

## Weighing the Stars I: The mass-luminosity relation from $0.70M_{\odot}$ to $0.075M_{\odot}$

ANDREW W. MANN,<sup>1, 2,\*</sup> TRENT DUPUY,<sup>3</sup> ADAM KRAUS,<sup>4</sup> ERIC GAIDOS,<sup>5</sup> AARON RIZZUTO,<sup>4, †</sup> MICHAEL IRELAND,<sup>6</sup> MEGAN ANSELL,<sup>7</sup> JEAN-LUC BEUZIT,<sup>8</sup> JASON DITTMANN,<sup>9</sup> CHAO-LING HUNG,<sup>10</sup> SAMUEL FACTOR,<sup>4</sup> THIERRY FORVEILLE,<sup>8</sup> MICHAEL LIU,<sup>11</sup> AND RAQUEL MARTINEZ<sup>4</sup>

<sup>1</sup> Columbia University, Department of Astronomy, 550 West 120th Street, New York, NY 10027

<sup>2</sup> Department of Physics and Astronomy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3255, USA

<sup>3</sup> Gemini Observatory, Northern Operations Center, 670 N. Aohoku Place, Hilo, HI 96720, USA

<sup>4</sup> Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA

<sup>5</sup> Department of Geology & Geophysics, University of Hawaii at Mānoa, Honolulu, Hawaii 96822 USA

<sup>6</sup> Research School of Astronomy & Astrophysics, Australian National University, Canberra ACT 2611, Australia

<sup>7</sup> Department of Astronomy, University of California at Berkeley, Berkeley, CA 94720, USA

<sup>8</sup> Observatoire de Grenoble, Institut de Planétologie et d’Astrophysique de Grenoble, BP 53, 38041 Grenoble Cedex 9, France

<sup>9</sup> Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA, 02138

<sup>10</sup> Department of Physics, Manhattan College, 4513 Manhattan College Parkway, Riverdale, NY 10471, USA

<sup>11</sup> Institute for Astronomy, University of Hawai‘i at Mānoa, Honolulu, HI 96822, USA

Submitted to The Astrophysical Journal

### ABSTRACT

The mass-luminosity relation for late-type stars has long been a critical tool for both estimating stellar masses and testing evolutionary models. New methods to measure the metallicities of M dwarfs, the availability of *Gaia* parallaxes, an increasing number of benchmark binaries near the hydrogen burning limit, and the need for higher-precision masses for M dwarfs in the Solar neighborhood all motivate revisiting existing relations. Here we present an empirical relation between  $M_{K_S}$  and  $M_*$  spanning  $0.70M_{\odot}$  down to the hydrogen burning limit ( $\simeq 0.075M_{\odot}$ ). The relation is calibrated from 61 nearby binary systems, whose orbits we determine using a combination of our own monitoring with NIRC2 on Keck, archival adaptive optics data, and astrometry from the literature. We use these data to determine total masses to better than 1% in the best cases (median precision of 6.4%). Joining these masses with resolved  $M_{K_S}$  magnitudes for binary component stars, we fit an empirical  $M_{K_S} - M_*$  relation that can be used to predict masses to 2-3% for  $0.18M_{\odot} < M_* < 0.70M_{\odot}$ , and 3-6% for  $0.075M_{\odot} < M_* < 0.18M_{\odot}$ . Our resulting relation is mostly consistent with earlier work, and deviations are explainable by differences in the input sample. We find [Fe/H] has a weak effect on the  $M_{K_S} - M_*$  for the range of metallicities expected in the Solar neighborhood ( $\simeq 1\%$  change in  $M_*$ ), but may exceed uncertainties in the relation ( $> 3\%$ ) for subdwarfs. Our results also suggest that evolutionary models likely overestimate the role of [Fe/H] on the  $M_{K_S} - M_*$  relation, but the difference is not significant with the current sample.

*Keywords:* stars: fundamental parameters; stars: low-mass; stars: luminosity function, mass function

Corresponding author: Andrew W. Mann  
[mann.andrew.w@gmail.com](mailto:mann.andrew.w@gmail.com)

\* Hubble Fellow

† 51 Peg b Fellow

## 1. INTRODUCTION

Over the past decade, M dwarfs have become critically important across a wide range of astrophysics. On the small scale, M dwarfs have become attractive targets for the identification and characterization of exoplanets. The small size, low mass, and low luminosity of late-type dwarfs facilitates the discovery of exoplanets, particularly planets that are small (e.g. Muirhead et al. 2012b; Martinez et al. 2017; Mann et al. 2018), young (e.g., Bowler et al. 2012; Mann et al. 2016), and/or in their circumstellar habitable-zone (e.g., Tarter et al. 2007; Shields et al. 2016; Dittmann et al. 2017). Close-in, rocky planets are also significantly more common around M dwarfs than their Sun-like counterparts (Dressing & Charbonneau 2013; Petigura et al. 2013; Mulders et al. 2015; Gaidos et al. 2016), and are more likely to harbor close-in multi-planet systems (e.g., Muirhead et al. 2015; Ballard & Johnson 2016).

Like planets, studies of Galactic properties are inexorably linked to the properties of their most numerous constituents (>70% of stars in the Solar neighborhood are M dwarfs, Henry et al. 1994; Reid et al. 2004). Late-type dwarfs weigh heavily on the Galactic mass function (e.g., Covey et al. 2008), and are useful probes of the Milky Way’s structure (e.g., Jurić et al. 2008; Ferguson et al. 2017), kinematics (e.g., Bochanski et al. 2007; Yi et al. 2015) and chemical evolution (Woolf & West 2012; Hejazi et al. 2015).

Late-type stars also play an important role on extragalactic and potentially cosmological scales. Although K and M dwarfs are much fainter than their higher-mass counterparts, they are numerous enough to impact the integrated light of many massive galaxies, making their properties an essential component of the initial mass function (e.g., Conroy & van Dokkum 2012; McConnell et al. 2016) and mass-to-light ratio (Spinelli et al. 2015) in nearby galaxies. Recent non-detection of Type Ia supernovae progenitors have shifted the field from single-degenerate origins (evolved star and a white dwarf) towards a double-degenerate model (two white dwarfs), however, single degenerate models with an M dwarf and a white dwarf are still plausible (Wheeler 2012).

For all these areas of research, it is essential that we have a method to estimate the masses of late type dwarfs, as well as a means to test methods on which M dwarf parameters are determined. For stars lacking a precise distance, most studies depend on evolutionary models to derive masses (e.g., Muirhead et al. 2012a; Conroy & van Dokkum 2012). However, differences between empirical and model-predicted mass-radius and radius-luminosity relations for late-type stars (e.g., Boyajian

et al. 2012; Feiden & Chaboyer 2012) raise concerns about the reliability of model-based masses. Further, the masses derived depend on both the model grid used (Spada et al. 2013; Choi et al. 2016), and the observed parameter over which the interpolation is done (e.g., color versus luminosity, Mann et al. 2012, 2015). Ultimately, these models need to be tested, and if necessary, refined.

