations. We summarize our main results below.

1. We measure the most precise masses to-date for the late-T dwarfs Gl 229 B ($71.4 \pm 0.6\,M_{\rm Jup}$) and Gl 758 B ($38.0 \pm 0.8\,M_{\rm Jup}$) and the M-dwarf Gl 229 A ($0.579 \pm 0.007\,M_{\odot}$). Notably, our masses for Gl 229 A and B have uncertainties $\lesssim$1%, despite the fact that the system has been observed for only 10% of the orbital period. We find good constraints on the masses of the other four BD companions, with a typical improvement in precision of a factor of two compared to previously published results. For the mid-T dwarfs HD 13724 B and HD 19467 B we infer masses of $36.2^{+1.6}_{-1.5}\,M_{\rm Jup}$ and $65^{+6}_{-5}\,M_{\rm Jup}$, respectively, while for the L/T transition companion HD 33632 Ab we find $50^{+6}_{-5}\,M_{\rm Jup}$. The mass of $72.5 \pm 1.3\,M_{\rm Jup}$ that we measure for the mid-L dwarf HD 72946 B places the object on the boundary between stars and BDs.
2. We perform mass–$L_{\rm bol}$–age benchmark tests of substellar evolutionary models and

compare the results with six other such systems from the literature. We identify a pattern of BDs being over-luminous at younger ages and lower masses and under-luminous (or over-massive) at older ages and higher masses.

3. The mass and luminosity of Gl 229 B is highly discrepant with modern substellar evolutionary models. We reaffirm that neither an unusually low metallicity for Gl 229 B nor a massive, interior companion are likely to reconcile this discrepancy. It seems more likely that Gl 229 B itself may be an unresolved binary.
4. While companion mass is most directly constrained by astrometric accelerations measured for the host stars, our joint orbital analysis also results in well-measured orbital parameters for most systems. Companion eccentricities range widely from near-circular, like $0.12^{+0.18}_{-0.09}$ for HD 33632 Ab, to moderately eccentric, like 0.335 ± 0.026 for HD 13724 B; and even as high as $0.851^{+0.002}_{-0.008}$ for Gl 229 B.
5. Our 1.2% mass measurement of Gl 229 A provides a strong validation of mass–magnitude relations for low-mass stars at the few-percent level. On the other hand, our dynamical mass for HD 13724 A is 2σ lower than expected from stellar evolution models given its $T_{\rm eff}$ and $L_{\rm bol}$. This illustrates how exoplanet dynamical analyses are seeping into the domain where they provide meaningful constraints on the astrophysics of their host stars.

Gaia EDR3 accelerations and the wealth of relative astrometry and RVs have allowed us to reach mass precisions on these six BDs where the model-testing error budget is dominated by the host star's age. In other words, substellar evolutionary models predict cooling ages (given our high-precision masses) that are comparable to or much better than the age constraints on the host star.

Sub-1% uncertainties in BD masses, like we have demonstrated for Gl 229 B, will be especially valuable for certain applications. High precision masses will aid studies where $\log g$ is important, like atmosphere modeling and retrievals (e.g., combining the analyses of Mollière et al., 2020 and Brandt et al., 2021b). High precision masses would also result in ultra-precise BD-cooling ages that would be especially useful for calibrating stellar age-dating methods at a wide range of ages, in between the usual benchmark stellar clusters. Sub-1% mass precision is also crucial for precisely identifying the mass boundary between stars and BDs, as this is a relatively sharp transition often with small, few-$M_{\rm Jup}$ variation between model predictions (e.g., see Section 7.1 of Dupuy & Liu, 2017).

Gaia EDR3 accelerations continue to broaden the applications of direct imaging studies, now allowing for a 1% dynamical mass for the host star, as well as the companion, in at least the case of the Gl 229 system. Dynamical primary masses may prove useful, in the coming decade, for constraining stellar evolutionary models of the host stars. As the baseline of the *Gaia* mission grows, and with the release of the individual observation epochs in *Gaia* DR4 (expected some years from now), dynamical constraints on BD's will continue to improve, test, and progress our understanding of brown dwarf and giant planet formation.

4.10 Posteriors and Priors of the orbital fits.

Table 4.11. Posteriors of the Gl229 system.

Parameter	Prior Distribution	Posterior $\pm 1\sigma$
Stellar mass	Uniform	$0.579 \pm 0.007\,M_\odot$
Parallax (ϖ)	$173.574 \pm 0.017\,$mas (*Gaia* eDR3)	$173.574 \pm 0.013\,$mas
Barycenter Proper Motions[a]	Uniform	$\mu_\alpha = -145.46 \pm 0.13,\ \ \mu_\delta = -705.83 \pm 0.15\,$mas/yr
HIRES RV Zero Point	Uniform	$-7^{+34}_{-36}\,$m/s
HARPS RV Zero Point	Uniform	$-10^{+34}_{-36}\,$m/s
Howard et. al. 2021 RV Zero Point	Uniform	$0.59^{+0.59}_{-0.15}\,$m/s
RV jitter	Log-flat over [0,300 m/s]	$0.828 \pm 0.045\,$m/s
Parameter	Prior Distribution	Posterior $\pm 1\sigma$
$\sqrt{e}\sin\omega$	Uniform over [-1, 1]	$-0.14^{+0.82}_{-0.21}$
$\sqrt{e}\cos\omega$	Uniform over [-1, 1]	$0.85^{+0.05}_{-1.4}$
Semi-major axis (a)	$1/a$ (log-flat)	$33.3^{+0.4}_{-0.3}\,$A.U.

Table 4.11 (cont'd)

Parameter	Prior Distribution	Posterior $\pm 1\sigma$
Inclination (i)	$\sin i$ (geometric)	$7.7^{+7.6}_{-4.4}$ degrees
PA of ascending node	Uniform	-29^{+13}_{-140} degrees
Mean Longitude at $t_{\rm ref}$ ($\lambda_{\rm ref}$)	Uniform	-57^{+140}_{-13} degrees
BD Mass (M)	$1/M$ (log-flat)	$71.4 \pm 0.6\, M_{\rm Jup}$
Eccentricity (e)	(derived quantity)	$0.851^{+0.002}_{-0.008}$
Argument of Periastron (ω)	(derived quantity)	-9^{+140}_{-13} degrees
Periastron Time (T_0)	(derived quantity)	2466912^{+97}_{-63} days
Period	(derived quantity)	86909^{+1900}_{-1700} days
		$237.9^{+5.1}_{-4.6}$ years
`orvara` Reference Epoch ($t_{\rm ref}$)	2455197.50 BJD	··· ···

Note. — Orbital elements all refer to orbit of the companion about the barycenter. The orbital parameters for the primary about each companion are identical except $\omega_A = \omega + \pi$. We use $\pm\sigma$ to denote the 1σ Gaussian error about the median when the posteriors are approximately symmetric. Otherwise, we denote the value by median^{+u}_{-l} where u and l denote the 68.3% confidence interval about the median. The reference epoch $t_{\rm ref}$ is not a fitted parameter and has no significance within the fit itself, it is the epoch at which the Mean Longitude ($\lambda_{\rm ref}$) is evaluated.

[a] μ_α and μ_δ refer to the proper motions in right-ascension and declination, respectively.

Table 4.12. Posteriors of the Gl758 system.

Parameter	Prior Distribution	Posterior $\pm 1\sigma$
Stellar mass	Uniform	$1.05^{+0.25}_{-0.23}\,M_\odot$
Parallax (ϖ)	64.07 ± 0.015 mas (*Gaia* eDR3)	64.0703 ± 0.00253 mas
Barycenter Proper Motions[a]	Uniform	$\mu_\alpha = 81.05^{+0.19}_{-0.27}$ mas/yr & $\mu_\delta = 162.15^{+0.51}_{-0.36}$ mas/yr
APF RV Zero Point	Uniform	110^{+20}_{-25} m/s
Tull/McD RV Zero Point	Uniform	90^{+20}_{-25} m/s
HIRES RV Zero Point	Uniform	$0.9^{+1.1}_{-0.58}$ m/s
RV jitter	Log-flat over [0,300 m/s]	0.852 ± 0.038 m/s
Parameter	Prior Distribution	Posterior $\pm 1\sigma$
$\sqrt{e}\sin\omega$	Uniform over [-1, 1]	0.17 ± 0.26
$\sqrt{e}\cos\omega$	Uniform over [-1, 1]	$-0.37^{+0.31}_{-0.18}$
Semi-major axis (a)	$1/a$ (log-flat)	$29.7^{+5.3}_{-4.2}$ A.U.

Table 4.12 (cont'd)

Parameter	Prior Distribution	Posterior $\pm 1\sigma$
Inclination (i)	$\sin i$ (geometric)	$51.6^{+4.4}_{-5.4}$ degrees
PA of ascending node	Uniform	$180.6^{+2.8}_{-3.9}$ degrees
Mean Longitude at $t_{\rm ref}$ ($\lambda_{\rm ref}$)	Uniform	48^{+14}_{-11} degrees
BD Mass (M)	$1/M$ (log-flat)	$38.04 \pm 0.74\,M_{\rm Jup}$
Eccentricity (e)	(derived quantity)	0.24 ± 0.11
Argument of Periastron (ω)	(derived quantity)	155^{+35}_{-56} degrees
Periastron Time (T_0)	(derived quantity)	2470102^{+2400}_{-4000} days
Period	(derived quantity)	56270^{+23000}_{-14000} days
		154^{+63}_{-39} years
`orvara` Reference Epoch ($t_{\rm ref}$)	2455197.50 BJD	$\cdots$ $\cdots$

Table 4.13. Posteriors of the HD13724 system.

Parameter	Prior Distribution	Posterior $\pm1\sigma$
Stellar mass	Uniform	$0.95^{+0.08}_{-0.07}\,M_\odot$
Parallax (ϖ)	$23.016 \pm 0.018\,\mathrm{mas}$ (*Gaia* eDR3)	$23.0159 \pm 0.00214\,\mathrm{mas}$
Barycenter Proper Motions[a]	Uniform	$\mu_\alpha = -31.47 \pm 0.019, \quad \mu_\delta = -67.77 \pm 0.016\,\mathrm{mas/yr}$
COR07 RV Zero Point	Uniform	$-20650.6 \pm 4.6\,\mathrm{m/s}$
COR98 RV Zero Point	Uniform	$-20605.9^{+9.4}_{-10}\,\mathrm{m/s}$
COR14 RV Zero Point	Uniform	$-20603 \pm 11\,\mathrm{m/s}$
HARPS RV Zero Point	Uniform	$0.8^{+1.4}_{-0.63}\,\mathrm{m/s}$
RV jitter	Log-flat over [0,300 m/s]	$1.888 \pm 0.057\,\mathrm{m/s}$
Parameter	Prior Distribution	Posterior $\pm1\sigma$
$\sqrt{e}\sin\omega$	Uniform over [-1, 1]	-0.029 ± 0.021
$\sqrt{e}\cos\omega$	Uniform over [-1, 1]	-0.578 ± 0.022

Table 4.13 (cont'd)

Parameter	Prior Distribution	Posterior $\pm 1\sigma$
Semi-major axis (a)	$1/a$ (log-flat)	$12.4^{+0.6}_{-0.5}$ A.U.
Inclination (i)	$\sin i$ (geometric)	$45.1^{+2}_{-1.8}$ degrees
PA of ascending node	Uniform	4.3 ± 2 degrees
Mean Longitude at $t_{\rm ref}$ ($\lambda_{\rm ref}$)	Uniform	161.1 ± 2 degrees
BD Mass (M)	$1/M$ (log-flat)	$36.2^{+1.6}_{-1.5}\,M_{\rm Jup}$
Eccentricity (e)	(derived quantity)	0.34 ± 0.03
Argument of Periastron (ω)	(derived quantity)	-177.1 ± 2.1 degrees
Periastron Time (T_0)	(derived quantity)	2456166 ± 53 days
Period	(derived quantity)	16027^{+830}_{-710} days $43.9^{+2.3}_{-1.9}$ years
`orvara` Reference Epoch ($t_{\rm ref}$)	2455197.50 BJD	$\cdots$ $\cdots$

Table 4.14. Posteriors of the HD19467 system.

Parameter	Prior Distribution	Posterior $\pm 1\sigma$
Stellar mass	$0.953 \pm 0.022\,M_\odot$	$0.953 \pm 0.022\,M_\odot$
Parallax (ϖ)	$31.219 \pm 0.024\,\mathrm{mas}$ (*Gaia* eDR3)	$31.219 \pm 0.013\,\mathrm{mas}$
Barycenter Proper Motions[a]	Uniform	$\mu_\alpha = -7.915^{+0.087}_{-0.08}$ mas/yr & $\mu_\delta = -261.335^{+0.095}_{-0.11}$ mas/yr
HIRES RV Zero Point	Uniform	-110^{+170}_{-88} m/s
HARPS RV Zero Point	Uniform	$2.9^{+1.3}_{-0.37}$ m/s
RV jitter	Log-flat over [0,300 m/s]	$0.92 \pm 0.068\,\mathrm{m/s}$
Parameter	Prior Distribution	Posterior $\pm 1\sigma$
$\sqrt{e}\sin\omega$	Uniform over [-1, 1]	$-0.67^{+0.11}_{-0.09}$
$\sqrt{e}\cos\omega$	Uniform over [-1, 1]	$0.0^{+0.3}_{-0.4}$
Semi-major axis (a)	$1/a$ (log-flat)	$47^{+18}_{-8.1}$ A.U.
Inclination (i)	$\sin i$ (geometric)	$133^{+14}_{-9.6}$ degrees

Table 4.14 (cont'd)

Parameter	Prior Distribution	Posterior $\pm 1\sigma$
PA of ascending node	Uniform	31^{+16}_{-98} degrees
Mean Longitude at $t_{\rm ref}$ ($\lambda_{\rm ref}$)	Uniform	186^{+20}_{-120} degrees
BD Mass (M)	$1/M$ (log-flat)	$65.4^{+5.9}_{-4.6}\,M_{\rm Jup}$
Eccentricity (e)	(derived quantity)	0.54 ± 0.11
Argument of Periastron (ω)	(derived quantity)	-90^{+24}_{-33} degrees
Periastron Time (T_0)	(derived quantity)	2485464^{+13000}_{-3100} days
Period	(derived quantity)	115283^{+73000}_{-29000} days 316^{+200}_{-78} years
`orvara` Reference Epoch ($t_{\rm ref}$)	2455197.50 BJD	$\cdots$ $\cdots$

Table 4.15. Posteriors of the HD33632 system.

Parameter	Prior Distribution	Posterior $\pm 1\sigma$
Stellar mass	$1.1 \pm 0.1\,M_\odot$	$1.086 \pm 0.092\,M_\odot$
Parallax (ϖ)	37.895 ± 0.026 mas (*Gaia* eDR3)	37.8952 ± 0.00547 mas
Barycenter Proper Motions[a]	Uniform	$\mu_\alpha = -144.935^{+0.074}_{-0.067}$ mas/yr & $\mu_\delta = -134.99^{+0.26}_{-0.24}$ mas/yr
Lick RV Zero Point	Uniform	$1^{+1.6}_{-0.74}$ m/s
RV jitter	Log-flat over [0,300 m/s]	$2.52^{+0.19}_{-0.17}$ m/s
Parameter	Prior Distribution	Posterior $\pm 1\sigma$
$\sqrt{e}\sin\omega$	Uniform over [-1, 1]	$0.01^{+0.2}_{-0.21}$
$\sqrt{e}\cos\omega$	Uniform over [-1, 1]	0.1 ± 0.37
Semi-major axis (a)	$1/a$ (log-flat)	$23.6^{+3.2}_{-4.5}$ A.U.
Inclination (i)	$\sin i$ (geometric)	$45.2^{+4.7}_{-11}$ degrees
PA of ascending node	Uniform	$39.3^{+5.7}_{-6.5}$ degrees

Table 4.15 (cont'd)

Parameter	Prior Distribution	Posterior $\pm 1\sigma$
Mean Longitude at t_{ref} (λ_{ref})	Uniform	$-158^{+14}_{-9.5}$ degrees
BD Mass (M)	$1/M$ (log-flat)	$50^{+5.6}_{-5}$ M_{Jup}
Eccentricity (e)	(derived quantity)	$0.12^{+0.18}_{-0.09}$
Argument of Periastron (ω)	(derived quantity)	-0^{+86}_{-140} degrees
Periastron Time (T_0)	(derived quantity)	2468815^{+18000}_{-5800} days
Period	(derived quantity)	39178^{+7900}_{-10000} days 107^{+21}_{-28} years
`orvara` Reference Epoch (t_{ref})	2455197.50 BJD	··· ···

Table 4.16. Posteriors of the HD 33632 system from a joint fit including the stellar companion HD 33632 B.