In a case where the target M dwarf has a precise distance, it is common to use a relation between mass and luminosity (e.g., Henry & McCarthy 1993; Delfosse et al. 2000). These relations are usually calibrated empirically using binaries, which have masses determined from long-term astrometric and/or spectroscopic monitoring of their orbits. Combined with a parallax and component magnitudes for each system, it is then possible to fit for a relation between absolute magnitude (a proxy for total luminosity) and stellar mass.

The empirical mass-luminosity relations of Henry & McCarthy (1993) and Delfosse et al. (2000) provide mass determinations with  $\simeq 10\%$  precision, with more recent improvements by Benedict et al. (2016). As fields that rely on M dwarf parameters have pushed to higher precision, there has been need for proportionately improved precision on M dwarf masses. With the arrival of *Gaia* parallaxes late-type dwarfs even beyond the Solar neighborhood will have  $< 1\%$  parallaxes, making scatter in the  $M_{K_S} - M_*$  relation the primary source of error. In addition to the need for better precision, all relations have gaps in their calibration sample, particularly below  $0.1M_\odot$ . Further, as these studies push to samples that are not well represented in the Solar neighborhood, it has become critical to understand systematics in the relation (e.g., metallicity, age, and activity). As methods to measure precise spectroscopic metallicities of M dwarfs have only recently become possible recently (e.g., Rojas-Ayala et al. 2010; Neves et al. 2014), previous studies were unable to empirically study the role metallicity on the mass-luminosity relation.

Here we present a revised empirical relation between  $M_*$  and  $M_{K_S}$  spanning almost an order of magnitude, from  $0.075M_\odot$  to  $0.70M_\odot$ . The relation is built on orbital fits to visual binaries from a combination of adaptive optics imaging and astrometric measurements in the literature. In Section 2 we detail our selection of nearby late-type binaries with orbits amenable to mass determinations. We overview our astrometric and spectroscopic observations in Section 3, including those from data archives and the literature. Our orbit-fitting procedure is explained in Section 4. We describe our method to determine other parameters of each system ([Fe/H], distance, and  $M_{K_S}$ ) in Section 5. Our method and results from fitting the  $M_{K_S} - M_*$  relation is described in Sec-

tion 6, including an analysis of the errors as a function of  $M_{K_S}$  and a detailed look at the effects of [Fe/H]. We compare our results to that of earlier relations in Section 6.4. We conclude in Section 7 with a brief summary, and a discussion of the important caveats and complications to consider when using our relation, as well as future directions we are taking to expand on the work given here.

*If you want to use the relations in this manuscript, we advise at least reading Section 7.3 to understand the potential limitations of our provided program and posteriors.*

## 2. TARGET SELECTION

Our initial list of late-type binaries was built by cross-matching catalogs of nearby M dwarfs (Lépine et al. 2013; Gaidos et al. 2014; Dittmann et al. 2014; Winters et al. 2015), with the fourth catalog of interferometric measurements of binary stars (INT4, Hartkopf et al. 2001), and adaptive optics (AO) images from the Keck Observatory Archive (KOA). As part of this cross-match, we also included targets matching the M dwarf selection criteria of Gaidos et al. (2014), but with a bluer ( $V - J > 2$ ) color cut to incorporate additional late-K dwarfs. We kept any binaries with separations less than 5''. We then added in other known late-type binaries from Law et al. (2008), Janson et al. (2012), Janson et al. (2014), and Ward-Duong et al. (2015). This provided a list of more than 300 systems.

From here we selected those for which we expect to be able to obtain a precise orbit on a reasonable (few year) timescale. We assumed the average of available separation measurements represented the typical semi-major axis. We then identified systems for which we would have at least 30% of the orbit within the two years of our orbital monitoring program with Keck. This cut accounted for existing data. As a result, long-period binaries with extensive previous observations were included, depending on the baseline available, while those with only a single epoch would generally need to have orbits of  $\ll 10$  years to be observed. After this cut we were left with 125 systems. With the exception of 38 systems in the far South ( $\delta < -30^\circ$ , outside the range of Keck), all of these were included in our initial observing queue.

We removed 21 systems from our analysis because of an unresolved tertiary (or quaternary) component noted in the literature (e.g., Law et al. 2010; Tokovinin & Smekhov 2002; Tokovinin 2018). In their current form, such systems are not useful for our analysis, as we have no  $\Delta K$  magnitudes or mass ratios for the unresolved components. Since many of these are double- or triple-

lined systems, it is possible to recover their parameters with multi-epoch radial velocities, and some systems already have the necessary data in the literature (e.g., Ségransan et al. 2000). We are actively monitoring these systems with high-resolution near-infrared spectrographs (Yuk et al. 2010; Park et al. 2014; Rayner et al. 2012), but they are excluded from the analysis done here. Triples where the third component is more distant (and well resolved) were retained.

A total of 18 systems were flagged as young, all of which were tagged to nearby young moving groups or clusters (Shkolnik et al. 2012; Kraus et al. 2014; Gagné et al. 2014; Malo et al. 2014; Gagné et al. 2015; Riedel et al. 2017; Shkolnik et al. 2017; Rizzuto et al. 2017; Lee & Song 2018). We continued to monitor these targets even after flagging them as young, but they were not included in the analysis for this paper. Many of these are expected to be pre-main-sequence stars and hence will not follow the same mass-luminosity relation. Even for those that are on the main-sequence, such young M dwarfs may have higher luminosities for a given mass (e.g., Kraus et al. 2015; Stassun et al. 2012; Somers & Stassun 2017). These young systems will be the subject of a future study into the effects of age and activity on the parameters of M dwarfs.

After the completion of our observing program, we removed targets for which we had  $< 30\%$  of the orbit or fewer than six unique measurement epochs. We attempted to fit orbits of the remaining 70 systems (Section 4). Seven of the orbits fits failed to converge, mostly because their semi-major axes were much larger than expected (e.g., eccentric systems caught near periastron). An additional 14 were removed because the derived mass for the system was too imprecise (errors  $> 25\%$ ) to be useful. Most of these 14 were systems with a poor or no parallax. This left us with 52 binaries (104 stars).

We added in nine targets with orbits from Dupuy & Liu (2017), primarily to fill in the sample around the end of the M dwarf sequence. These nine were selected because they are massive enough to sustain hydrogen fusion and satisfy all our other selection criteria. Systems from Dupuy & Liu (2017) also had their orbits fit using a nearly identical method to our own, often using similar or identical data sources and observational methods (primarily Keck/NIRC2).

Basic parameters of the final 61 systems included in our analysis are given in Table 3.

## 3. OBSERVATIONS AND DATA REDUCTION

### 3.1. Adaptive Optics Imaging

We analyzed a mix of AO data from our own program with NIRC2/Keck, and images from the archives of Keck,

the Canada France Hawaii Telescope (CFHT), the Very Large Telescope, and Gemini observatory. In general we analyzed all usable (e.g., non saturated, components resolved) regardless of observing mode and filter. For our analysis, we considered a single ‘epoch’ all data taken in a single night for a given target and filter.

We separate the observations and reduction by instrument/telescope below. The full list of astrometry and contrast measurements is given in Table ??, sorted by target.

### 3.1.1. Keck II/NIRC2 Imaging and Masking

Between March 2002 and November 2017, 48 of our targets were observed for a total of 166 epochs, all using the facility AO imager, NIRC2, on Keck II atop Maunakea. The majority of these data ( $> 100$  observations) were taken as part of our long-term monitoring program from June 2015 to November 2017. The rest of these observations were downloaded from the Keck Observatory Archive (KOA). For our observations, we always used the vertical angle mode and the narrow camera. Depending on brightness and conditions, imaging was usually taken with either the  $K'$  or  $K_{\text{cont}}$  filters, and non-redundant aperture masking (NRM) was always taken using the 9-hole mask and  $K'$  filter. After acquiring the target and allowing the AO loops to close, we took four to 10 images or 6-8 interferograms, adjusting coadds and integration time based on the brightness of the target.

NIRC2 data taken from the archive comprised a wide range of observing modes, filters, and cameras, although the majority was taken with the narrow camera and one of the  $H$ - or  $K$ -band filters. We included nearly all data with a clear detection of both binary components independent of the observing setup. Archive data was taken between March 2002 up to November 2015. We discarded saturated images, those taken with the coronagraph atop one or both binary components, and those with extremely poor AO correction.

The same basic data reduction was applied to both observations from our own program and those from the archive, following the procedure from Kraus et al. (2016). To briefly summarize, we linearized each frame then dark and flat corrected using calibration data taken the same night. In cases where no appropriate darks or flats were taken in the same night, we use a set from the nearest available night. We interpolated over “dead” and “hot” pixels, which were identified from superflats and superdark built from data spanning 2006 to 2016. Because flats are rarely taken in narrowband filters, we used superflats built from the nearest (in wavelength) broadband filter where appropriate. Pixels with flux levels  $> 10\sigma$  above the median of the eight adjacent pixels

(primarily cosmic rays) were replaced with the median. Images were visually inspected as part of identifying the binary location, and a handful ( $< 1\%$ ) of images were negatively impacted by our cosmic ray removal (e.g., removal of part of the source). These images were manually corrected.

### 3.1.2. CFHT/PUEO Imaging

In total, 31 of our targets were observed with the 3.6m Canada-France-Hawaii Telescope (CFHT) using the Adaptive Optics Bonnette (AOB, often referred to as PUEO after the Hawaiian owl, Arsenault et al. 1994) with the KIR infrared camera (Doyon et al. 1998). After removing bad observations (e.g., saturated, unresolved, or poor AO correction), 187 unique epochs were used, all of which we downloaded from the CFHT archive at the Canadian Astronomy Data Centre<sup>1</sup>. Observations spanned December 1997 to January 2007, covering most of the time PUEO was in use at CFHT (1996 to 2011). Most observations were taken in a few year period from 2000 to 2003. Observations spanned a range of filters across  $JHK$  bands, but the majority used either the narrowband  $[\text{Br}\gamma]$  or  $[\text{FeII}]$  filters. All observations used a 3-5 point dither pattern and took at least two images at each dither location.

Data reduction for PUEO data followed the same basic steps as our NIRC2 data. We first applied flat-fielding and dark correction using a set of superflats and superdark built by splitting the datasets into 6 month blocks and combining calibration data within the block. We identified bad pixels by making large median stacks of science data taken across dozens of objects, which were then replaced with the median of nearby pixels. To identify cosmic rays, we median stacked consecutive images of each target (at a fixed location), then compared the stack to individual images. Data points  $> 10\sigma$  above the robust mean of the stacks are replaced with the median of the eight surrounding points. Because PUEO/KIR data were taken in sets of  $> 5$  images before the object was dithered, this median stacking was effective for removing nearly all cosmic rays.

### 3.1.3. VLT/NaCo Imaging

We downloaded AO-assisted images from the ESO archive taken with the Nasmyth Adaptive Optics System Near-Infrared Imager and Spectrograph (NAOS-CONICA, or NaCo) instrument on VLT<sup>2</sup>. Data spanned November 2001 to November 2017, with about half the epochs taken from 2001 to 2005. Data include 42 unique

<sup>1</sup> CADC archive

<sup>2</sup> ESO Archive

epochs over 19 of our targets, excluding saturated or otherwise unusable data. Observations were taken with a wide range of filters, cameras, and observing patterns, but the majority were taken using the S13 camera with either broadband  $K_s$  and  $L$ , or narrowband [FeII] and [Br $\gamma$ ] filters, and always following a 2-4 point dither pattern.

Basic data reduction was applied to NaCo data following a similar procedure as with the PUEO and NIRC2 data. We applied flat-fielding and dark corrections to each observation using the standard set of calibrations taken each night as part of the VLT queue. In the case where calibration (dark or flat) data was missing or unusable, we used the nearest (in time) set of calibration images matching the filter (for flats) and exposure setup (for darks). Flats taken in broadband filters were used for flat-fielding narrow-band data at similar wavelengths. We built bad pixel masks using median stacks of all images taken within a night after applying flat and dark corrections. To identify and remove cosmic rays we used the L.A. Cosmic software (van Dokkum 2001).

### 3.1.4. Gemini/NIRI Imaging

We retrieved 14 epochs for 7 of our targets from the Gemini archive, all taken with the AO imager NIRI (Hodapp et al. 2003) on Gemini North. All observations were taken between August 2008 and February 2011 and were taken with the assistance of the ALTtitude conjugate Adaptive optics for the InfraRed (ALTAIR). Most observations were taken with the f32 camera using broadband  $J$ -,  $H$ -, or  $K$ -band filters. All observations followed a 2-4 point dither pattern and took at least two images at each dither location.