Parameter	Prior Distributions	Posteriors $\pm 1\sigma$
Stellar mass	$1.1 \pm 0.1\,M_\odot$	$1.084 \pm 0.085\,M_\odot$
Parallax (ϖ)	$37.895 \pm 0.026\,\mathrm{mas}$ (*Gaia* eDR3)	$37.896 \pm 0.026\,\mathrm{mas}$
Barycenter Proper Motions[a]	Uniform	$\mu_\alpha = -144.88^{+0.063}_{-0.059}\,\mathrm{mas/yr}$ & $\mu_\delta = -135.71^{+0.21}_{-0.2}\,\mathrm{mas/yr}$
Lick RV Zero Point	Uniform	$3^{+2.9}_{-1.7}\,\mathrm{m/s}$
RV jitter	Log-flat over [0,300 m/s]	$2.5^{+0.2}_{-0.18}\,\mathrm{m/s}$

Parameter	Prior Distributions	Posterior $\pm 1\sigma$ on BD companion	Posterior $\pm 1\sigma$ on stellar companion
$\sqrt{e}\sin\omega$	Uniform over [-1, 1]	$0^{+0.18}_{-0.2}$	$-0.32^{+0.46}_{-0.22}$
$\sqrt{e}\cos\omega$	Uniform over [-1, 1]	$0.18^{+0.31}_{-0.36}$	$-0.61^{+0.4}_{-0.22}$
Semi-major axis (a)	$1/a$ (log-flat)	$23.3^{+3}_{-3.9}$ A.U.	832^{+250}_{-220} A.U.
Inclination (i)	$\sin i$ (geometric)	$44.8^{+4.8}_{-10}$ degrees	$74.6^{+4.1}_{-11}$ degrees

Table 4.16 (cont'd)

Parameter	Prior Distributions	Posteriors $\pm 1\sigma$	
PA of ascending node	Uniform	39.4 ± 5.3 degrees	$8.6^{+7.4}_{-11}$ degrees
Mean Longitude at $t_{\rm ref}$ ($\lambda_{\rm ref}$)	Uniform	$-158^{+10}_{-8.8}$ degrees	60^{+32}_{-24} degrees
Mass (M)	$1/M$ (log-flat) $0.22 \pm 0.03\,M_\odot$	$49.8^{+5.5}_{-4.7}\,M_{\rm Jup}$	$0.215 \pm 0.029\,M_\odot$
Eccentricity (e)	(derived quantity)	$0.12^{+0.17}_{-0.09}$	$0.56^{+0.27}_{-0.35}$
Argument of Periastron (ω)	(derived quantity)	-1^{+63}_{-120} degrees	-1^{+63}_{-120} degrees
Periastron Time (T_0)	(derived quantity)	2469028^{+15000}_{-4900} days	$6934301^{+4100000}_{-2800000}$ days
Period	(derived quantity)	38522^{+7300}_{-9000} days	$7715574^{+3800000}_{-2900000}$ days
		105^{+20}_{-25} years	21124^{+10000}_{-7900} years
`orvara` Reference Epoch ($t_{\rm ref}$)	2455197.50 BJD	⋯	⋯

Table 4.17. Posteriors of the HD72946 system.

Parameter	Prior Distribution	Posterior $\pm 1\sigma$
Stellar mass	$0.986 \pm 0.027\,M_\odot$	$0.987 \pm 0.026\,M_\odot$
Parallax (ϖ)	$38.981 \pm 0.041\,\mathrm{mas}$ (*Gaia* eDR3)	$38.9803 \pm 0.00707\,\mathrm{mas}$
Barycenter Proper Motions[a]	Uniform	$\mu_\alpha = -133.97 \pm 0.09, \quad \mu_\delta = -136.89 \pm 0.04\,\mathrm{mas/yr}$
SOPHIE RV Zero Point	Uniform	$-29524^{+12}_{-11}\,\mathrm{m/s}$
ELODIE RV Zero Point	Uniform	$0.32^{+0.11}_{-0.082}\,\mathrm{m/s}$
RV jitter	Log-flat over [0,300 m/s]	$2.64 \pm 0.10\,\mathrm{m/s}$
Parameter	Prior Distribution	Posterior $\pm 1\sigma$
$\sqrt{e}\sin\omega$	Uniform over [-1, 1]	0.6616 ± 0.0080
$\sqrt{e}\cos\omega$	Uniform over [-1, 1]	0.226 ± 0.014
Semi-major axis (a)	$1/a$ (log-flat)	$6.445 \pm 0.056\,\mathrm{A.U.}$
Inclination (i)	$\sin i$ (geometric)	$59.5^{+1.2}_{-1.1}$ degrees

Table 4.17 (cont'd)

Parameter	Prior Distribution	Posterior $\pm 1\sigma$
PA of ascending node	Uniform	167.9 ± 2.6 degrees
Mean Longitude at $t_{\rm ref}$ ($\lambda_{\rm ref}$)	Uniform	23.93 ± 0.66 degrees
BD Mass (M)	$1/M$ (log-flat)	$72.5 \pm 1.3\,M_{\rm Jup}$
Eccentricity (e)	(derived quantity)	0.4889 ± 0.0074
Argument of Periastron (ω)	(derived quantity)	71.1 ± 1.2 degrees
Periastron Time (T_0)	(derived quantity)	2455960 ± 8 days
Period	(derived quantity)	5813 ± 38 days
		15.92 ± 0.1 years
`orvara` Reference Epoch ($t_{\rm ref}$)	2455197.50 BJD	··· ···

Table 4.18. Posteriors of the HD 72946 system from a joint fit including the stellar companion HD 72945.

Parameter	Prior Distributions	Posteriors $\pm 1\sigma$
Stellar mass	$0.986 \pm 0.027\,M_\odot$	$0.997 \pm 0.026\,M_\odot$
Parallax (ϖ)	38.981 ± 0.041 mas (*Gaia* eDR3)	38.982 ± 0.041 mas
Barycenter Proper Motions[a]	Uniform	$\mu_\alpha = -132.12^{+0.13}_{-0.12}$ mas/yr & $\mu_\delta = -135.03 \pm 0.11$ mas/yr
SOPHIE RV Zero Point	Uniform	-30347^{+1800}_{-380} m/s
ELODIE RV Zero Point	Uniform	$1.6^{+2.3}_{-1.1}$ m/s
RV jitter	Log-flat over [0,300 m/s]	$2.65^{+0.11}_{-0.097}$ m/s

Parameter	Prior Distributions	Posterior $\pm 1\sigma$ on BD companion	Posterior $\pm 1\sigma$ on stellar companion
$\sqrt{e}\sin\omega$	Uniform over [-1, 1]	0.651 ± 0.00803	$0.25^{+0.21}_{-0.32}$
$\sqrt{e}\cos\omega$	Uniform over [-1, 1]	$0.251^{+0.012}_{-0.013}$	$0.22^{+0.5}_{-0.91}$
Semi-major axis (a)	$1/a$ (log-flat)	6.514 ± 0.056 A.U.	200^{+52}_{-41} A.U.

Table 4.18 (cont'd)

Parameter	Prior Distributions	Posteriors $\pm 1\sigma$	
Inclination (i)	$\sin i$ (geometric)	62.8 ± 1.4 degrees	$95.5^{+2.7}_{-1.2}$ degrees
PA of ascending node	Uniform	$174.2^{+2.7}_{-2.9}$ degrees	$27^{+180}_{-1.5}$ degrees
Mean Longitude at $t_{\rm ref}$ ($\lambda_{\rm ref}$)	Uniform	$23.01^{+0.66}_{-0.61}$ degrees	190^{+14}_{-160} degrees
Mass (M)	$1/M$ (log-flat) $1.1 \pm 0.1\,M_\odot$	$72.1^{+1.4}_{-1.3}\,M_{\rm Jup}$	$1.089 \pm 0.094\,M_\odot$
Eccentricity (e)	(derived quantity)	$0.4871^{+0.007}_{-0.00786}$	$0.44^{+0.31}_{-0.26}$
Argument of Periastron (ω)	(derived quantity)	68.9 ± 1.1 degrees	68.9 ± 1.1 degrees
Periastron Time (T_0)	(derived quantity)	2455947.1 ± 8.3 days	$2835898^{+280000}_{-320000}$ days
Period	(derived quantity)	5880 ± 43 days	$715396^{+300000}_{-210000}$ days
		16.1 ± 0.12 years	1959^{+820}_{-570} years
`orvara` Reference Epoch ($t_{\rm ref}$)	2455197.50 BJD	$\cdots$	$\cdots$

Chapter 5

Towards Better Masses: A Recalibration of *Hipparcos* 2

5.1 The Need for a Recalibration

The first reduction of the *Hipparcos* mission (hereafter, Hip1; ESA, 1997) provided proper motions, parallaxes, and positions for over 100,000 stars across the sky. These parameters were reduced from repeated observations of the instantaneous position of the stars. For this, *Hipparcos* scanned the sky over its four year mission in a revolving pattern called the scanning law (van Leeuwen & Evans, 1998). The photon-multiplier tube and slit arrangement of the satellite meant that each observation is precise in the direction along the scan, and imprecise perpendicular to the scan. In practice, each observation is thus considered one dimensional. The Intermediate Astrometric Data (IAD) are required to reconstruct the final catalog parameters and errors. The IAD collectively consist of: the one-dimensional residuals (with respect to the catalog skypath) in the along-scan directions; the on-sky angle of the scan; the time each scan was collected; and the associated standard error of each residual measurement (hereafter, along-scan error). Each version of the *Hipparcos* catalog has its own IAD.

The van Leeuwen (2007a) *Hipparcos* reduction (hereafter, Hip2), was a new reduction of the same observational data, which reported significant improvements, in part due to an improved satellite attitude reconstruction (van Leeuwen & Fantino, 2005). This led to formal errors on the final parameters (positions, parallaxes, and proper motions) in the catalog better than Hip1 by up to a factor of four. The Hip2 IAD have slightly different residuals (compared to Hip1), and importantly: dramatically smaller along-scan errors, such that some scans are weighted substantially more than other scans during the fitting of the skypath parameters.

In this chapter, we show that including a constant residual offset and cosmic dispersion in the Hip2 IAD creates a new catalog with significantly better agreement with *Gaia* EDR3. The agreement is improved at a level that is nearly sixty Gaussian sigma above random chance, even after accounting for the addition of these two free parameters. The contents of this chapter have been submitted for publication in the Royal Astronomical Society Techniques and Instruments (RASTI), under the title “Statistical properties of Hipparcos 2, caveats on its use, and a recalibration of the Intermediate Astrometric Data” (G. Brandt, D. Michalik, & T. Brandt, 2022).

5.2 Motivation

Both *Hipparcos* reductions are distributed as two sets of astrometric data: the catalog which describes the best-fit skypath via five (or more) astrometric parameters; and the IAD, which gives the difference (residual) between the observed on-sky position and the best-fit skypath. The five astrometric parameters reported by the catalog are those of the χ^2 (chi-squared) minimizing solution (see e.g., Section 2.1 of Brandt et al., 2021c), where $\chi^2 = \sum_i r_i^2/\sigma_i^2$. r_i is the along-scan residual for the $i^{\rm th}$ observation, and σ_i is the corresponding along-scan error. If the data and errors are well-behaved, the residuals with respect to the best-fit solution across every star should have mean zero. There are nearly fourteen million residuals across the entire IAD and so the mean can be easily computed with high formal precision.

The mean of the Hip1 residuals is consistent with zero. It is -0.002 mas with a formal uncertainty on the mean of 0.01 mas: it is displaced from zero by only 0.2 Gaussian sigma. However, the Hip2 residuals are surprising. The mean of the distribution of residuals is -0.12 mas while the formal uncertainty on this mean is 0.003 mas. This is an offset that is significant at 40 Gaussian sigma due to the millions of samples. The fact that Hip2 has an unexpectedly non-zero residual offset motivates the following ostensibly absurd proposal: could adding this residual offset, or a similar value fit as a free parameter, improve the catalog?

Additionally, we investigate a second free parameter to improve the catalog: inflating each individual along-scan error with a global cosmic dispersion added in quadrature (cosmic dispersion is referred to as astrometric excess noise in *Gaia*). Inflating the errors is motivated by 1: the evidence that the errors in *Hipparcos* 2 are underestimated (discussed extensively in Brandt, 2018); and 2. the fact that for many stars, *Hipparcos* 2 deflated the along-scan errors so that the goodness-of-fit metric for the astrometric skypath became desirable (van Leeuwen, 2007b; Michalik et al., 2014) – a process often referred to as error renormalization but which is typically only used to inflate errors. Makarov (2022) and Gould et al. (2016) (see Section 5) additionally hint to subtleties in the *Hipparcos* error profile. Note that inflating the uncertainties on the fitted astrometric parameters is not equivalent to inflating the individual IAD along-scan errors, because the latter results in a more even weighting of each scan to which the skypath is fitted (and therefore *different* astrometric parameters). We stress that an investigation of whether the along-scan errors should be inflated is, a-priori, difficult without *Gaia* EDR3 parallaxes and proper motions (or other ground-truth values). ESA (1997) were able to estimate a 0-20% error inflation for Hip1 through various investigations such as comparing the distribution of parallaxes to the expected distribution from a model of the galaxy (see their Table 2). Yet, as we show, *Gaia* allows one to precisely calculate the error inflation needed to match external uncertainties. In the next section, we describe how we recalibrate the *Hipparcos* 2 IAD by inflating errors and adding a residual offset.

5.3 Methods and Data

When we refer to Hip2 as-is, we mean the IAD corresponding to the van Leeuwen (2007a) catalog when it was published[1]. Adding a residual offset and cosmic dispersion to the Hip2 IAD creates an entirely new set of IAD (residuals and along scan errors) for each star. Performing a standard five-parameter fit to these new IAD results in new, recalibrated astrometric parameters. We refer to this process as recalibration and the new catalog and new IAD (together) as Hip2 recalibrated. We recalibrate *Hipparcos* with code[2] based on the open-source astrometric-fitting package `htof` (Brandt et al., 2021c; Brandt & Michalik, 2020). We seek the cosmic dispersion and residual offset that *best* recalibrate Hip2. We define the *best* Hip2 recalibrated catalog as follows:

1. The best catalog has the most accurately estimated uncertainties on parallaxes and proper motions.
2. The best catalog has the best agreement with ground-truth parallaxes and proper motions.

Ideally both 1. and 2. occur simultaneously but we do not enforce that explicitly. In practice, "ground-truth" means a modern measure of parallaxes and proper motions that are substantially more precise than Hip2. Additionally, any ground-truth parameters must be such that lingering systematics affect the data at a level much smaller than the

---

[1] For Hip2 as-is, we use the IAD recently published for the 2014 Java Tool. But these IAD are, for the vast majority of stars, nearly identical to the IAD published in 2007. See comments on this in Brandt et al. (2021c).

[2] Code available by request.

Hip2 errors on those same parameters.

We use *Gaia* EDR3 for the ground-truth parallaxes. The *Gaia* EDR3 parallaxes are on average 40 times more precise than the Hip1 parallaxes (and Hip2 after our recalibration). For ground-truth proper motions, we adopt the long-term proper motions between *Hipparcos* and *Gaia*, from the HGCA (Brandt, 2021). These proper motions are based on the positional displacement between the two missions. For our sample, these long-term proper motions are on average 30 times more precise than the *Hipparcos* proper motions.