Data from NIRI were reduced using the same basic tools as for all other adaptive optics data. First, we applied flat and dark corrections to each set of images using the standard calibration data taken as part of the Gemini queue, usually within 24h of the target observations. In most cases, flats taken in broadband filters were used for narrowband flat-fielding. We then identified bad pixels from median stacking all images within a given night. Observations with a target near or on top of a heavily impacted pixel were discarded (identified with the mask). We used the L.A. Cosmic software for the identification and removal of cosmic rays.

## 3.2. Astrometry and photometry

Extracting separations and position angle measurements followed a similar multi-step procedure across all instruments, excluding the NRM data (which is described below). Our method is based largely on that described in Dupuy et al. (2016) and Dupuy & Liu (2017), which is built on the techniques from Liu et al. (2008) and Dupuy

et al. (2010). Users can consult these papers for more details, but we describe most of the process here.

We first cross-correlated each image with a model Gaussian PSF to identify the most significant peaks. The cross-correlation peak occasionally centers on instrumental artifacts, often struggles to separate out partially overlapping binaries, and can easily identify the wrong source for triple systems. So this step was checked by eye and manually updated as needed. The eye-check phase also allowed us to manually remove data of poor quality: e.g., no or poor AO correction, saturated data, or an unresolved system. We used these centers as the initial guess for the position in the next phase.

We then fit for the center of the PSFs on each image (separately) by either: 1) fitting the PSF with two triple 2D Gaussians using the least-squared code MPFIT (Markwardt 2009), or 2) running *Starfinder*, as described in Diolaiti et al. (2000). Although the results of these two methods generally agreed, *Starfinder* was preferred, as it used a more realistic model of the PSF and worked even with mediocre AO correction provided the component PSFs were well-separated. *Starfinder*, however, struggled with the tightest binaries, as it failed to build a realistic PSF. So the Gaussian fit was used for any case where *Starfinder* failed. A subset of the data (10 targets) were analyzed using both methods to confirm consistency between the techniques. Both methods included an estimate for the flux of each component as part of the fit, from which we determined the contrast ratio.

The measured pixel positions were converted into separations and position angles based on the distortion correction (for NIRC2), measured pixel scale, and relevant header information. For the NIRC2 narrow camera, we used the Yelda et al. (2010) distortion solution for data taken before April 2015, and Service et al. (2016) for data taken after 2015. This also included an estimate of the pixel scale and orientation of  $9.952 \pm 0.002$  mas pixel $^{-1}$  and  $0.252 \pm 0.009^\circ$  for the former, and  $9.952 \pm 0.004$  mas pixel $^{-1}$  and  $0.262 \pm 0.020^\circ$  for the latter. For the NIRC2 wide camera we used the solution from Fu et al. (2012, priv. comm.)<sup>3</sup>, with a pixel scale of  $39.686 \pm 0.008$  mas pixel $^{-1}$  and the same orientation as the narrow camera.

Distortion solutions were not available for the other instruments, but data was always taken following a dither pattern to sample the CCD distortion pattern. For KIR, we adopted a pixel scale of  $34.8 \pm 0.1$  mas pixel $^{-1}$  (Stapelfeldt et al. 2003) and an orientation of

<sup>3</sup> <http://homepage.physics.uiowa.edu/~haifu/idl/nirc2wide/>

$0\pm2^{\circ}$ <sup>4</sup>. For NaCo, we assumed a pixel scale of  $13.24\pm0.05$  mas pixel $^{-1}$  for the S13 camera (Masciadri et al. 2003; Neuhauser et al. 2005) and the values given in the ESO documentation for all others<sup>5</sup> (with the same error). The rotation taken from the NaCo headers was assumed to be correct to  $0.4^{\circ}$  (Seifahrt et al. 2008). For NIRI observations we applied a distortion solution and used a pixel scale provided by Gemini<sup>6</sup> for each camera, with a global uncertainty of 0.05 mas pixel $^{-1}$  on the pixel scale and  $0.7^{\circ}$  on the orientation. The NIRI orientation error was determined using observations of the same target between observing runs, but over which no orbital motion should have been visible.

Extraction of separations and position angles from of masking observations followed the appendix of Kraus et al. (2008) with the aid of the ‘Sydney’ aperture-masking interferometry code<sup>7</sup>. To remove systematics, the observation of each target observed with NRM was paired with a calibration observation of another nearby star taken in the same night with a similar setup and observational parameters. Binary system profiles were then fit to the closure phase to produce detection limits and uncertainties. More details on the analysis of masking data can be found in Lloyd et al. (2006), Kraus et al. (2008) and Evans et al. (2012).

Data from a single night and with the same filter and target were combined into a single measurement, with measurement errors estimated using the scatter in the individual images within a night. The scatter across images was combined with the uncertainty in the rotation and pixel scale (and/or error from the distortion solution) in quadrature. This treated image scale and rotation uncertainty as correlated within a night and filter (so they do not decrease with  $\sqrt{N}$ ), and uncorrelated with measurement errors.

We applied an additional correction for differential atmospheric refraction (Lu et al. 2010) using filter wavelength information and weather data from the header (for VLT) or from the CFHT archive<sup>8</sup> (for Keck, CFHT, and Gemini).

### 3.3. Literature Astrometry

We adopted astrometric measurements (position angles and separations) from the literature for 50 of the 52 targets analyzed here. To help identify sources of astrometry, we used the fourth catalog of interferomet-

ric observations of binary stars (INT4, Hartkopf et al. 2001). We only used measurements that included both a separation and position angle. In cases of literature data also available in one of the archives above, we adopted our own measurements over the literature data to ensure more homogenous measurements and errors.

One complication using the broad literature astrometry was the issue of inhomogeneous reporting of separation and position angle errors. Many references provided measurements without errors, and many more reported errors only based on the PSF measurement. These often did not account for more complicated terms, e.g., pixel scale, offsets in the position angle of the detector on the sky, or time-dependent changes in the optics and detector.

We mitigated this problem by generating new errors common to a given data reference. For each literature source, we identified a set of binaries where an orbit can be fit (with  $<5\%$  errors on the angular separation) without data from that particular reference. Some literature sources were merged for this comparison, provided they used the same instrument and basic analysis. We included binaries not specifically analyzed in this paper. We fit the orbit following the method outlined in Section 4, using only the least-squares method for efficiency. We then compared the expected position angle and separation to the measurements from the reference in question. From this difference we computed the required missing error term, which, when added in quadrature with the reported errors (zero for those without errors) would yield a reduced  $\chi^2$  ( $\chi_{\nu}^2$ ) of 1. This process was repeated twice, each time updating the errors where appropriate. The repeats are necessary as many orbit fits change as the adopted errors are adjusted. References where we could not test the reported errors (or derive a new term) and those with extremely large required error terms were not used.

All literature astrometry used in this paper is listed alongside our own measurements in Table ??.

### 3.4. Near-infrared Spectra with IRTF/SpeX

To estimate the metallicities of our targets we obtained near-infrared spectra for 60 of 61 targets using the SpeX spectrograph (Rayner et al. 2003) on the NASA Infrared Telescope Facility (IRTF) atop Maunakea. Observations were taken between May 2011 and November 2017. Most data were taken as part of programs to characterize the fundamental properties of nearby M dwarfs (e.g., Mann et al. 2013b; Gaidos et al. 2014; Terrien et al. 2015). All spectra were taken in SXD mode, providing simultaneous wavelength coverage from 0.9 to  $2.5\mu\text{m}$ . For 57 of the targets, observations were taking using the  $0.3\times15''$  slit,

<sup>4</sup> KIR specifications

<sup>5</sup> NaCo Documentation

<sup>6</sup> NIRI documentation

<sup>7</sup> <https://github.com/mikeireland/idlnrm>

<sup>8</sup> CFHT weather archive

which yielded a resolution of  $R \simeq 2000$ . Spectra for three targets were taken from Dupuy & Liu (2012), Dupuy et al. (2009), or Dupuy & Liu (2017), using the  $0.5 \times 15''$  or  $0.8 \times 15''$  slits, which gave spectral resolutions of  $\simeq 750$ -1500.

For Gl 65 and HD 239960 the SpeX slit was aligned to get spectra of both targets simultaneously. For all other targets the binary was unresolved or too poorly resolved to separate in data reduction, and instead the slit was aligned with the parallactic angle. Each target was nodded between two positions along the slit to help remove high background sky levels. Depending on the target brightness and conditions, between 6 and 30 individual exposures were taken following this nodding pattern, with exposure times varying from 8s to 180s. An A0V-type star was observed immediately before or after each target to help with the removal of telluric lines and flux calibrate the spectrum. The final stacked spectra had S/N of  $> 100$  per resolving element in the  $K$ -band for all but the four faintest targets (which had  $S/N > 50$ ).

Basic data reduction was performed with *SpeXTool* package (Cushing et al. 2004). This included flat-fielding, sky subtraction, extraction of the one-dimensional spectrum, wavelength calibration, stacking of individual exposures, and merging of separate orders. Telluric lines were removed and the spectrum was flux calibrated using the A0V star observations and the *xteLLcor* software package (Vacca et al. 2003). When possible, the same A0V star was used for multiple targets taken near each other in time.

The only target missing a SpeX spectrum was Gl 54, which was too far south to be observed with IRTF.

#### 4. ORBIT FITTING

We fit the astrometry with the Monte Carlo Markov Chain (MCMC) software *emcee* (Foreman-Mackey et al. 2013), a python implementation of the affine-invariant ensemble sampler (Goodman & Weare 2010). For each system, we fit for seven orbital elements: the orbital period ( $P$ ), combined angular semi-major axis ( $a''$ ), eccentricity ( $e$ ), inclination ( $i$ ), epoch of periastron passage ( $T_0$ ), argument of periastron ( $\omega$ ), and position angle of the line of nodes ( $\Omega$ ). Parameters were limited by physical constraints, e.g.,  $P > 0$ ,  $0 < e < 1$ , and  $0 < i < \pi$ , but were given no additional boundaries.

We applied priors of  $1/P$ ,  $1/a''$ , and  $\sin(i)$  to  $P$ ,  $a''$ , and  $i$ , respectively. All other parameters evolved under uniform priors. Walkers were initialized with the best-fit orbit determined using *MPFIT* (Markwardt 2009) with a spread in starting values based on the *MPFIT* estimated errors. Each MCMC chain was initially run for  $10^5$  steps

with 100 walkers. We tested to see if a chain converged using the autocorrelation time (Goodman & Weare 2010). In about 1/3 of the systems the fit did not converge after the initial  $10^5$  steps, so we ran these for  $10^6$  total steps. Those that did not converge after  $10^6$  steps were not included in our final sample (these are not counted in the 61 system count, see Section 2), as these were also the least-well characterized systems. We saved every 100 steps in the chain, and the first 5% of each chain was removed for burn-in.

Systems of near-equal mass may have the primary and companion misidentified, both in our own measurements and also those taken from the literature. We identified such points by eye during the *MPFIT* stage and manually adjusted the position angles before starting the MCMC run. In total  $\simeq 11$  epochs were corrected this way, almost all of which were for three systems (GJ 3421, Gl 831, and Gl 913). GJ 3421 in particular, has  $\Delta K = 0.001 \pm 0.018$ , and required adjustment for nearly half of the observations. A more robust solution to this problem would be to feed a double-peaked posterior at the reported value and  $\pm 180^\circ$  into the likelihood function. However, in all cases the problematic points were obvious by eye, had reported  $\Delta$ mags consistent with zero, and a simple  $180^\circ$  completely fixed the orbit. So there was no advantage to using a more sophisticated analysis.

We show some example orbital fits in Figure 1 (with the full set in the Appendix B) and provide the best-fit orbital parameters in Table 4. Orbits span a wide range in orbital period; the tightest binaries have  $P < 1$  yr, while the widest  $P > 50$  yr. The two systems with  $P > 50$  yr (Gl 301 and Gl 277) were also some of the least-well characterized.

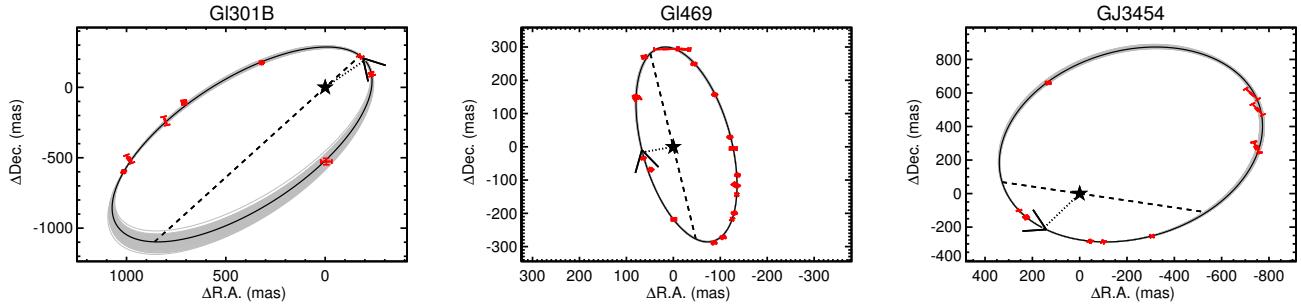
Resulting fits all have  $0.1 < \chi_\nu^2 < 2$ , within the expected distribution given the number measurements per orbit. This was in part a consequence of our method of assigning errors to literature data (Section 3.3). However, we achieved reasonable  $\chi_\nu^2$  values even when all or nearly all the data were from our own measurements, indicating that our astrometry measurements have reasonable assigned errors.