Because we use long-term, ∼25-year, proper motions as the ground-truth values for the 3.4-year *Hipparcos* proper motions, we restrict our analysis to the sources that are not accelerating. We identify these with the HGCA, by choosing stars with an acceleration χ^2 less than 10. Furthermore, we limit ourselves to only five-parameter sources (solution type 5), and we remove the 6617 sources where the initially published IAD was flagged as corrupted by Brandt et al. (2021c). Our final sample consists of 67,256 stars and nearly 8 million individual IAD observations.

A natural metric to measure agreement is the likelihood $\mathcal{L}$ of the ground-truth astrometric parameters compared to the *Hipparcos* parameters. For parallax ϖ this is,

$$-2\ln\mathcal{L}_{\varpi} = \sum_i \left(\ln\sigma^2_{\varpi,i} + \underbrace{\frac{(\varpi_{H,i} - \varpi_{G,i})^2}{\sigma^2_{\varpi,i}}}_{\chi^2} \right). \tag{5.1}$$

with

$$\sigma^2_{\varpi,i} = \sigma^2_{\varpi,H,i} + \sigma^2_{\varpi,G,i} \tag{5.2}$$

where $\varpi_{H,i}$ and $\varpi_{G,i}$ are the parallaxes in *Hipparcos* and *Gaia* for star i, and $\sigma_{\varpi,H,i}$ and $\sigma_{\varpi,G,i}$ are their uncertainties. The metric, Equation (5.1), is in essence the formal χ^2 (the right-hand term) with an added regularization (left-hand term). Similar equations describe the likelihood for the two directions of proper motion:

$$-2\ln\mathcal{L}_{\mu\alpha} = \sum_i \left(\ln\sigma^2_{\mu\alpha,i} + \frac{(\mu_{\alpha,H,i} - \mu_{\alpha,G,i})^2}{\sigma^2_{\mu\alpha,i}} \right) \tag{5.3}$$

$$-2\ln\mathcal{L}_{\mu\delta} = \sum_i \left(\ln\sigma^2_{\mu\delta,i} + \frac{(\mu_{\delta,H,i} - \mu_{\delta,G,i})^2}{\sigma^2_{\mu\delta,i}} \right) \tag{5.4}$$

where α and δ refer to right-ascension[3] and declination, respectively.

For any constant cosmic dispersion and constant residual offset, we perturb the *Hipparcos* catalog IAD by those two values, then re-solve for a new, slightly perturbed catalog: each cosmic dispersion and residual offset yields new $\varpi_{H,i}$, $\mu_{\alpha,H,i}$, and $\mu_{\delta,H,i}$ (and errors) for each star. Thus for every recalibration, there is an associated larger or smaller likelihood relative to as-is Hip2.

An important side-effect is that maximizing the likelihood as a function of the cosmic dispersion will enforce a reduced χ^2 near unity. Although a reduced χ^2 of unity is desirable, it is not necessarily desirable to enforce it outright. Thus, instead of adopting that constrained metric exactly, we hold fixed the uncertainties for every refit; and thus a reduced χ^2 of 1 is not explicitly enforced. Therefore, *the only way* that a cosmic dispersion can improve the agreement with the reference astrometry is by bringing the Hipparcos astrometry closer to the Gaia values. Adding a uniform cosmic dispersion to

---

[3] α includes the standard $\cos\delta$ factor. It is often denoted α^*.

all IAD gives each along-scan measurement a more uniform contribution to the fit than in Hip2 as-is, where the uncertainties on different scans can differ more substantially. This different relative weighting of the IAD produces small differences in the best-fit astrometric sky paths. The aim is that these small differences budge the values naturally closer to the *Gaia* EDR3 "ground-truths".

We therefore minimize the right hand terms of Equations (5.1)-(5.4), or in other words, the formal χ^2. Minimizing the three χ^2 (with fixed uncertainties) each separately will yield three different sets of best-fit cosmic dispersions and residual offsets. Additionally we compute the sum of the three individual χ^2 as a combined metric. The best-fit values could differ significantly from one metric to another. Differing best-fit values would be evidence against the hypothesis that a cosmic dispersion and residual offset can improve the catalog. Ideally, there exists a single residual offset and cosmic dispersion who naturally result in a reduced χ^2 of unity across *all four* metrics.

We solve for the best residual offset (Δr) and cosmic dispersion by brute force minimization of the metric. We compute the χ^2 metrics at each trial value of Δr and cosmic dispersion. We test values of Δr over a fine grid between -0.4 mas and 0.8 mas, and cosmic dispersions between 0 mas and 5 mas. The boundaries are such because the global minimum occurs well within that grid; all trial values outside those boundaries yielded steadily worse catalogs. In the following section, we discuss the results of this optimization.

5.4 The Recalibrated Hipparcos 2007 Catalog

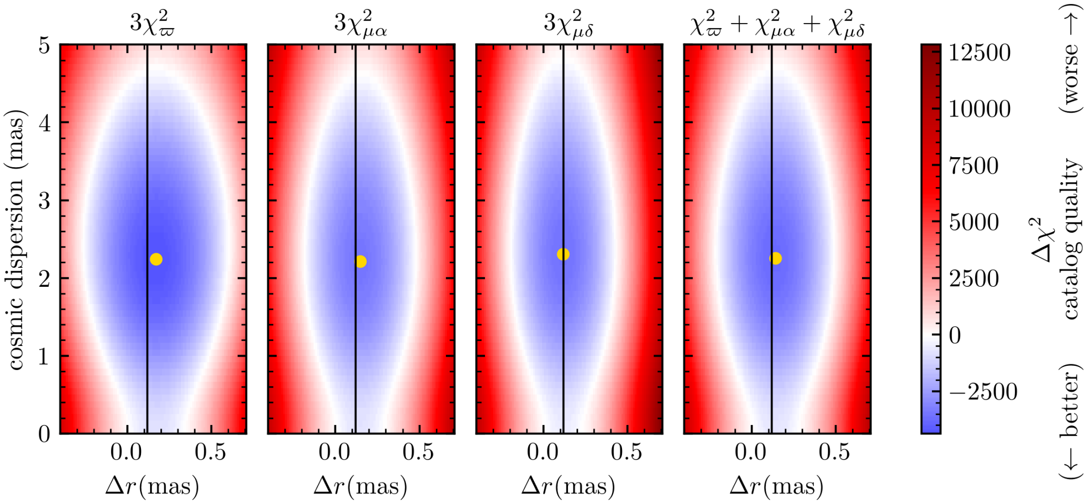

Figure 5.1 The χ^2 between *Hipparcos* and *Gaia* parallaxes (first panel, counting from the left) and *Hipparcos* and *Gaia-Hipparcos* long-term proper motions (second and third panels), by adding a residual offset (Δr) and cosmic dispersion to the Hip2 IAD. We have subtracted the χ^2 of the catalog as-published (0 cosmic dispersion and 0 residual offset) from each panel. The far right panel is the sum of the three metrics. The first three panels have been multiplied by three to match the scale of the far-right sum panel. $\Delta\chi^2 = 0$ represents no improved agreement with *Gaia*. $\Delta\chi^2 < 0$ corresponds to an improved final catalog, and a positive $\Delta\chi^2$ indicates degradation. The thin, vertical line in each panel is the negative of the observed residual offset from the van Leeuwen (2007a) IAD. The yellow dot denotes the minimum metric value in each panel. We adopt the far-right panel as the best metric, yielding $\Delta r = 0.141$ mas and a cosmic dispersion of 2.25 mas as the optimal parameters.

Figure 5.1 shows the four χ^2 metrics (one for parallax, one for each direction of proper motion, and the average of all three metrics) over the evaluation grid of Δr and cosmic dispersion. Each parcel of the grid requires refitting all 67,000 stars in our sample. We computed four hundred million astrometric solutions to construct Figure 5.1.

The best agreement with *Gaia* EDR3 occurs at a $\Delta r = 0.141$ mas and a cosmic dispersion of 2.25 mas (for the sum of all three metrics). Remarkable is that all four metrics yield roughly the same best-fit values, a fact that strongly validates the procedure. Comparing the panels of Figure 5.1 with each other, the best-fit Δr and cosmic dispersion from one metric are within three standard deviations of those from any other metric. The agreement between the *Hipparcos* and *Gaia* catalogs is improved by 3800 points of χ^2 in total (roughly sixty Gaussian sigma), by adopting the quoted Δr and cosmic dispersion while holding the *Hipparcos* uncertainties fixed.[4] Because our metric is χ^2, we can easily calculate the formal errors on our best-fit values from Figure 5.1. We use the χ^2 survival function for 2 degrees of freedom. The best-fit values with 1σ confidence intervals, are $\Delta r = 0.141 \pm 0.008$ mas and a cosmic dispersion of 2.25 ± 0.04 mas.

As another consistency check, we computed the constant offset that brings the residual mean to zero. The residual mean in the as-is catalog is -0.12 mas, but adding back in $+0.12$ mas and refitting does not cause the residuals to be centered at zero. The offset that centers the residuals at 0 is 0.145 mas, in excellent statistical agreement with our best-fit Δr, despite the fact that we never enforced balanced residuals in the metric. Thus, fitting for the best external statistical agreement with *Gaia* EDR3 also improved the internal statistical properties of the IAD.

Figure 5.2 showcases the improved accuracy of the errors on the astrometric parameters. In particular, the figure displays histograms of the three sets of z-scores, for Hip2

---

[4] The 3800 points of improvement is with respect to the parameter errors after inflating the *Hipparcos* IAD by 2.25 mas. Adopting the Hip2 as-is catalog errors results in a much larger improvement of 6600 points of χ^2.

recalibrated, Hip2 as-is, and Hip1:

$$z_{\varpi,i} = \frac{\varpi_{H,i} - \varpi_{G,i}}{\sqrt{\sigma^2_{\varpi,H,i} + \sigma^2_{\varpi,G,i}}} \tag{5.5}$$

$$z_{\mu\alpha,i} = \frac{\mu_{\alpha,H,i} - \mu_{\alpha,\mathrm{HG},i}}{\sqrt{\sigma^2_{\mu\alpha,H,i} + \sigma^2_{\mu\alpha,\mathrm{HG},i}}} \tag{5.6}$$

$$z_{\mu\delta,i} = \frac{\mu_{\delta,H,i} - \mu_{\delta,\mathrm{HG},i}}{\sqrt{\sigma^2_{\mu\delta,H,i} + \sigma^2_{\mu\delta,\mathrm{HG},i}}} \tag{5.7}$$

where HG denotes the long-baseline HGCA proper motions[5] from the HGCA and i indexes the stars in the sample (all ~67,000).

The middle column of Figure 5.2 displays the Hip2 as-is parameters with respect to the *Gaia* EDR3 values. There, the distribution of z-scores is over-dispersed relative to the unit-Gaussian. This is clear evidence for an underestimation of the Hip2 as-is parameter errors. For comparison, we show Hip1 in the right column. It too has evidence for error-underestimation, but to a noticeably lesser degree. Some degree of underestimation is not surprising and simply hints at non-negligible external uncertainties in Hip1 with respect to *Gaia* EDR3. Indeed, the Hip1 consortia were aware of this possibility, and estimated the ratio of external to internal uncertainties to be between 1 and 1.2 (ESA, 1997) – e.g., an error inflation between 0 and 20%. Figure 5.2 (Hip1 panel) shows that the true ratio is 1.07 for the parallaxes and 1.12 for the proper motions, which agrees remarkably well with ESA (1997).

---

[5] The long-baseline proper motions are sometimes referred to as $\Delta\alpha/\Delta t$ in r.a. and $\Delta\delta/\Delta t$ in decl., for example in Brandt, 2018.

The left column of Figure 5.2 shows Hip2 recalibrated. The distribution of the z-scores matches extremely well the unit-Gaussian. This demonstrates the reliability of the recalibrated *parameter* errors. This is noteworthy. The recalibrated parameter z-scores *did not* have to result in a unit-Gaussian, as we never fit for that behavior. Our recalibration was tuned only for the best agreement with *Gaia* EDR3 in the values of the parameters.

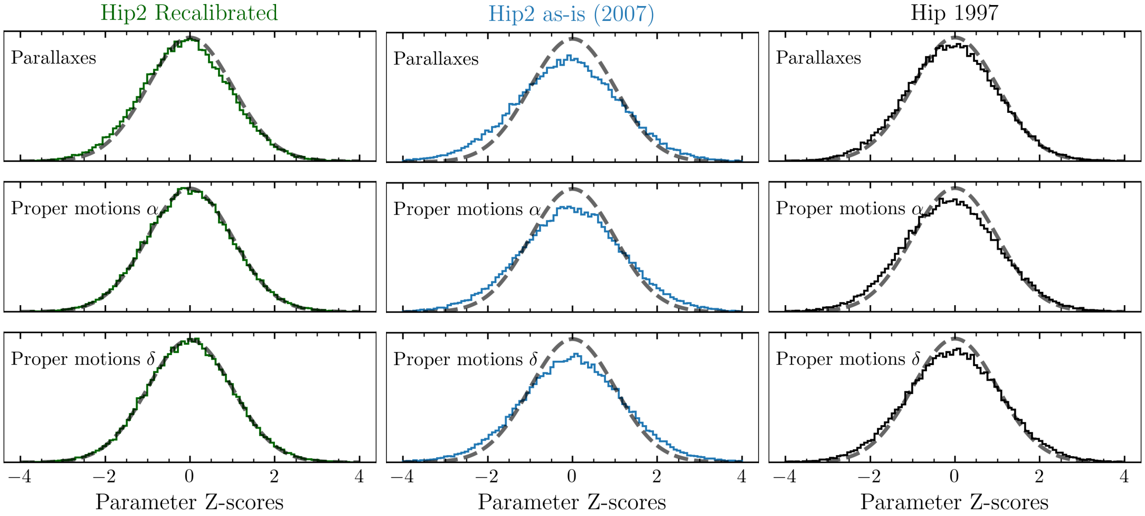

Figure 5.2 Histograms of z-scores of the fitted astrometric parameters relative to *Gaia* EDR3 parallaxes (top row) and HGCA long-baseline proper motions (middle and bottom row). The vertical axis is probability density. Left shows the Hip2 recalibrated IAD (green), and middle is the IAD as-is (blue; showing the less-than-optimal agreement with *Gaia* EDR3). Right is the Hip1 IAD (black), showing how there is still an overdispersion (error underestimation), but to a lesser degree than in Hip2. A unit-Gaussian distribution (grey dashed line) is plotted to guide the eye.

Thus we have the following remarkable summary. We did *not* tune the recalibration to: 1. improve the internal statistical agreement of Hip2 (i.e., remove the residual offset); nor 2. ensure more accurate errors on the *Hipparcos* astrometric parameters. However, both 1. and 2. are improved by the recalibration, even though neither were explicitly optimized for. The residual offset is removed, and the estimate of the astrometric errors is improved (Figure 5.2). Finally, note that a parallax offset of -0.09 mas is visible in the Hip2 as-is, Hip2 recalibrated, and Hip1 data (top row of Figure 5.2). This offset is 19 sigma significant (in terms of the standard error on the estimate of the mean of the ϖ z-scores). We do not investigate this offset in detail.