The nine systems from Dupuy & Liu (2017) are labeled separately in Table 3.

#### 5. STELLAR PARAMETERS

##### 5.1. Parallaxes

Parallaxes for 54 of the targets were drawn from the literature. In general we adopted the most precise parallax available, although preference is given to parallaxes that accounted for the fact that the target is a binary and those with a baseline significantly different from the binary's orbital period. For  $\simeq 1/3$  of the sample, parallaxes



**Figure 1.** Example results of our orbit fitting procedure for Gl301AB (left), Gl469AB (middle), and GJ3454AB (right). Red points are individual measurements of the separation and position angle. Black solid line shows the best-fit (highest likelihood). Grey lines are drawn by randomly sampling 100 orbit fits from the MCMC fit to give an estimate of the errors. The dotted line marks the periastron passage, with an arrow for the orbital direction, and the dashed line indicates the line of nodes. For Gl301AB, there are no data for a large section of the orbit, but the final fit is relatively well constrained. For Gl469AB, high-quality astrometry is available for the full orbit, and the resulting errors on orbital parameters are so small that the grey lines cannot be seen. All other orbits are shown in Appendix A

were drawn from the new reduction of the Hipparcos data (van Leeuwen 2007). A few systems are in hierarchical triples with a wider companion that also has an independent parallax measurement, which we used where possible. We drew parallaxes from Dupuy & Liu (2017) along with their orbital solutions, and parallaxes from Benedict et al. (2016) for those overlapping with our own sample. Most of the remaining parallaxes were drawn from the general catalogue of trigonometric parallaxes (van Altena et al. 1995).

Many of our targets do not have entries in the second data release of *Gaia* (DR2, Lindegren et al. 2018; Gaia Collaboration et al. 2018). They were most likely excluded because centroid shifts due to orbital motion prevented a single-star solution (a requirement to be included in DR2). While many targets do have entries, we found significant differences between the *Gaia* values and earlier measurements (including from *Gaia* DR1, Gaia Collaboration et al. 2016). This is most evident when we examine a few wide triples in *Gaia* DR2 (where the third star is several arc seconds away, and hence easily resolved); many such systems have significantly different parallaxes reported for the unresolved pair than for the wider companion. This strongly indicates that orbital motion on our targets is impacting the final parallax values or uncertainties even when a value is reported. We found no such issues with our other parallax sources, although some of this issue may be hidden by larger uncertainties.

We ultimately used *Gaia* DR2 parallaxes for only two systems, GJ1245AC and GJ277AC, where we adopted the parallax of the wider (resolved) B components. In both cases the B components are not known to harbor another unresolved star, and hence they should not be impacted by the same binarity issue. GJ 570BC, Gl

695BC, GJ 2005BC, and Gl 22AC also have wider companions. However, Gl 570A is extremely bright ( $G \simeq 5$ ) and hence may have an unreliable parallax, GJ 2005A has no entries in *Gaia* DR2, and the available parallaxes (from *HST* and Hipparcos) for Gl 695A,BC and Gl 22AC,B are more precise than the ones for Gl 695A and Gl 22B from *Gaia* DR2.

For seven systems we derived new parallaxes using MEarth astrometry (Nutzman & Charbonneau 2008). Updated parallaxes were measured following the procedure from Dittmann et al. (2014). The only difference was we used a dataset with a baseline  $\simeq$ two years longer, which helps average out systematic errors arising from centroid motion due to the binary orbit, and significantly reduces the overall uncertainties.

We list all parallaxes and references in Table 3.

## 5.2. [Fe/H]

We estimated [Fe/H] using our SpeX spectra and the empirical relations from Mann et al. (2013a) for K5-M6 dwarfs, and Mann et al. (2014) for M6-M9 dwarfs. For the one target lacking a SpeX spectrum (Gl 54), we derive [Fe/H] using the optical calibration of Mann et al. (2013a) and a moderate-resolution optical spectrum taken from Gaidos et al. (2014). These relations use a series of atomic lines (primarily Na, Ca, and K features) in the optical or NIR (e.g., Rojas-Ayala et al. 2010; Terrien et al. 2012), empirically calibrated using wide binaries containing a Sun-like primary and an M-dwarf companion (e.g., Bonfils et al. 2005; Johnson & Apps 2009; Neves et al. 2012). The calibrations were based on the assumption that components of such binaries have similar or identical metallicities (e.g., Teske et al. 2015). Similar methods have been used extensively to assign metallicities across the M dwarf sequence (e.g., Terrien

et al. 2015; Muirhead et al. 2015; Dressing et al. 2017; Van Grootel et al. 2018; Mace et al. 2018). Final adopted [Fe/H] values are given in Table 3.

Uncertainties on [Fe/H] from such spectroscopic determinations are generally 0.08-0.12 dex. This exceeds the measurement uncertainties by a factor of 1.5-3 (Newton et al. 2014). The difference may be due to inhomogeneities in the assigned primary stars' metallicities (e.g., Brewer et al. 2015, 2016; Hinkel et al. 2016), as well as complications in using atomic lines as a proxy, which are not expected to scale perfectly with iron abundances (Rojas-Ayala et al. 2012; Veyette et al. 2016). The existing data suggest that relative abundances (one M dwarf relative to another) are far more precise than the overall scale (relative the Sun).

For all but two targets (Gl 65 and HD 239960), our NIR spectra are for the combined flux of the binary components. Mann et al. (2014) explored the issue of measuring metallicities of binaries with unresolved data by combining spectra of single-stars with similar metallicities and re-applying the same calibration. The bias introduced is negligible ( $\lesssim 0.02$  dex) when compared to overall uncertainties. The added scatter is smaller than the measurement uncertainties, and can be explained entirely by the Poisson noise introduced in the addition of component spectra.

Three targets (LP415-20, 2M2140+16, and 2M2206-20) have SpeX spectra taken with a wider slit, yielding lower spectral resolution. The bands in Mann et al. (2014) are defined using a homogenous dataset taken with the narrow ( $0.3''$ ) slit, so this difference may impact the derived [Fe/H]. We tested this by convolving a set of single-star SpeX spectra taken with the  $0.3''$  slit with a Gaussian to put them at the appropriate lower resolution. The median of the derived [Fe/H] values changed by  $< 0.01$  dex, but the change varies between targets. Based on the resulting scatter, we estimate the errors on [Fe/H] from the lower-resolution spectra to be 0.12 dex on a Solar scale and 0.08 dex on a relative scale. These targets are marked separately in Table 3.