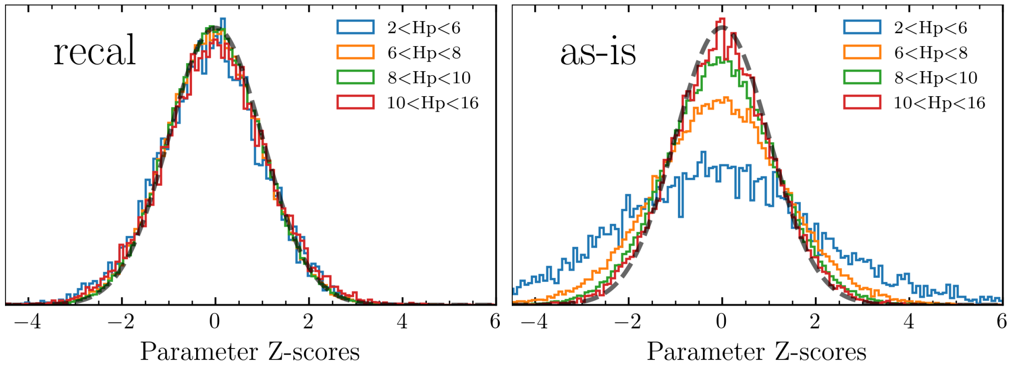

Figure 5.3 Histograms of z-scores of the fitted astrometric parameters relative to *Gaia* EDR3, binned by magnitude. Parallax and proper motions have been binned together. **Left:** the recalibrated data, showing how all stars, independent of magnitude, have recalibrated parameters that agree excellently with *Gaia* EDR3. Note that the same cosmic dispersion and residual offset have been used for *all* stars. **Right:** the as-is data, showing how the errors are most under-estimated for bright stars and have the worst agreement with *Gaia* EDR3. The vertical axis is probability density. The expected unit-Gaussian distribution (grey dashed line) is overplotted. The as-is panel is substantially unchanged even if the error transformation from the published catalog is kept (see Michalik et al., 2014 for an explanation of the transformation).

Figure 5.2 showed that the catalog is improved on average by recalibration. Figure 5.3 shows these improvements across magnitudes. The right panel shows how Hip2 as-is has a varying over-dispersion in the parameter residuals; the overdispersion (and therefore error underestimation) increases with brightness. Despite that complex variation with magnitude, Figure 5.3 (left panel) shows that recalibration unambiguously improves agreement with *Gaia* EDR3 for all magnitudes. The cosmic dispersion, although a constant for all stars, remarkably and naturally compensates perfectly for the varying amount of error underestimation.

5.5 A Focused Discussion of the Hip2 IAD Residuals

The previous section demonstrated that a recalibration of the Hip2 IAD, including a substantial error inflation, improves the agreement of the resulting Hip2 catalog with Gaia. After this recalibration, Hip2 astrometry is then ≈5% more precise, on average, than Hip1 astrometry. In this section we turn our attention to the Hip2 IAD themselves: their statistics, and their suitability for direct orbit fitting. We especially focus on the statistics of their z-scores.

A well-fit model of any data set will have a reduced chi-squared (χ^2/dof) of ≈1, where dof is the degrees of freedom: the number of measurements (hereafter, N) minus the number of fitting parameters. This directly implies that the z-scores of the measurements should have a variance of about dof/N, which is always less than one. Thus, fitting any model to any dataset results in residuals that are technically overfit. To quantify the magnitude of overfitting, we define the overfitting fraction as 1 minus the variance of the distribution of the z-scores; we expect a value $\approx 1 - \mathrm{dof}/N$. We usually quote the overfitting fraction as a percent.

In Hip1, each source in our sample has, on average, 64 observations. Fitting a five-parameter astrometric skypath should result in a z-score variance of $(64-5)/64 = 0.92$, or overfitting of 8%. For Hip2, we would also expect the residual distribution to be narrower than a unit Gaussian, but less so than for Hip1 because each scan in an orbit was included separately[6]. The average number of scans per source is 114, and so fitting a five-parameter astrometric skypath would result in a z-score variance of $1-5/114=0.96$, for an overfitting of $\approx$4%. For the subsequent subsections, it is important to keep in mind that figure of 4% for Hip2.

5.5.1 The residuals and overfitting

The left panel of Figure 5.4 displays the z-scores of the residuals for three hundred bright stars from Hip2. The z-score width of Hip2 as-is (blue) is slightly higher than the expected value of 0.96: the z-scores have a variance of 1.01. However, the distribution is much narrower in the recalibrated case (green) for these bright magnitudes. The distribution has a variance of 0.15, suggesting that a model with roughly 80 to 90 free-parameters was fit – not five – if this variance is taken at face-value. The recalibration procedure does not induce the overfitting. Figure 5.3 shows that the recalibrated errors are correct at every magnitude. Thus, the recalibration merely reveals the overfitting in the along-scan residuals. Instead, in Hip2 as-is, the overfitting was apparent in the astrometric parameters (Figure 5.2), and we show in this section that it is hidden in the residuals because of the deflated along-scan errors. The apparent overfitting is larger for brighter stars. The middle and right panels of Figure 5.4 show these variances for all magnitudes in our sample. They are binned by Hipparcos magnitude (Hp) in constant bins of 300

---

[6] One would expect the 2-4 scans within a single orbit to be correlated, but for the construction of the Hip2 as-is catalog they were assumed to be uncorrelated. Thus the overfitting fraction expected is 4%.

stars. The recalibrated panel reveals that bright stars are significantly more overfit than the catalog on-average. The overfitting fraction varies from ≈85% (Hp<4) to non-existent at faint magnitudes (Hp>11). Averaging across all magnitudes, the overfitting fraction is 22%, much larger than the 4% expected.

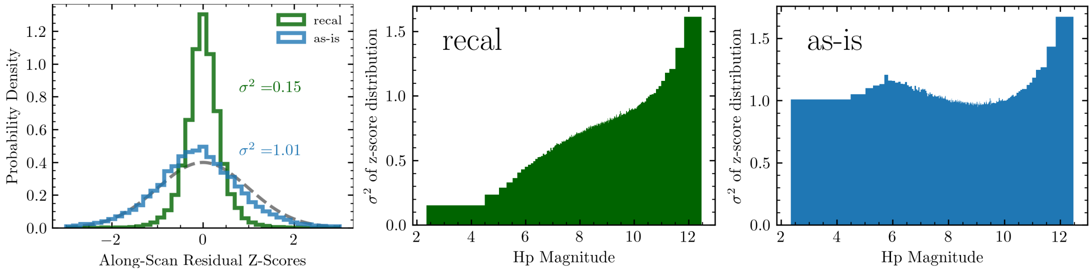

Figure 5.4 **Left:** the z-scores of the residuals of 300 bright stars near Hp magnitude 3. The recalibrated distribution is highly narrowed ($\sigma^2 = 0.15$). **Middle:** the variance of the z-score distribution across all *Hipparcos* magnitudes in our sub-sample (i.e., the green σ^2 in the left-panel, but for the many different magnitude bins). Each bin contains 300 stars. The apparent overfitting revealed by the recalibration increases with increasing brightness (decreasing magnitude). **Right:** the same as the middle panel but for Hip2 as-is. **Note:** the first magnitude bin (magnitudes ≈2 to ≈4) in both middle/right panels correspond to the recal/as-is distributions in the left panel. The as-is data do not include the error transformation (see Michalik et al., 2014), which forced the reduced χ^2 of every star to unity.

An overfitting of 22% instead of 4% implies either that the residuals are suppressed (i.e., structure has been removed by overfitting), or that the recalibrated along-scan errors for the on-sky positions are overestimated. But the recalibrated along-scan errors cannot be deflated without degrading the agreement of the resulting astrometric catalog with *Gaia* EDR3. The presence of modest overfitting neatly explains the residuals' narrow distribution of z-scores. In other words, significantly more parameters were fit to Hip2 than five astrometric parameters per star; how many additional parameters exactly is unknown to us. This fact has important consequences for the expected behavior of the residuals and in what contexts they can be used.

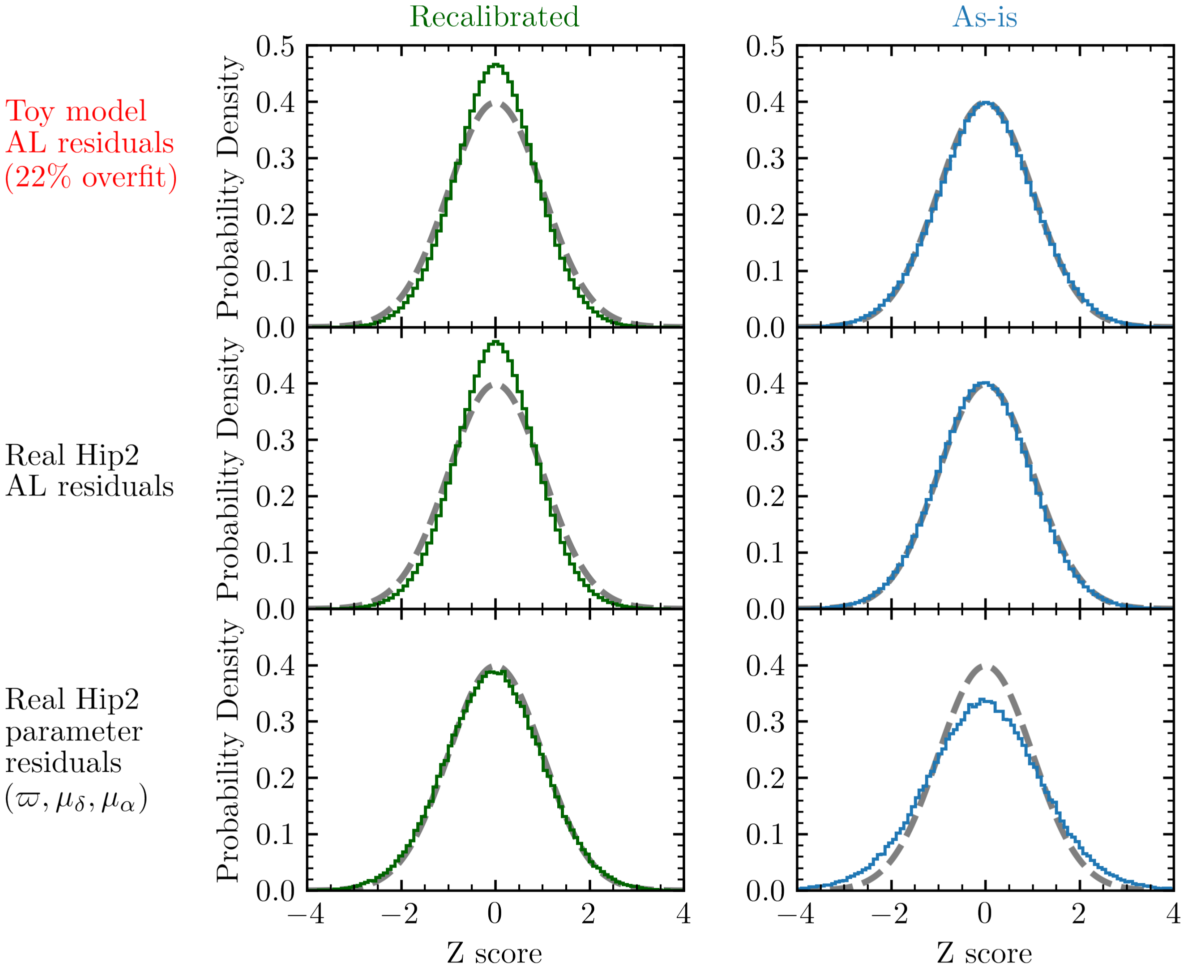

Figure 5.5 A pictographic table where the left column is recalibrated, and the right column is as-is. **Top row:** The "toy model" along-scan residuals. These are the z-scores of the recalibrated toy model (left; green) and the toy model as-is (right; blue). Each panel has a light-grey unit-Gaussian overplotted. Note that the ideal *residuals* distribution is not exactly the unit-Gaussian plotted, rather it is one with a variance of 0.96. **Middle row:** the actual Hip2 residuals, with the recalibrated Hip2 IAD on the left (green) and the IAD as-is on the right (blue). The width, in terms of the z-score, of the recalibrated distribution is 0.881 (variance $0.881^2 \approx 0.78$), revealing the presence of 22% over-fitting inherent to the 2007 IAD. **Bottom row:** The real Hip2 parameter residuals, as in Figure 5.2, but with parallaxes and proper motions binned together. Note how the recalibrated parameters (bottom left) match the unit-Gaussian, while the recalibrated residuals (middle left) deviate from expectation. The situation is reversed for the as-is data.

In the search for an explanation of the apparent overfitting, we turn to the attitude reconstruction. (The additional two parameters fit to 7.7 million residuals in our recalibration cannot explain the 22% overfitting.) One reason that Hip2 achieved smaller formal errors was because of an improved reconstruction of the satellite attitude (van Leeuwen, 2007b). The satellite attitude reconstruction operation fit one free parameter (called a node) roughly every sixty seconds, over the entire 3.4-year baseline of *Hipparcos*. van Leeuwen (2007b) states that the average interval between nodes is 64 seconds (section 9.2.3), and Section 3.3. of van Leeuwen & Fantino, 2005 implies 66 seconds. A 64 or 66 second interval over 3.4 years means that 1.6 to 1.7 million points were fit for the satellite attitude reconstruction. There are 13.6 million residuals in the entire Hip2 dataset. 1.6 million points fit corresponds to 12%. The satellite attitude reconstruction, together with the five astrometric parameters per star, can account for ≈3/4 of the 22% overfitting fraction present. Hip1, which lacks this detailed attitude reconstruction, has an overfitting of 11%. The five-parameter astrometric fits account for 8% overfitting, again ≈3/4 of the total overfitting seen. Additionally, in Hip2, the bright stars were weighted more heavily in the satellite attitude reconstruction compared to faint stars. Thus this would also naturally explain why bright stars appear much more overfit than the faint stars.

In Hip2 as-is, the IAD residuals approximate the expected distribution, while the resulting astrometric catalog does not agree with *Gaia* EDR3. Our recalibration produces excellent agreement between the Hip2 astrometric catalog and *Gaia* EDR3, while it results in an IAD distribution considerably narrower than expected. Figure 5.5 shows these facts in a pictographic table: the middle and bottom rows show the distribution of IAD residuals and astrometric catalog residuals, respectively; the left column is recalibrated, while the right column shows our recalibration. The top row, a toy model whose behavior matches that of Hip2 (as-is and recalibrated), is the subject of the next subsection.

Figure 5.5 shows empirically that with overfit residuals, one can never have the z-scores of the fitted parameters and the z-scores of the along-scan residuals simultaneously follow the expected distributions (a unit-Gaussian and a slightly narrower Gaussian, respectively). To understand further, we present an informative toy-model replica of the *Hipparcos* 2007 data.

5.5.2 A toy model of overfit data

We address the hypothesis that a large number of free parameters were fit to the Hip2 IAD, such that some unknown function was subtracted from the residuals. We illustrate this hypothesis on toy data and aim to show the following two behaviors.

a. Astrometric parameters fit to overfit residuals can be well-behaved even if the residuals themselves (the on-sky positions) have had some modestly-significant unknown function subtracted away.

b. Astrometric parameters fit to overfit residuals can be improved by refitting with the correct along-scan uncertainties.

Point a., which on its surface is oxymoronic, is in fact expected *if* the overfitting procedure itself leads to a better underlying model of the data. This is always possible, especially if the overfitting procedure is done simultaneously with the fitting of the astrometric skypath (or in some iterative procedure that is approximately simultaneous). In the case of Hip2, the process of developing a better underlying model (skypath) was had at the cost of suppressing structure in the residuals. A non-parametric attitude correction based on the IAD themselves will inevitably include, and project out, a component of the actual sky paths. This reduces the utility of the residuals themselves; but leads to improved astrometric parameters (roughly 5% better on average compared to Hip1).

We create a single set of toy model data that is analagous to Hip2 in order to illustrate this behavior. The toy model is comprised of 30,000 fictitious stars with 108 scans each, for roughly 3 million residuals total. Each star has a sky-path described by a single parameter: the mean of that star's position. Their ground-truth positions are all taken to be 0 mas, and measurement error causes the observations to jitter about that zero position. We take each scan to uniformly have an uncertainty of 1 mas. The residuals therefore are Gaussian random variables with standard deviations of 1 mas. The total dataset is thus 6 million entries long: 3 million residuals with 3 million (identical) along-scan errors. We refer to this dataset as the "perfect Hip2" toy model IAD. d We replicate overfitting in Hip2 by doing the following to the "perfect Hip2" toy IAD. We chunk the residuals into 27 point segments, and fit a fifth order polynomial to every segment. This causes a 22% overfitting fraction. This is intentionally identical to the 22% in the real Hip2 data. We leave the along-scan errors set to 1 mas. We refer to this now-overfit dataset, yet with correct along-scan errors, as the "Hip2 recalibrated" toy IAD. The top-left panel of Figure 5.5 shows the distribution of z-scores (the residuals all divided by 1 mas) in green for the "Hip2 recalibrated" toy IAD. It is narrower than 1, owing to the overfitting. The z-score distribution has a standard deviation of 0.881. Note that $1 - 0.881^2 = 0.22$, equal to the input overfitting fraction.

The 'correct' along-scan errors cannot be known to be 1 mas a-priori. They must be estimated from the data. As well, a metric for 'correct' must be established, which is a-priori difficult. We consider the following metric of 'correct': the residuals' z-score distribution matches a unit Gaussian. It makes little difference whether we adopt the slightly more-exact definition of 'correct' as being a Gaussian of variance 0.96 – not 1. The observer who was handed the "Hip2 as-is" toy IAD, would have found that a (deflated) along-scan error of 0.86 mas achieves correctness[7]. Deflating the along-scan errors to 0.86 mas gives a distribution of residual z-scores that best visually matches a unit-Gaussian. Combining the overfit residuals with the deflated along-scan errors of 0.86 mas, yields our second dataset. We call this dataset the "Hip2 as-is" toy IAD. The top-right panel of Figure 5.5 shows the distribution of z-scores for the "Hip2 as-is" toy IAD. The width is 1, showing nearly perfect agreement with the desired normal distribution.

An along-scan error of 0.86 mas is an underestimate. What has occurred is that structure has been removed from the residuals, and the along-scan errors have been deflated to compensate. But it is nearly impossible to know how much structure has been removed without a comparison dataset: a toy "*Gaia* EDR3" to match the toy

---

[7] Or, slightly more than 0.86 mas if a final distribution width of 0.96 was desired.

“Hip2 as-is” IAD.

For all three toy datasets (“as-is”, “recalibrated”, and “perfect”), we perform a 1-parameter astrometric fit to each star: fitting for the mean of the residuals. We then examine the statistics of the distribution of the fitted parameter for all 30,000 stars. Remember that the toy dataset (prior to overfitting) was constructed so that each star’s position was scattered around 0. So our toy “*Gaia* EDR3” astrometric parameters are simple: positions of 0 mas for every star.

The z-scores produced by fitting astrometric parameters to the toy IAD are nearly identical to those seen in the real data in the bottom panels of Figure 5.5. For the as-is toy data, the distribution of parameter z-scores is over-dispersed. The scatter of the fitted parameters is nearly 0.1 mas (equalling the expected 1 mas/$\sqrt{N}$), but the inferred formal error on each parameter is about 0.086 mas. So fitting parameters to the toy “Hip2 as-is” IAD ostensibly yields “improved” formal errors that are better by roughly 12%. Yet, because of toy “Gaia EDR3” we know that there is no real improvement: the scatter relative to this comparison data is not any better than 0.1 mas.

Recall that the toy “Hip2 recalibrated” IAD has identical residuals as the toy “Hip2 as-is” IAD. Yet the former has 1 mas along-scan errors, which are inflated relative to the as-is data. Fitting astrometric parameters to the toy “Hip2 recalibrated” IAD gives formal uncertainties that match the actual 0.1 mas scatter observed about zero. The distribution of z-scores is accurately Gaussian, just like with the real recalibrated Hip2 data.

An equivalent way to interpret this result is: one can recalibrate and fix the toy "as-is" data by inflating each along-scan error with a cosmic dispersion of 0.5 mas. This brings the total along-scan error back to the correct value of 1 mas. As we showed in Sections 5.3 and 5.4, solving for the correct cosmic dispersion is straightforward if one has access to a very precise comparison dataset of parameters like *Gaia* EDR3. We argue that this toy-model has provided understanding of why our recalibration on actual Hip2 data was successful.

5.6 The Recalibrated IAD in the Context of Orbit Fitting.

We have so far reserved our discussion to the benefits of the recalibration to the IAD as a whole. The recalibrated Hip2 data provide better agreement with *Gaia* EDR3, but they also inform and prove useful for individual planetary systems, in particular, the orbits of stars with known substellar companions. For any planetary system, one way to constrain the orbits of the constituents is to fit directly the reflex motion of the star to the on-sky positions given by *Hipparcos*– as in Zucker & Mazeh (2001); Sozzetti & Desidera (2010); Reffert & Quirrenbach (2011); Sahlmann et al. (2011); Snellen & Brown (2018); Nowak et al. (2020). One can reconstruct the on-sky positions by adding the residuals to the best fitting skypath.

The on-sky positions may be statistically well-behaved for stars where the errors are not severely underestimated. These are stars whose average along-scan error is significantly larger than 2.25 mas, mainly those fainter than $Hp \approx 9$. However, for bright stars like 51 Eri (along-scan errors ≈1 mas) or β Pic (along-scan errors ≈0.8 mas) with very large differences between as-is and recalibrated formal errors, these on-sky positions are poorly behaved statistically. Caution should be heeded in these cases. We quantitatively illustrate why with 51 Eri. Table 5.6 shows the first five rows of the Hip2 as-is IAD and the Hip2 recalibrated IAD. On the right of the dividing vertical line is the IAD as-is published with the Java Tool (see footnote 1), and on the left is the IAD after recalibration. IORB, EPOCH, RES, and SRES are the orbit number, the time in years from 1991.25, the residual, and the along-scan error, respectively. Many columns have been omitted for clarity because, like IORB and EPOCH, they do not change with recalibration. The table highlights how the errors of the on-sky positions of 51 Eri, after recalibration, are roughly 2.5 times larger than the as-is data. The best-fitting sky-path to 51 Eri after recalibration therefore has a much smaller χ^2 of only 24 – a factor of 5 smaller than what is expected from 104 observations and 5 free parameters. It correspondingly exerts much less influence on an orbital fit than with the IAD as-published. Such a potentially hazardous situation is common because using *Hipparcos* and *Gaia* astometry, alongside radial velocities, is the norm for breaking the mass-inclination degeneracy of giant-planet orbits.

Beyond the effects of underestimated uncertainties, the use of the IAD residuals themselves remains problematic for orbit fitting. Hip2 used observed stellar positions to re-

Table 5.1. The first five entries of the Hip2 IAD for 51 Eri (Hip 21547).

IORB	EPOCH	RES(recal)	SRES(recal)	RES(as-is)	SRES(as-is)
167	−1.2032	0.50	2.42	0.11	0.88
167	−1.2032	1.58	2.42	1.20	0.88
208	−1.1533	−0.97	2.43	−1.01	0.91
208	−1.1533	0.06	2.48	0.02	1.04
208	−1.1533	−0.87	2.44	−0.91	0.93

construct the satellite attitude. A physical perturbation to the stellar positions will, in turn, perturb the attitude correction. Real stellar motion residuals cannot be distinguished from noise and will be suppressed as a byproduct of correcting the satellite attitude. A similar effect is well-known in high-contrast imaging, resulting in substantial suppression of planet light as a byproduct of suppressing starlight using postprocessing algorithms (Lafrenière et al., 2007; Soummer et al., 2012). Correcting for this bias can be done by reprocessing the data with synthetic sources added (Lafrenière et al., 2007; Marois et al., 2010a) or by forward-modeling the effects of postprocessing (Brandt et al., 2013; Pueyo, 2016). Unfortunately, in the context of Hipparcos, both of these approaches would require repeating the attitude correction, which is not feasible. We are left with a bias whose sign is clear (residuals from the the best-fit sky path will be suppressed) but whose magnitude is uncertain.

Overall, this implies important caveats when constraining the orbits of giant planets with Hip2. For the brightest (most precise) sources in *Hipparcos*, their errors are really a factor of 2-3 worse than reported in the published Hip2 catalog – reducing their utility.

Moreover, their on-sky positions, even if one wanted to fit to them, are potentially affected by a partial subtraction of real residuals in the sky path due to overfitting. There is, however, a clear way forward: fitting to recalibrated astrometric parameters[8] and not the individual positions implied by the IAD. Indeed, by combining error-inflated Hip2 astrometric parameters with *Gaia* EDR3, Dupuy et al. (2022) derived the strongest mass upper limit for 51 Eri b. If one fits to the recalibrated *astrometric parameters* instead, then the left column of Figure 5.2 proves that the weighting is correct. The complicated magnitude-dependent overfitting fraction of Figure 5.4 can be safely ignored. But, fitting to the astrometric parameters comes with the unfortunate drawback that nearly all of the information about short period planets (periods less than the *Hipparcos* 3.5 year baseline) is lost. However, for long-period planets, fitting to astrometric parameters offers the same constraining power as fitting to the on-sky positions themselves.

5.7 A Merger of Hip1 and Hip2

We now address one final question: if the Hip2 recalibrated astrometric parameters are well behaved, is there a way to combine Hip1 and Hip2 (recalibrated) to yield a final set that is superior to either individually? This is particularly useful for the problem of fitting orbits of massive companions around stars. One would ideally want to fit to a single, final, set of merged astrometry. We investigate two ways to create this final merged reduction.

---

[8] The new astrometric parameters for any star are retrievable from the header of its recalibrated data file produced by `htof`.

5.7.1 An attempted merger between NDAC, FAST, and Hip2 at the level of the IAD.

For Hip1, two independent *Hipparcos* IAD were merged to produce a final set of IAD. These were the NDAC and FAST reductions by the NDAC and FAST data consortia, respectively (Bernstein, 1994; ESA, 1997). So, can we merge the three IAD from Hip2 (recal), NDAC, and FAST, to produce a final set of *Hipparcos* IAD, in the same way that NDAC and FAST were merged together to produce the first *Hipparcos* IAD? We refer to this way of merging as “a merger at the level of the IAD”.

We treat NDAC, FAST, and Hip2 as three independent yet correlated data reductions. We show in this section that a merger at the level of the IAD of all three reductions, because of the overfitting in Hip2, results in residuals whose statistical properties are polluted by overfitting in the same way as Hip2. We found that the level of overfitting was less, because of dilution from the less-overfit NDAC and FAST. But overfitting is still present. We conclude that a three-way merger does not yield useful IAD, and so such a merger is not useful. We now detail the methods of a three-way merger and describe the null result.

We start by formatting the Hip2 IAD so that each observation matches up with each NDAC and FAST observation from Hip1. Within a single data reduction, each star is comprised of many observations, each marked by a time. In a given Hip1 IAD file, NDAC, and FAST reported at most one observation per time (per-star). However, for Hip2 there are usually two to four observations at any one time stamp (see, for instance, the IAD of 51 Eri in Table 1). Hip 27321 has 111 observations in the Hip2 IAD, while there are roughly 33 observations for NDAC and FAST each in Hip1. Before attempting the merger, we average together the Hip2 data at a single time (per-star) with inverse variance weighting. This yields IAD that are equivalent when fitted, yet with only 1 observation per time instead of two to four. We performed this averaging and verified that the resulting astrometric parameters fitted to these IAD are identical to the published catalog. Meaning, the averaging of Hip2 data occurring at identical time-stamps (per-star) makes no difference, yet it crucially allows us to perform a statistical merger per-orbit with NDAC and FAST.

For a single observation of a star, we now have at-most three residuals: one from NDAC, one from FAST, and one from Hipparcos 2. For Hip 27321, we now have about 99 observations: one from Hip2, NDAC, and FAST for every time, across 33 different times. A merger at the level of IAD means combining, at each time, those three residuals together according to the Best-Linear-Unbiased-Estimator (BLUE) weights (see Brandt et al. 2021c, or Section 3.2 of van Leeuwen & Evans 1998). Such weights are calculated uniquely from the covariance matrix for the three observations. In the covariance matrix, we denote NDAC as N, FAST by F, and Hip2 by the letter V. Without loss of generality, we order the matrix such that the indices 1, 2, and 3 refer to NDAC (N), FAST (F), and Hip2 (V), respectively. E.g., $C_{1,2} = \rho_{\rm NF}\sigma_{\rm N}\sigma_{\rm F}$, and $C_{1,3} = \rho_{\rm NV}\sigma_{\rm N}\sigma_{\rm V}$, where ρ denotes the correlation coefficient and σ the along-scan error for that observation. The BLUE weights are

$$w_i = \frac{\sum_j (C^{-1})_{i,j}}{\sum_k \sum_j (C^{-1})_{k,j}} \tag{5.8}$$

with the three weights w_0, w_1, w_2 for combining the three residuals, the along-scan error of the merged residual is

$$\sigma_{\mathrm{MERGED}} = \sqrt{\sum_i \sum_j w_i w_j C_{i,j}} \tag{5.9}$$

So merging the NDAC, FAST, and Hip2 residuals is straightforward given the covariance matrix C for every observation (and its inverse C^{-1}). However, we do not know all the elements of this matrix.

We know the along-scan errors and thus the diagonal elements. The errors are available from the IAD for every observation and reduction. For Hip2, we use the recalibrated along-scan errors. ρ_{NF} is supplied with the Hip1 IAD, and usually falls between 0.2 and 0.8. Unfortunately however, we are missing two of the three correlation coefficients: ρ_{VF} and ρ_{VN}. So to perform the merger, we need to solve for those two unknown correlation coefficients.

Easiest is to calculate those coefficients empirically from the residuals themselves.

That trivial calculation yields:

$$\rho_{\mathrm{VF}} = 0.69 \tag{5.10}$$

$$\rho_{\mathrm{VN}} = 0.66 \tag{5.11}$$

We can merge the vast majority of orbits using the provided NDAC-FAST correlation coefficient (ρ_{NF}) and 0.69, 0.66 for ρ_{VF} and ρ_{VN}. However, for some values of ρ_{NF}, the triplet of correlations is disallowed by the Cauchy-Shwarz (CS) inequality and so the merger is not possible. This will be especially relevant if we want to try other values of ρ_{VF}. The CS inequality implies that given two correlation coefficients (ρ_{NF} and ρ_{VN}) then the third (ρ_{VF}) must lie within the range

$$\rho_{\min} \leq \rho_{\mathrm{VF}} \leq \rho_{\max} \tag{5.12}$$

$$\text{where } \rho_{\min} = \rho_{\mathrm{NF}}\rho_{\mathrm{NV}} - \sqrt{(1-\rho_{\mathrm{NF}}^2)(1-\rho_{\mathrm{NV}}^2)} \tag{5.13}$$

$$\& \ \rho_{\max} = \rho_{\mathrm{NF}}\rho_{\mathrm{NV}} + \sqrt{(1-\rho_{\mathrm{NF}}^2)(1-\rho_{\mathrm{NV}}^2)} \tag{5.14}$$

For problematic orbits, one can set ρ_{VF} to the maximal or minimal value, or any value in-between. We characterize this range as follows, using a CS magnitude, dubbed A_{CS}, that ranges between -1 and 1.

$$\rho_{\mathrm{VF}} = \rho_{\mathrm{NF}}\rho_{\mathrm{NV}} + A_{\mathrm{CS}}\sqrt{(1-\rho_{\mathrm{NF}}^2)(1-\rho_{\mathrm{NV}}^2)} \tag{5.15}$$

One solution for problematic orbits is to replace one of the correlation coefficients with the result of Equation (5.15) (for some value of A_{CS}). Replacing problematic orbits however is just one way to merge the three reductions. This method also relies on trusting the empirical correlation coefficients we calculated. Because we have resolved the issue of disallowed triplets, we could replace the empirical estimates with blind guesses. We outline below the three most reasonable algorithms to find these correlation coefficients. They are ordered in terms of computational complexity, from least to most. Or equivalently, ordered by how much we trust our empirical estimates of ρ_{VF} and ρ_{VN}, from most to least.