Two of the targets in our sample are L dwarfs (2M0746+20 and 2M1017+13). These both land above hydrogen burning limit, and hence are included in our analysis. However, the Mann et al. (2014) method contains no L dwarf calibrators. We report our derived [Fe/H] values for these systems, but advise treating the assigned values with skepticism until an L dwarf calibration becomes available.

Two targets (HIP 95995 and HIP 96656) are too warm (earlier than K5) for the calibration of Mann et al. (2013a). For these we take [Fe/H] values from Casagrande et al. (2011) and Torres et al. (2010), respectively. These

[Fe/H] measurements are not necessarily on the same scale as those from Mann et al. (2013a), which are calibrated against abundances of Sun-like stars from Brewer et al. (2015, 2016). Given reported variations in [Fe/H] for these stars, as well as [Fe/H] determination differences (Hinkel et al. 2014, 2016) we adopt conservative 0.08 dex uncertainties on both systems.

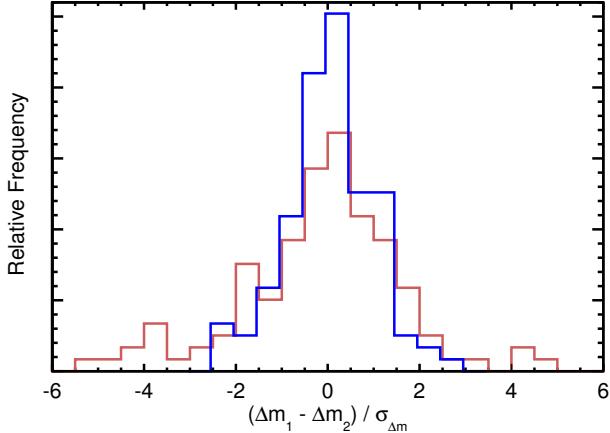
### 5.3. $K_S$ -magnitudes

We initially adopted unresolved  $K_S$  magnitudes from the Two Micron All Sky Survey (2MASS, Skrutskie et al. 2006). For the brightest stars in our sample, some 2MASS magnitudes were saturated, and others may have systematically underestimated errors as they are near saturation. For these targets we recalculated  $K_S$  magnitudes using available optical and NIR spectra, following the method of Mann & von Braun (2015) and Mann et al. (2015). Optical spectra were taken from Gaidos et al. (2014). Synthetic magnitudes were broadly consistent with 2MASS  $K_S$  magnitudes (and at similar precision) for fainter targets ( $K_S > 7$ ). We only updated  $K_S$  magnitudes for bright systems where our synthetic photometry differed from the 2MASS value by more than  $2\sigma$  (five systems). We mark these systems in Table 3.

To compute  $\Delta K_S$ , we calculated the robust weighted mean of all contrast ratios derived from our AO data (Section 3.2) taken with filters centered somewhere within the  $K$  band. Errors on contrasts for a given epoch were based on the scatter in flux measurements between consecutive measurements. This may have underestimated errors arising from imperfect flat-fielding, uncorrected non-linearities in the detector, or wavelength-dependent flux losses from the optics or atmosphere, as well as intrinsic variability of the star. We compared  $\Delta m$  measurements of the same star using the same filter and instrument but on different nights (Figure 2). The comparison suggested a missing error term of 0.017 magnitudes. There were not enough data to test this for all filters and instruments, but we get a consistent value across all filters with sufficient data for a test, suggesting the problem is relatively independent of the filter used. We included this term as an additional error term in our computation of  $\Delta K_S$ .

While all targets considered here had at least one measurement in a filter centered with the  $K$  band, none of the filters used were an exact match for 2MASS  $K_S$ . Much of the data were taken with narrow-band  $K_{\text{continuum}}$  or  $B\gamma$  filters, for example. We transformed these into 2MASS magnitudes before computing final  $\Delta K_S$ . These corrections were usually small ( $\lesssim 0.1$  magnitudes), and details on their derivation are given in Appendix A.

## 6. THE MASS-LUMINOSITY RELATION



**Figure 2.** Distribution of contrast ratio differences (in units of standard deviations) for data taken on the same targets but in different nights. The red line is before adding a missing error term of 0.017 magnitudes, while the blue line shows the distribution after adding this. The histograms are offset slightly for clarity, although identical bins are used as input.

### 6.1. Methodology

For main-sequence stars, the mass-luminosity relation traditionally takes the form:

$$\frac{L_*}{L_\odot} = C \left( \frac{M_*}{M_\odot} \right)^\alpha, \quad (1)$$

where  $\alpha$  depends on the dominant energy transport mechanism (e.g., radiative versus convective) and internal structure of the star (Hansen et al. 2004). This relation can depend significantly on chemical abundance and age, which produces a large scatter in the empirical mass-luminosity relation (e.g., Torres et al. 2010). For M dwarfs on the main sequence, we can effectively ignore age, as they change negligibly on Gyr timescales (Baraffe et al. 2015; Choi et al. 2016). We discuss issues of metallicity further in Section 6.3.

It is convenient to write this in terms of  $M_{K_S}$  instead of  $L_*$ . Absolute magnitudes are more easily measured than overall luminosity, and this avoids introducing errors from uncertain bolometric corrections or the need to take flux-calibrated spectra in order to measure the bolometric flux directly. Switching to  $M_{K_S}$  also mitigates effects of metallicity, as the the  $K$ -band is heavily dominated by metal-insensitive CO and H<sub>2</sub>O absorption bands. This has been seen both in stellar spectra (Rojas-Ayala et al. 2012) and model atmospheres (Figure 3). The lower metallicity sensitivity of  $K$ -band is also a major reason  $M_{K_S}$  yields a tighter relation with  $M_*$  and  $R_*$  when compared to similar relations using an optical magnitudes

(e.g., Henry & McCarthy 1993; Delfosse et al. 2000; Mann et al. 2015; Benedict et al. 2016).

Since absolute magnitudes are proportional to the log of the luminosity, we put Equation 1 in terms of  $M_{K_S}$  and  $\log(M_*/M_\odot)$ . Our sample encompasses almost an order of magnitude in mass and hence a range of underlying stellar physics. No single  $\alpha$  value is expected to fit over the full sequence. Instead, we assumed that  $\alpha$  depends on  $M_{K_S}$ , which we approximated as a polynomial of the form:

$$\log \left( \frac{M_*}{M_\odot} \right) = \sum_{i=0}^n a_i (M_{K_S} - zp)^i, \quad (2)$$

where  $a_i$  are the  $n$  fit coefficients. The constant  $zp$  is a zero-point (or anchor) magnitude, which we fixed at 7.5. This approximately corresponded to the logarithmic central mass of stars in our sample. Shifting  $zp$  did not change the overall fit, but did make the shape more sensitive to small changes in the  $a_i$  values.