1. Use the empirical values $\rho_{\mathrm{VF}} = 0.69$ and $\rho_{\mathrm{VN}} = 0.66$. In the rare edge-case where the correlations form a triplet that is incompatible with the CS-inequality, replace ρ_{VF} with equation (5.15) for some global value of A_{CS}. Try full mergers with all possible values of A_{CS} (i.e, varying between -1 and 1), and take the merger with the best statistical properties.

2. Fix only one of the two unknown correlation coefficients to the empirical value. E.g., $\rho_{\mathrm{VN}} = 0.66$. Then solve for ρ_{VF} with equation (5.15) for some global value of A_{CS}. Try full mergers with all possible values of A_{CS}, and take the merger with the best statistical properties.

3. Trust only the Hip1 provided ρ_{NF}. Then, let ρ_{VN} vary over all possible values (-1 to 1). For each trial value of ρ_{VN}, solve for ρ_{VF} with equation (5.15). Try full mergers with all possible values of A_{CS} and ρ_{VN}. Note that this requires evaluating a grid

of mergers. E.g., 900 mergers are required for 30 values of A_{CS} and 30 of ρ_{VN}.

We implemented all three methods above. For all constructions, the final set of merged data was polluted by overfitting in the same way as Hip2, meaning that the merger yielded residuals with undesirable statistical properties. The amount of overfitting varied but was never reduced to a satisfactory level.

Although we went through great pains to exhaust the three reasonable options outlined, this conclusion is no surprise. As described in Section 5.5.2, when residuals are overfit, real signal is removed and cannot be regained. The underlying model (skypath) may be improved, but those residuals are only useful relative to that model (Hip2) and not necessarily relative to another model (e.g., Hip1). We conclude that it is presently impossible to merge Hip1 and Hip2 together at the level of the intermediate data and individual orbits, in a way that yields new IAD with the desired statistical properties.

A second, albeit higher-level, way of merging Hip1 and Hip2 turned out to be viable and trivial: combining the reductions at the level of the parameters. This does not result in a new set of IAD, but it does result in an improved set of fitted astrometric parameters. We detail these efforts and the resulting merger in the following subsection.

5.7.2 A weighted combination of Hip1 and recalibrated Hip2

Brandt (2018) showed that a weighted combination of Hip1 and Hip2 yields a catalog that is better than either reduction on their own. We perform a weighted combination of Hip1 and Hip2 just as in the HGCA (Brandt, 2018, 2021), except that we use the new, recalibrated Hip2 data. The astrometric parameters (parallax, proper motions) for a star are the weighted combination of the parameters Hip1 and the recalibrated Hip2. So $\mu_{\alpha\,\mathrm{merged}} = (1-f)\mu_{\alpha\mathrm{Hip1}} + f\mu_{\alpha\mathrm{Hip2,\,recal}}$, with f being the weight applied to Hip2. We do a simple grid search for the weight that maximizes agreement with *Gaia* EDR3 proper motions.

The result of that minimization is $f = 0.63$. The left-hand panel of Figure 5.6 shows the resulting distribution of proper motion residuals, $\mu_{\alpha\,\mathrm{merged}} - \mu_{\alpha\,\mathrm{Gaia\,EDR3}}$ and $\mu_{\delta\,\mathrm{merged}} - \mu_{\delta\,\mathrm{Gaia\,EDR3}}$, binned together. We display the raw parameter differences, not normalizing by standard error. This way, it is easy to see by-eye that the 37%/63% weighting of Hip1/Hip2 recalibrated (black) is more performant than either reduction on their own. The combination is 5% better, on average, than Hip2 recalibrated on its own (green); and 11% better than Hip1 on its own (red). This improvement is highly statistically significant, as discussed by Brandt (2018).

For comparison, we repeat the exercise with Hip2 as-is – just as was done in Brandt (2018). The right-hand panel of Figure 5.6 shows the result of the mixture of Hip1 and Hip2. We find, just as Brandt (2018) found, that the optimal mixture is 40/60 Hip1/Hip2 as-is. The slightly smaller weight on Hip2 (when using the as-is reduction) is expected. Given a mixture of Hip2 and Hip1, more weight should be applied to Hip2 recalibrated because it is improved over Hip2 as-is.

The Hip1/Hip2 recalibrated mixture is barely better (only 0.2% better) than a Hip1/Hip2 as-is mixture. This is a direct consequence of Hip2 (recal) being 0.5% better than

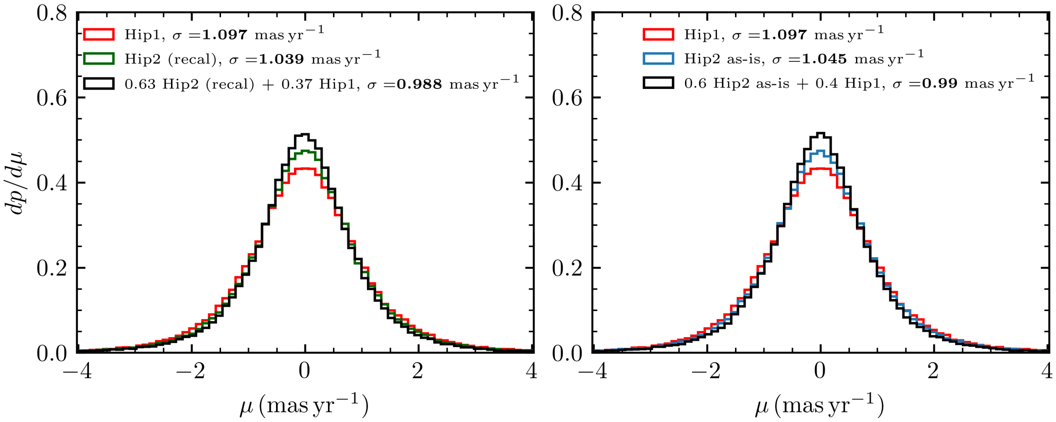

Figure 5.6 The distributions of the differences between *Hipparcos*-*Gaia* EDR3 long-baseline proper motions and *Hipparcos* proper motions, for variations of mixing the *Hipparcos* 1997 and 2007 catalogs. Both right-ascension and declination proper motions are binned together. In the **left panel**, we show these differences using the original ESA (1997) *Hipparcos* catalog (labelled Hip1), the recalibrated version of the van Leeuwen (2007a) *Hipparcos* (labelled Hip2 recal) catalog, and a mixture of the two catalogs at the level of the best-fit parameters. The mixture of the two catalogs has lower residuals than either catalog on its own. In the legend, σ is the standard deviation of the residuals for all points within the plot's limits. The **right panel** shows the same comparisons, but using the original Hip2 catalog (not recalibrated). This Figure is in the same style as Figure 2 in Brandt (2018), where this comparison was made for *Gaia* DR2. The right hand side here slightly differs from Figure 2 in Brandt (2018), because of the slightly different subset of *Hipparcos* sources used.

Hip2 as-is, where one would expect an improvement of $0.3\% = f \cdot (0.5\%) = 0.63 \cdot (0.5\%)$. This yields an important and highly-convenient corollary: Previous works that mix Hip1/Hip2 40/60 with inflated parameter errors need not take into account this new recalibration because the improvement (on average) is not significant. It is possible though that Hip1/Hip2 recalibrated is superior for some sources and use cases. It is possible as well, that for some sources Hip2 recal (alone) is the optimal choice.

5.8 Conclusions

We showed that *Hipparcos* 2007 is a valuable and improved record of the *Hipparcos* mission; and is improved further by including an offset to the IAD residuals and a cosmic dispersion in quadrature to the along-scan errors. The best residual offset is +0.141 mas and the best cosmic dispersion is 2.25 mas. This yields a recalibrated IAD, and fitting skypaths to those recalibrated IAD yields recalibrated astrometric parameters. The recalibration of Hip2 can easily be done on-the-fly by adding those best fit values to and refitting the IAD. Instead of producing an entirely new catalog and recalibrated IAD, we provide a python interface for users to easily recalibrate the data on-demand. We have updated `htof` with the new capability to recalibrate the IAD. `htof` can apply the residual offset and cosmic dispersion to the IAD, perform a refit to those data, and write to file the new best-fit parameters and the recalibrated IAD. This operation takes negligible time to complete. We encourage the reader to consult the recalibration example jupyter notebook[9]. The recalibrated astrometric parameters have better agreement with *Gaia* EDR3 astrometric parameters. This improvement is significant at roughly sixty Gaussian sigma.

We revealed two important caveats with using *Hipparcos* 2, both with the recalibrated and 2007 as-is data. These are

1. The individual position measurements (i.e., the residuals) show strong evidence for being overfit by 22% on average (Figure 5.5), and the overfitting is more severe for

---

[9] `https://github.com/gmbrandt/HTOF/blob/main/examples_recalibrating_hip2.ipynb`

brighter stars. The satellite attitude reconstruction in Hip2 accounts for about half of the overfitting.

2. Because the individual position measurements of Hip2 are overfit, one should exercise caution when fitting an orbit directly to those positions.

We offer recommendations for the major use case of fitting orbits to the intermediate astrometric data. We recommend (in order from most preferred to least):

1. We recommend fitting orbits to the astrometric parameters and proper motion anomalies provided by the *Hipparcos*-*Gaia* Catalog of Accelerations, which uses a mixture of Hip1/Hip2. We showed that a mixture of Hip1/Hip2 recalibrated is more or less identical to Hip1/Hip2 as-is, and so the HGCA is in line with the improvements to Hip2 presented here.

2. In cases where one finds they do not want to use the HGCA, then the next best option is to use the recalibrated Hip2 astrometric parameters.

3. If one needs to fit to the on-sky positions directly (i.e., the IAD residuals + the best-fit skypath), then take caution. Check the statistical properties of the recalibrated residuals and errors for that particular source, or consider using the Hip1 (which is less overfit than Hip2) residuals and skypath. One should perform the same statistical checks to Hip1.

Following our recommendations means that one will be least likely to bias the inferred parameters of companions in an orbital fit. There exist stars where fitting the on-sky positions may be okay. That is the case if the recalibrated residuals happen to be statistically well behaved for that source. As with all data and all recommendations, there is no replacement for individual assessment of a data's validity on a case-by-case basis. The *Hipparcos* 2007 recalibration, in addition to its statistical improvements, can serve as a valuable tool in that assessment.

Bibliography

Ackerman, A. S., & Marley, M. S. 2001, ApJ, 556, 872, doi: 10.1086/321540

Aguilera-Gómez, C., Ramírez, I., & Chanamé, J. 2018, A&A, 614, A55, doi: 10.1051/0004-6361/201732209

Ahmic, M., Croll, B., & Artymowicz, P. 2009, ApJ, 705, 529, doi: 10.1088/0004-637X/705/1/529

Ahuir, J., Brun, A. S., & Strugarek, A. 2020, A&A, 635, A170, doi: 10.1051/0004-6361/201936974

Allard, F., Hauschildt, P. H., Alexander, D. R., Tamanai, A., & Schweitzer, A. 2001a, ApJ, 556, 357, doi: 10.1086/321547

—. 2001b, ApJ, 556, 357, doi: 10.1086/321547

Anderson, E., & Francis, C. 2012, Astronomy Letters, 38, 331, doi: 10.1134/S1063773712050015

Arriagada, P. 2011, ApJ, 734, 70, doi: 10.1088/0004-637X/734/1/70

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: 10.1051/0004-6361/201322068

Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., & Hauschildt, P. H. 2003, A&A, 402, 701, doi: 10.1051/0004-6361:20030252

Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, A&A, 577, A42, doi: 10.1051/0004-6361/201425481

Baranne, A., Queloz, D., Mayor, M., et al. 1996, A&AS, 119, 373

Bardalez Gagliuffi, D. C., Burgasser, A. J., Gelino, C. R., et al. 2014, ApJ, 794, 143, doi: 10.1088/0004-637X/794/2/143

Barrado y Navascués, D., Stauffer, J. R., Song, I., & Caillault, J. P. 1999, ApJL, 520, L123, doi: 10.1086/312162

Beatty, T. G., Morley, C. V., Curtis, J. L., et al. 2018, AJ, 156, 168, doi: 10.3847/1538-3881/aad697

Bell, C. P. M., Mamajek, E. E., & Naylor, T. 2015, MNRAS, 454, 593, doi: 10.1093/mnras/stv1981

Benedict, G. F., Henry, T. J., Franz, O. G., et al. 2016, AJ, 152, 141, doi: 10.3847/0004-6256/152/5/141

Berardo, D., & Cumming, A. 2017, ApJL, 846, L17, doi: 10.3847/2041-8213/aa81c0

Berardo, D., Cumming, A., & Marleau, G.-D. 2017, ApJ, 834, 149, doi: 10.3847/1538-4357/834/2/149

Bernstein, H. H. 1994, A&A, 283, 293

Bertelli, G., Bressan, A., Chiosi, C., Fagotto, F., & Nasi, E. 1994, A&AS, 106, 275

Bertelli, G., Girardi, L., Marigo, P., & Nasi, E. 2008, A&A, 484, 815, doi: 10.1051/0004-6361:20079165

Bertelli, G., Nasi, E., Girardi, L., & Marigo, P. 2009, A&A, 508, 355, doi: 10.1051/0004-6361/200912093

Beuzit, J. L., Vigan, A., Mouillet, D., et al. 2019, A&A, 631, A155, doi: 10.1051/0004-6361/201935251

Bildsten, L., Brown, E. F., Matzner, C. D., & Ushomirsky, G. 1997, ApJ, 482, 442, doi: 10.1086/304151

Binks, A. S., & Jeffries, R. D. 2014, MNRAS, 438, L11, doi: 10.1093/mnrasl/slt141

Boley, A. C., & Durisen, R. H. 2010, ApJ, 724, 618, doi: 10.1088/0004-637X/724/1/618

Bonnefoy, M., Lagrange, A. M., Boccaletti, A., et al. 2011, A&A, 528, L15, doi: 10.1051/0004-6361/201016224

Bonnefoy, M., Boccaletti, A., Lagrange, A. M., et al. 2013, A&A, 555, A107, doi: 10.1051/0004-6361/201220838

Bouchy, F., & Sophie Team. 2006, in Tenth Anniversary of 51 Peg-b: Status of and prospects for hot Jupiter studies, ed. L. Arnold, F. Bouchy, & C. Moutou, 319–325

Bouchy, F., Ségransan, D., Díaz, R. F., et al. 2016, A&A, 585, A46, doi: 10.1051/0004-6361/201526347

Bowler, B. P., Blunt, S. C., & Nielsen, E. L. 2020, The Astronomical Journal, 159, 63, doi: 10.3847/1538-3881/ab5b11

Bowler, B. P., Dupuy, T. J., Endl, M., et al. 2018, AJ, 155, 159, doi: 10.3847/1538-3881/aab2a6

Bowler, B. P., Endl, M., Cochran, W. D., et al. 2021, ApJL, 913, L26, doi: 10.3847/2041-8213/abfec8

Brandt, G. M., Brandt, T. D., Dupuy, T. J., Li, Y., & Michalik, D. 2021a, AJ, 161, 179, doi: 10.3847/1538-3881/abdc2e

Brandt, G. M., Brandt, T. D., Dupuy, T. J., Michalik, D., & Marleau, G.-D. 2021b, ApJL, 915, L16, doi: 10.3847/2041-8213/ac0540

Brandt, G. M., & Michalik, D. 2020, gmbrandt/HTOF: Paper submission release, zenodo, 0.3.1, Zenodo, doi: 10.5281/zenodo.4118572

Brandt, G. M., Michalik, D., Brandt, T. D., et al. 2021c, AJ, 162, 230, doi: 10.3847/1538-3881/ac12d0

—. 2021d, AJ, 162, 230, doi: 10.3847/1538-3881/ac12d0

Brandt, G. M., Dupuy, T. J., Li, Y., et al. 2021e, arXiv e-prints, arXiv:2109.07525. https://arxiv.org/abs/2109.07525

Brandt, T. D. 2018, The Astrophysical Journal Supplement Series, 239, 31, doi: 10.3847/1538-4365/aaec06

—. 2021, arXiv e-prints, arXiv:2105.11662. https://arxiv.org/abs/2105.11662

Brandt, T. D., Dupuy, T. J., & Bowler, B. P. 2018, arXiv e-prints, arXiv:1811.07285. https://arxiv.org/abs/1811.07285

Brandt, T. D., Dupuy, T. J., Bowler, B. P., et al. 2019, A Dynamical Mass of 70 ± 5 Jupiter Masses for Gliese 229B, the First Imaged T Dwarf. https://arxiv.org/abs/1910.01652

Brandt, T. D., Dupuy, T. J., Li, Y., et al. 2021f, AJ, 162, 186, doi: 10.3847/1538-3881/ac042e

Brandt, T. D., McElwain, M. W., Turner, E. L., et al. 2013, ApJ, 764, 183, doi: 10.1088/0004-637X/764/2/183

Brandt, T. D., Kuzuhara, M., McElwain, M. W., et al. 2014, ApJ, 786, 1, doi: 10.1088/0004-637X/786/1/1

Burningham, B., Marley, M. S., Line, M. R., et al. 2017, MNRAS, 470, 1177, doi: 10.1093/mnras/stx1246

Burrows, A., Hubbard, W. B., Lunine, J. I., & Liebert, J. 2001, Reviews of Modern Physics, 73, 719, doi: 10.1103/RevModPhys.73.719

Burrows, A., & Liebert, J. 1993, Reviews of Modern Physics, 65, 301, doi: 10.1103/RevModPhys.65.301

Burrows, A., Marley, M., Hubbard, W. B., et al. 1997, ApJ, 491, 856, doi: 10.1086/305002

Butler, R. P., Vogt, S. S., Laughlin, G., et al. 2017, AJ, 153, 208, doi: 10.3847/1538-3881/aa66ca

Calissendorff, P., & Janson, M. 2018, ArXiv e-prints. https://arxiv.org/abs/1806.07899

Cantat-Gaudin, T., & Brandt, T. D. 2021, arXiv e-prints, arXiv:2103.07432. https://arxiv.org/abs/2103.07432

Cardoso, C. V. V. 2012, PhDT Univ. Exeter

Carmichael, T. W., Latham, D. W., & Vanderburg, A. M. 2019, AJ, 158, 38, doi: 10.3847/1538-3881/ab245e

Casagrande, L., Schönrich, R., Asplund, M., et al. 2011, A&A, 530, A138, doi: 10.1051/0004-6361/201016276

Chauvin, G., Lagrange, A. M., Beust, H., et al. 2012, A&A, 542, A41, doi: 10.1051/0004-6361/201118346

Cheetham, A., Ségransan, D., Peretti, S., et al. 2018, A&A, 614, A16, doi: 10.1051/0004-6361/201630136

Chen, M., Li, Y., Brandt, T. D., et al. 2022, AJ, 163, 288, doi: 10.3847/1538-3881/ac66d2

Chilcote, J., Barman, T., Fitzgerald, M. P., et al. 2015, ApJL, 798, L3, doi: 10.1088/2041-8205/798/1/L3

Chilcote, J., Pueyo, L., De Rosa, R. J., et al. 2017, AJ, 153, 182, doi: 10.3847/1538-3881/aa63e9

Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102, doi: 10.3847/0004-637X/823/2/102

Crepp, J. R., Gonzales, E. J., Bechter, E. B., et al. 2016, ApJ, 831, 136, doi: 10.3847/0004-637X/831/2/136

Crepp, J. R., Johnson, J. A., Howard, A. W., et al. 2014, ApJ, 781, 29, doi: 10.1088/0004-637X/781/1/29

Crepp, J. R., Johnson, J. A., Fischer, D. A., et al. 2012a, ApJ, 751, 97, doi: 10.1088/0004-637X/751/2/97

—. 2012b, ApJ, 751, 97, doi: 10.1088/0004-637X/751/2/97

Crepp, J. R., Rice, E. L., Veicht, A., et al. 2015a, ApJL, 798, L43, doi: 10.1088/2041-8205/798/2/L43

—. 2015b, ApJL, 798, L43, doi: 10.1088/2041-8205/798/2/L43

Currie, T., Thalmann, C., Matsumura, S., et al. 2011, ApJL, 736, L33, doi: 10.1088/2041-8205/736/2/L33

Currie, T., Thalmann, C., Matsumura, S., et al. 2011, The Astrophysical Journal, 736, L33, doi: 10.1088/2041-8205/736/2/l33

Currie, T., Brandt, T. D., Kuzuhara, M., et al. 2020, ApJL, 904, L25, doi: 10.3847/2041-8213/abc631

Currie, T., Brandt, T. D., Kuzuhara, M., et al. 2021, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 11823, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1182304, doi: 10.1117/12.2595001

Datson, J., Flynn, C., & Portinari, L. 2015, A&A, 574, A124, doi: 10.1051/0004-6361/201425000

Delgado Mena, E., Israelian, G., González Hernández, J. I., et al. 2014, A&A, 562, A92, doi: 10.1051/0004-6361/201321493

Delgado Mena, E., Moya, A., Adibekyan, V., et al. 2019, A&A, 624, A78, doi: 10.1051/0004-6361/201834783

Demarque, P., Woo, J.-H., Kim, Y.-C., & Yi, S. K. 2004, ApJS, 155, 667, doi: 10.1086/424966

Dieterich, S. B., Weinberger, A. J., Boss, A. P., et al. 2018, ApJ, 865, 28, doi: 10.3847/1538-4357/aadadc

Dotter, A., Chaboyer, B., Jevremović, D., et al. 2008, The Astrophysical Journal Supplement Series, 178, 89, doi: 10.1086/589654

Dupuy, T. J., Brandt, G. M., & Brandt, T. D. 2022, MNRAS, 509, 4411, doi: 10.1093/mnras/stab3148

Dupuy, T. J., Brandt, T. D., Kratter, K. M., & Bowler, B. P. 2019a, ApJL, 871, L4, doi: 10.3847/2041-8213/aafb31

Dupuy, T. J., & Liu, M. C. 2017, ApJS, 231, 15, doi: 10.3847/1538-4365/aa5e4c

Dupuy, T. J., Liu, M. C., & Ireland, M. J. 2009, ApJ, 692, 729, doi: 10.1088/0004-637X/692/1/729

—. 2014, ApJ, 790, 133, doi: 10.1088/0004-637X/790/2/133

Dupuy, T. J., Liu, M. C., Allers, K. N., et al. 2018, AJ, 156, 57, doi: 10.3847/1538-3881/aacbc2

Dupuy, T. J., Liu, M. C., Best, W. M. J., et al. 2019b, AJ, 158, 174, doi: 10.3847/1538-3881/ab3cd1

Egeland, R., Soon, W., Baliunas, S., et al. 2017, ApJ, 835, 25, doi: 10.3847/1538-4357/835/1/25

Ekström, S., Georgy, C., Eggenberger, P., et al. 2012, A&A, 537, A146, doi: 10.1051/0004-6361/201117751

ESA, ed. 1997, The HIPPARCOS and TYCHO catalogues. Astrometric and photometric star catalogues derived from the ESA HIPPARCOS Space Astrometry Mission, ESA Special Publication

Fabrycky, D. C., & Murray-Clay, R. A. 2010, The Astrophysical Journal, 710, 1408, doi: 10.1088/0004-637x/710/2/1408

Feiden, G. A., & Chaboyer, B. 2013, ApJ, 779, 183, doi: 10.1088/0004-637X/779/2/183

Feng, F., Anglada-Escudé, G., Tuomi, M., et al. 2019, MNRAS, 490, 5002, doi: 10.1093/mnras/stz2912

Feng, F., Butler, R. P., Shectman, S. A., et al. 2020, ApJS, 246, 11, doi: 10.3847/1538-4365/ab5e7c

Filippazzo, J. C., Rice, E. L., Faherty, J., et al. 2015, ApJ, 810, 158, doi: 10.1088/0004-637X/810/2/158

Fischer, D. A., Marcy, G. W., & Spronck, J. F. P. 2014, ApJS, 210, 5, doi: 10.1088/0067-0049/210/1/5

Fontanive, C., Mužić, , K., Bonavita, M., & Biller, B. 2019, MNRAS, 490, 1120, doi: 10.1093/mnras/stz2587

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

Fortney, J. J., Marley, M. S., Saumon, D., & Lodders, K. 2008, ApJ, 683, 1104, doi: 10.1086/589942

Frantseva, K., Mueller, M., Pokorný, P., van der Tak, F. F. S., & ten Kate, I. L. 2020, A&A, 638, A50, doi: 10.1051/0004-6361/201936783

Gagné, J., Mamajek, E. E., Malo, L., et al. 2018, ApJ, 856, 23, doi: 10.3847/1538-4357/aaae09

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1, doi: 10.1051/0004-6361/201629272

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1, doi: 10.1051/0004-6361/201833051

—. 2021, A&A, 649, A1, doi: 10.1051/0004-6361/202039657

Gaidos, E., & Mann, A. W. 2014, ApJ, 791, 54, doi: 10.1088/0004-637X/791/1/54

Gaidos, E., Mann, A. W., Lépine, S., et al. 2014, MNRAS, 443, 2561, doi: 10.1093/mnras/stu1313

Gelman, A., & Rubin, D. B. 1992, Statistical Science, 7, 457, doi: 10.1214/ss/1177011136

Georgy, C., Ekström, S., Eggenberger, P., et al. 2013, A&A, 558, A103, doi: 10.1051/0004-6361/201322178

Gomes da Silva, J., Santos, N. C., Adibekyan, V., et al. 2021, A&A, 646, A77, doi: 10.1051/0004-6361/202039765

Götberg, Y., Davies, M. B., Mustill, A. J., Johansen, A., & Church, R. P. 2016, A&A, 592, A147, doi: 10.1051/0004-6361/201526309

Gould, A., Kollmeier, J. A., & Sesar, B. 2016, arXiv e-prints, arXiv:1609.06315. https://arxiv.org/abs/1609.06315

Goździewski, K., & Migaszewski, C. 2018, ApJS, 238, 6, doi: 10.3847/1538-4365/aad3d3

—. 2020, ApJL, 902, L40, doi: 10.3847/2041-8213/abb881

Gravity Collaboration, Lacour, S., Nowak, M., et al. 2019, A&A, 623, L11, doi: 10.1051/0004-6361/201935253

Gravity Collaboration, Nowak, M., Lacour, S., et al. 2020, A&A, 633, A110, doi: 10.1051/0004-6361/201936898

Gray, R. O., Corbally, C. J., Garrison, R. F., McFadden, M. T., & Robinson, P. E. 2003, AJ, 126, 2048, doi: 10.1086/378365

Grevesse, N., & Sauval, A. J. 1998, SSRv, 85, 161, doi: 10.1023/A:1005161325181

Groff, T. D., Peters, M. A., Kasdin, N. J., et al. 2013, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8864, Techniques and Instrumentation for Detection of Exoplanets VI, ed. S. Shaklan, 88640H, doi: 10.1117/12.2025081

Groff, T. D., Kasdin, N. J., Limbach, M. A., et al. 2015, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9605, Techniques and Instrumentation for Detection of Exoplanets VII, ed. S. Shaklan, 96051C, doi: 10.1117/12.2188465

Haghighipour, N., & Boss, A. P. 2003, ApJ, 598, 1301, doi: 10.1086/378950

Heap, S. R., Lindler, D. J., Lanz, T. M., et al. 2000, The Astrophysical Journal, 539, 435, doi: 10.1086/309188

Helled, R., & Bodenheimer, P. 2010, Icarus, 207, 503, doi: 10.1016/j.icarus.2009.11.023

Herwig, F. 2000, A&A, 360, 952. https://arxiv.org/abs/astro-ph/0007139

Høg, E., Fabricius, C., Makarov, V. V., et al. 2000, A&A, 355, L27

Holmberg, J., Nordström, B., & Andersen, J. 2009, A&A, 501, 941, doi: 10.1051/0004-6361/200811191

Ireland, M. J., Kraus, A., Martinache, F., Lloyd, J. P., & Tuthill, P. G. 2008, ApJ, 678, 463, doi: 10.1086/529578

Isaacson, H., & Fischer, D. 2010, ApJ, 725, 875, doi: 10.1088/0004-637X/725/1/875

Johansson, E. M., van Dam, M. A., Stomski, P. J., et al. 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7015, Adaptive Optics Systems, ed. N. Hubin, C. E. Max, & P. L. Wizinowich, 70153E, doi: 10.1117/12.790198

Kervella, P., Arenou, F., Mignard, F., & Thévenin, F. 2019, A&A, 623, A72, doi: 10.1051/0004-6361/201834371

Kharchenko, N. V. 2001, Kinematika i Fizika Nebesnykh Tel, 17, 409

Kharchenko, N. V., Scholz, R. D., Piskunov, A. E., Röser, S., & Schilbach, E. 2007, Astronomische Nachrichten, 328, 889, doi: 10.1002/asna.200710776

King, R. R., McCaughrean, M. J., Homeier, D., et al. 2010, A&A, 510, A99, doi: 10.1051/0004-6361/200912981

Kozai, Y. 1962, AJ, 67, 591, doi: 10.1086/108790

Kraus, S., Le Bouquin, J.-B., Kreplin, A. e., et al. 2020, ApJL, 897, L8, doi: 10.3847/2041-8213/ab9d27

Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007, ApJ, 660, 770, doi: 10.1086/513180

Lagrange, A. M., Bonnefoy, M., Chauvin, G., et al. 2010, Science, 329, 57, doi: 10.1126/science.1187187

Lagrange, A. M., Meunier, N., Rubini, P., et al. 2019a, Nature Astronomy, 421, doi: 10.1038/s41550-019-0857-1

Lagrange, A. M., Boccaletti, A., Langlois, M., et al. 2019b, A&A, 621, L8, doi: 10.1051/0004-6361/201834302

Lagrange, A. M., Rubini, P., Nowak, M., et al. 2020, A&A, 642, A18, doi: 10.1051/0004-6361/202038823

Lazorenko, P. F., & Sahlmann, J. 2018, A&A, 618, A111, doi: 10.1051/0004-6361/201833626

Lee, J., & Song, I. 2019, MNRAS, 489, 2189, doi: 10.1093/mnras/stz2290

Leggett, S. K. 1992, ApJS, 82, 351, doi: 10.1086/191720

Li, Y., Brandt, T. D., Brandt, G. M., et al. 2021, arXiv e-prints, arXiv:2109.10422. https://arxiv.org/abs/2109.10422

Lindegren, L., & Dravins, D. 2021, arXiv e-prints, arXiv:2105.09014. https://arxiv.org/abs/2105.09014

Lindegren, L., Mignard, F., Söderhjelm, S., et al. 1997, A&A, 323, L53

Lindegren, L., Hernandez, J., Bombrun, A., et al. 2018, arxiv. https://arxiv.org/abs/1804.09366

Lindegren, L., Klioner, S. A., Hernández, J., et al. 2020, arXiv e-prints, arXiv:2012.03380. https://arxiv.org/abs/2012.03380

Liu, M. C., Dupuy, T. J., & Ireland, M. J. 2008, ApJ, 689, 436, doi: 10.1086/591837

Lorenzo-Oliveira, D., Freitas, F. C., Meléndez, J., et al. 2018, A&A, 619, A73, doi: 10.1051/0004-6361/201629294

Luck, R. E. 2017, AJ, 153, doi: 10.3847/1538-3881/153/1/21

Macintosh, B., Graham, J. R., Ingraham, P., et al. 2014, Proceedings of the National Academy of Science, 111, 12661, doi: 10.1073/pnas.1304215111

Maire, A. L., Molaverdikhani, K., Desidera, S., et al. 2020a, A&A, 639, A47, doi: 10.1051/0004-6361/202037984

Maire, A. L., Baudino, J. L., Desidera, S., et al. 2020b, A&A, 633, L2, doi: 10.1051/0004-6361/201937134

Makarov, V. V. 2022, arXiv e-prints, arXiv:2207.12975. https://arxiv.org/abs/2207.12975

Maldonado, J., Eiroa, C., Villaver, E., Montesinos, B., & Mora, A. 2012, A&A, 541, doi: 10.1051/0004-6361/201218800

Males, J. R., Close, L. M., Morzinski, K. M., et al. 2014, ApJ, 786, 32, doi: 10.1088/0004-637X/786/1/32

Mamajek, E. E., & Bell, C. P. M. 2014, MNRAS, 445, 2169, doi: 10.1093/mnras/stu1894

Mamajek, E. E., & Hillenbrand, L. A. 2008, ApJ, 687, 1264, doi: 10.1086/591785

Mann, A. W., Dupuy, T., Kraus, A. L., et al. 2019, ApJ, 871, 63, doi: 10.3847/1538-4357/aaf3bc

Marleau, G. D., & Cumming, A. 2014, MNRAS, 437, 1378, doi: 10.1093/mnras/stt1967

Marleau, G.-D., Mordasini, C., & Kuiper, R. 2019, ApJ, 881, 144, doi: 10.3847/1538-4357/ab245b

Marley, M., & Robinson, T. 2015, Annual Review of Astronomy and Astrophysics, 53, 279, doi: 10.1146/annurev-astro-082214-122522

Marley, M. S., Fortney, J. J., Hubickyj, O., Bodenheimer, P., & Lissauer, J. J. 2007a, ApJ, 655, 541, doi: 10.1086/509759

—. 2007b, ApJ, 655, 541, doi: 10.1086/509759

Marley, M. S., & Robinson, T. D. 2015, ARA&A, 53, 279, doi: 10.1146/annurev-astro-082214-122522

Marois, C., Macintosh, B., Barman, T., et al. 2008, Science, 322, 1348, doi: 10.1126/science.1166585

Marois, C., Macintosh, B., & Véran, J.-P. 2010a, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7736, Adaptive Optics Systems II, ed. B. L. Ellerbroek, M. Hart, N. Hubin, & P. L. Wizinowich, 77361J, doi: 10.1117/12.857225

Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B., & Barman, T. 2010b, Nature, 468, 1080, doi: 10.1038/nature09684

Mayor, M., Pepe, F., Queloz, D., et al. 2003, The Messenger, 114, 20

McLean, I. S., & Sprayberry, D. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, ed. M. Iye & A. F. M. Moorwood, 1–6, doi: 10.1117/12.461785

Michalik, D., Lindegren, L., Hobbs, D., & Lammers, U. 2014, A&A, 571, A85, doi: 10.1051/0004-6361/201424606

Mints, A., & Hekker, S. 2017, A&A, 604, A108, doi: 10.1051/0004-6361/201630090

Miret-Roig, N., Galli, P. A. B., Brandner, W., et al. 2020, A&A, 642, A179, doi: 10.1051/0004-6361/202038765

Mollière, P., Stolker, T., Lacour, S., et al. 2020, A&A, 640, A131, doi: 10.1051/0004-6361/202038325

Mullan, D. J., & MacDonald, J. 2010, ApJ, 713, 1249, doi: 10.1088/0004-637X/713/2/1249

Neves, V., Bonfils, X., Santos, N. C., et al. 2013, A&A, 551, A36, doi: 10.1051/0004-6361/201220574

Nielsen, E. L., Liu, M. C., Wahhaj, Z., et al. 2014, The Astrophysical Journal, 794, 158, doi: 10.1088/0004-637x/794/2/158

Nielsen, E. L., De Rosa, R. J., Wang, J. J., et al. 2020, AJ, 159, 71, doi: 10.3847/1538-3881/ab5b92

Nissen, P. E., Christensen-Dalsgaard, J., Mosumgaard, J. R., et al. 2020, A&A, 640, A81, doi: 10.1051/0004-6361/202038300

Nowak, M., Lacour, S., Lagrange, A.-M., et al. 2020, A&A, 642, L2, doi: 10.1051/0004-6361/202039039

Noyes, R. W., Hartmann, L. W., Baliunas, S. L., Duncan, D. K., & Vaughan, A. H. 1984, ApJ, 279, 763, doi: 10.1086/161945

Oelkers, R. J., Rodriguez, J. E., Stassun, K. G., et al. 2018, VizieR Online Data Catalog, J/AJ/155/39

Oliphant, T. 2006, NumPy: A guide to NumPy, USA: Trelgol Publishing. `http://www.numpy.org/`

Pace, G. 2013, A&A, 551, L8, doi: `10.1051/0004-6361/201220364`

pandas development team, T. 2020, pandas-dev/pandas: Pandas, 1.0.5, Zenodo, doi: `10.5281/zenodo.3509134`

Paunzen, E. 2015, A&A, 580, A23, doi: `10.1051/0004-6361/201526413`

Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3, doi: `10.1088/0067-0049/192/1/3`

Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208, 4, doi: `10.1088/0067-0049/208/1/4`

Paxton, B., Marchant, P., Schwab, J., et al. 2015, ApJS, 220, 15, doi: `10.1088/0067-0049/220/1/15`

Paxton, B., Schwab, J., Bauer, E. B., et al. 2018, ApJS, 234, 34, doi: `10.3847/1538-4365/aaa5a8`

Paxton, B., Smolec, R., Schwab, J., et al. 2019, ApJS, 243, 10, doi: `10.3847/1538-4365/ab2241`

Pecaut, M. J., & Mamajek, E. E. 2013, ApJS, 208, 9, doi: `10.1088/0067-0049/208/1/9`

Perryman, M. 2012, European Physical Journal H, 37, 745, doi: `10.1140/epjh/e2012-30039-4`

Phillips, M. W., Tremblin, P., Baraffe, I., et al. 2020, A&A, 637, A38, doi: `10.1051/0004-6361/201937381`

Pietrinferni, A., Cassisi, S., Salaris, M., & Castelli, F. 2004, ApJ, 612, 168, doi: `10.1086/422498`

—. 2006, ApJ, 642, 797, doi: `10.1086/501344`

Pietrinferni, A., Cassisi, S., Salaris, M., Percival, S., & Ferguson, J. W. 2009, ApJ, 697, 275, doi: `10.1088/0004-637X/697/1/275`

Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, AJ, 156, 123, doi: `10.3847/1538-3881/aabc4f`

Pueyo, L. 2016, ApJ, 824, 117, doi: `10.3847/0004-637X/824/2/117`

Pueyo, L., Soummer, R., Hoffmann, J., et al. 2015, ApJ, 803, 31, doi: 10.1088/0004-637X/803/1/31

Quanz, S. P., Meyer, M. R., Kenworthy, M. A., et al. 2010, ApJL, 722, L49, doi: 10.1088/2041-8205/722/1/L49

Queloz, D., Mayor, M., Naef, D., et al. 2000, in From Extrasolar Planets to Cosmology: The VLT Opening Symposium, ed. J. Bergeron & A. Renzini, 548, doi: 10.1007/10720961_79

Ramírez, I., Fish, J. R., Lambert, D. L., & Allende Prieto, C. 2012, ApJ, 756, 46, doi: 10.1088/0004-637X/756/1/46

Reffert, S., & Quirrenbach, A. 2011, A&A, 527, A140, doi: 10.1051/0004-6361/201015861

Rein, H., & Liu, S. F. 2012, A&A, 537, A128, doi: 10.1051/0004-6361/201118085

Rein, H., & Spiegel, D. S. 2015, MNRAS, 446, 1424, doi: 10.1093/mnras/stu2164

Rice, W. K. M., Lodato, G., Pringle, J. E., Armitage, P. J., & Bonnell, I. A. 2006, MNRAS, 372, L9, doi: 10.1111/j.1745-3933.2006.00215.x

Rickman, E. L., Ségransan, D., Hagelberg, J., et al. 2020, A&A, 635, A203, doi: 10.1051/0004-6361/202037524

Rickman, E. L., Ségransan, D., Marmier, M., et al. 2019, A&A, 625, A71, doi: 10.1051/0004-6361/201935356

Rosenthal, L. J., Fulton, B. J., Hirsch, L. A., et al. 2021, ApJS, 255, 8, doi: 10.3847/1538-4365/abe23c

Roy, V. 2019, arXiv e-prints, arXiv:1909.11827. https://arxiv.org/abs/1909.11827

Sahlmann, J., Ségransan, D., Queloz, D., et al. 2011, A&A, 525, A95, doi: 10.1051/0004-6361/201015427

Sahlmann, J., Burgasser, A. J., Bardalez Gagliuffi, D. C., et al. 2020, MNRAS, 495, 1136, doi: 10.1093/mnras/staa1235

Sahlmann, J., Dupuy, T. J., Burgasser, A. J., et al. 2021, MNRAS, 500, 5453, doi: 10.1093/mnras/staa3577

Saumon, D., & Marley, M. S. 2008, ApJ, 689, 1327, doi: 10.1086/592734

Scholz, R. D. 2016, A&A, 587, A51, doi: 10.1051/0004-6361/201527965

Ségransan, D., Udry, S., Mayor, M., et al. 2010, A&A, 511, A45, doi: 10.1051/0004-6361/200912136

Service, M., Lu, J. R., Campbell, R., et al. 2016, PASP, 128, 095004, doi: 10.1088/1538-3873/128/967/095004

Sheppard, S. S., & Cushing, M. C. 2009, AJ, 137, 304, doi: 10.1088/0004-6256/137/1/304

Smith, B. A., & Terrile, R. J. 1984, Science, 226, 1421, doi: 10.1126/science.226.4681.1421

Snellen, I. A. G., Brandl, B. R., de Kok, R. J., et al. 2014, Nature, 509, 63, doi: 10.1038/nature13253

Snellen, I. A. G., & Brown, A. G. A. 2018, Nature Astronomy, 2, 883, doi: 10.1038/s41550-018-0561-6

Soderblom, D. R. 2010, Annual Review of Astronomy and Astrophysics, 48, 581, doi: 10.1146/annurev-astro-081309-130806

Soubiran, C., Le Campion, J.-F., Brouillet, N., & Chemin, L. 2016, A&A, 591, A118, doi: 10.1051/0004-6361/201628497

Soummer, R., Pueyo, L., & Larkin, J. 2012, ApJL, 755, L28, doi: 10.1088/2041-8205/755/2/L28

Sozzetti, A., & Desidera, S. 2010, A&A, 509, A103, doi: 10.1051/0004-6361/200912717

Spiegel, D. S., & Burrows, A. 2012, ApJ, 745, 174, doi: 10.1088/0004-637X/745/2/174

Spiegel, D. S., Burrows, A., & Milsom, J. A. 2011, ApJ, 727, 57, doi: 10.1088/0004-637X/727/1/57

Spina, L., Meléndez, J., Karakas, A. I., et al. 2017, Monthly Notices of the Royal Astronomical Society, 474, 2580, doi: 10.1093/mnras/stx2938

Stanford-Moore, S. A., Nielsen, E. L., De Rosa, R. J., Macintosh, B., & Czekala, I. 2020, ApJ, 898, 27, doi: 10.3847/1538-4357/ab9a35

Sudol, J. J., & Haghighipour, N. 2012, ApJ, 755, 38, doi: 10.1088/0004-637X/755/1/38

Suzuki, R., Kudo, T., Hashimoto, J., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III, ed. I. S. McLean, S. K. Ramsay, & H. Takami, 773530, doi: 10.1117/12.857361

Takeda, Y. 2007, Publications of the Astronomical Society of Japan, 59, 335, doi: 10.1093/pasj/59.2.335

Tal-Or, L., Trifonov, T., Zucker, S., Mazeh, T., & Zechmeister, M. 2019, MNRAS, 484, L8, doi: 10.1093/mnrasl/sly227

Thalmann, C., Carson, J., Janson, M., et al. 2009, ApJL, 707, L123, doi: 10.1088/0004-637X/707/2/L123

Thoul, A. A., Bahcall, J. N., & Loeb, A. 1994, ApJ, 421, 828, doi: 10.1086/173695

Torres, G., Curtis, J. L., Vanderburg, A., Kraus, A. L., & Rizzuto, A. 2018, ApJ, 866, 67, doi: 10.3847/1538-4357/aadca8

Triaud, A. H. M. J., Burgasser, A. J., Burdanov, A., et al. 2020, Nature Astronomy, 4, 650, doi: 10.1038/s41550-020-1018-2

Trifonov, T., Tal-Or, L., Zechmeister, M., et al. 2020, A&A, 636, A74, doi: 10.1051/0004-6361/201936686

Tsantaki, M., Sousa, S. G., Adibekyan, V. Z., et al. 2013, A&A, 555, A150, doi: 10.1051/0004-6361/201321103

Tull, R. G., MacQueen, P. J., Sneden, C., & Lambert, D. L. 1995, PASP, 107, 251, doi: 10.1086/133548

Tuomi, M., Jones, H. R. A., Barnes, J. R., Anglada-Escudé, G., & Jenkins, J. S. 2014, MNRAS, 441, 1545, doi: 10.1093/mnras/stu358

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science and Engineering, 13, 22, doi: 10.1109/MCSE.2011.37

van Leeuwen, F. 2007a, A&A, 474, 653, doi: 10.1051/0004-6361:20078357

—. 2007b, Hipparcos, the New Reduction of the Raw Data, Vol. 350 (Springer), doi: 10.1007/978-1-4020-6342-8

van Leeuwen, F., & Evans, D. W. 1998, Astronomy and Astrophysics Supplement Series, 130, 157, doi: 10.1051/aas:1998218

van Leeuwen, F., & Fantino, E. 2005, A&A, 439, 791, doi: 10.1051/0004-6361:20053193

van Saders, J. L., Ceillier, T., Metcalfe, T. S., et al. 2016, Nature, 529, 181, doi: 10.1038/nature16168

Vandal, T., Rameau, J., & Doyon, R. 2020, arXiv e-prints, arXiv:2009.09276. https://arxiv.org/abs/2009.09276

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: https://doi.org/10.1038/s41592-019-0686-2

Voges, W., Aschenbach, B., Boller, T., et al. 1999, A&A, 349, 389. https://arxiv.org/abs/astro-ph/9909315

—. 2000, IAUC, 7432, 3

Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, in Proc. SPIE, Vol. 2198, Instrumentation in Astronomy VIII, ed. D. L. Crawford & E. R. Craine, 362, doi: 10.1117/12.176725

Vousden, W. D., Farr, W. M., & Mandel, I. 2016, MNRAS, 455, 1919, doi: 10.1093/mnras/stv2422

Wahhaj, Z., Milli, J., Romero, C., et al. 2021, arXiv e-prints, arXiv:2101.08268. https://arxiv.org/abs/2101.08268

Wang, J., Wang, J. J., Ma, B., et al. 2020, AJ, 160, 150, doi: 10.3847/1538-3881/ababa7

Wang, J. J., Kulikauskas, M., & Blunt, S. 2021, whereistheplanet: Predicting positions of directly imaged companions. http://ascl.net/2101.003

Wang, J. J., Graham, J. R., Pueyo, L., et al. 2016, The Astronomical Journal, 152, 97, doi: 10.3847/0004-6256/152/4/97

Wang, J. J., Graham, J. R., Dawson, R., et al. 2018, AJ, 156, 192, doi: 10.3847/1538-3881/aae150

Wes McKinney. 2010, in Proceedings of the 9th Python in Science Conference, ed. Stéfan van der Walt & Jarrod Millman, 56 – 61, doi: 10.25080/Majora-92bf1922-00a

West, A. A., Hawley, S. L., Bochanski, J. J., et al. 2008, AJ, 135, 785, doi: 10.1088/0004-6256/135/3/785

Wizinowich, P., Acton, D. S., Shelton, C., et al. 2000, PASP, 112, 315, doi: 10.1086/316543

Zucker, S., & Mazeh, T. 2001, ApJ, 562, 549, doi: 10.1086/322959

Zuckerman, B., Rhee, J. H., Song, I., & Bessell, M. S. 2011, ApJ, 732, 61, doi: 10.1088/0004-637X/732/2/61

Zuckerman, B., Song, I., Bessell, M. S., & Webb, R. A. 2001, ApJL, 562, L87, doi: 10.1086/337968