For the left hand side, we computed the total system mass ( $M_{\text{tot}}$ ) for each binary by combining the orbital period ( $P$ ) and angular semi-major axes ( $a''$ ) from our fits to the orbital parameters (Section 4) with the parallaxes ( $\pi$ , Section 5.1), following:

$$M_{\text{tot}} = M_1 + M_2 = \frac{(a''/\pi)^3}{P^2}, \quad (3)$$

where  $P$  is in years,  $a''$  and  $\pi$  are in arcseconds, and  $M_{\text{tot}}$  is in Solar masses.

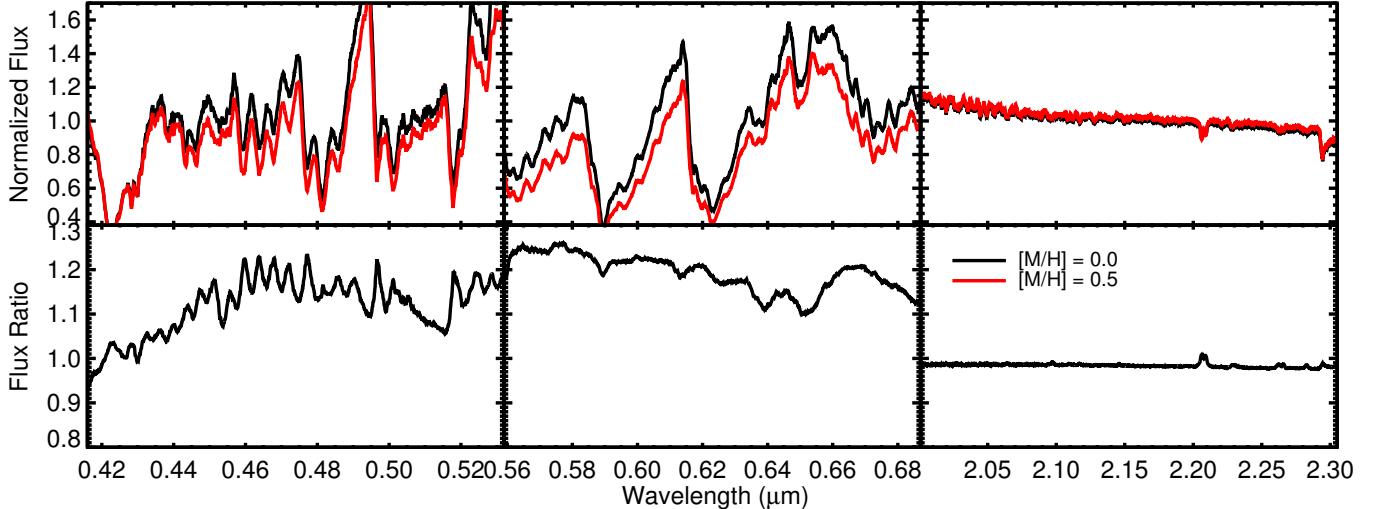
Because Equation 3 does not provide component star masses, when fitting for the  $a_i$  coefficients we performed the comparison between the predicted and dynamical *total* mass for each system. For this, we rewrote Equation 2 in terms of total mass:

$$M_{\text{tot}} = 10^{b_1} + 10^{b_2}, \quad (4)$$

$$b_j = \sum_{i=0}^n a_i (M_{K_S,j} - zp)^i, \quad (5)$$

where  $M_{K_S1}$  and  $M_{K_S2}$  are the primary and companion absolute  $K$ -band magnitudes, which were calculated from our measured  $\Delta K_S$  and unresolved  $K_S$  magnitudes (Section 5.3). Equation 4 could easily be modified for higher-order star systems, providing individual  $M_{K_S}$  magnitudes and the total mass of the system are known.

We fit for the  $a_i$  terms in Equation 4 using the MCMC code *emcee*, which lets us fit for the strong covariance between coefficients. Each coefficient evolved under uniform priors without limits, and was initialized to the best-fit value derived from *MPFIT*. We ran the MCMC chain with 800 walkers for 750,000 steps after a burn-in of 50,000 steps. We tested values of  $n$  (number of  $a_i$



**Figure 3.** Effect of changes in  $[M/H]$  on a model spectrum at  $T_{\text{eff}}=3200$  K,  $\log(g)=5$  in  $g$ -band (left),  $r$  (middle), and  $K$  (right). The top panel shows  $[M/H]=0$  (black) and  $[M/H]=+0.5$  (red) spectra from the CFIST BT-SETTL models (Allard et al. 2012). The bottom panel shows the ratio of the two, highlighting how small an effect  $[M/H]$  has in the  $K$ -band compared to optical regions. The one feature that stands out in the  $K$ -band is the Na doublet, which is a commonly used as a metallicity diagnostic for dwarfs (Rojas-Ayala et al. 2010; Terrien et al. 2012; Newton et al. 2014), and a gravity diagnostic for pre-main-sequence stars (e.g., Schlieder et al. 2012).

terms) from three to six. We found significant improvement in both  $\chi^2$  and the Bayes factor for the best-fit relation using up to five constants, but both metrics gave negligible improvement when we move to a sixth.

Errors on the dynamical mass are correlated to the errors on  $M_{K_S}$ . The system mass scales with  $a^3$ , which depends linearly with the parallax. As a result, the parallax was a major, if not the largest, source of uncertainty on the mass for most systems. Similarly, our  $K_S$  magnitudes had relatively small errors (0.01-0.03 mag), so  $M_{K_S}$  errors tended to be dominated by the parallax. Because this correlation is usually along (parallel to) the direction of the  $M_{K_S} - M_*$  relation, it can improve the fit if properly taken into account (when compared to assuming uncorrelated errors).

Our strategy was to treat distances of each system as a free parameter, letting them evolve under a strong prior from the observed parallax. The MCMC fit to Equation 4 was provided  $a''^3/P^2$  and  $K_S$  (with uncertainties) for each system, from which total mass and  $M_{K_S}$  were recomputed from the varying parallax of each system at each step in the MCMC before making a comparison. Since the orbital information provides no constraint on the distance, this effectively forced the parallax to explore a distribution along the input prior. It is possible for the parallax posterior to shift off this prior if it provided better agreement with the relation as determined by the rest of the sample. However, in practice this does not

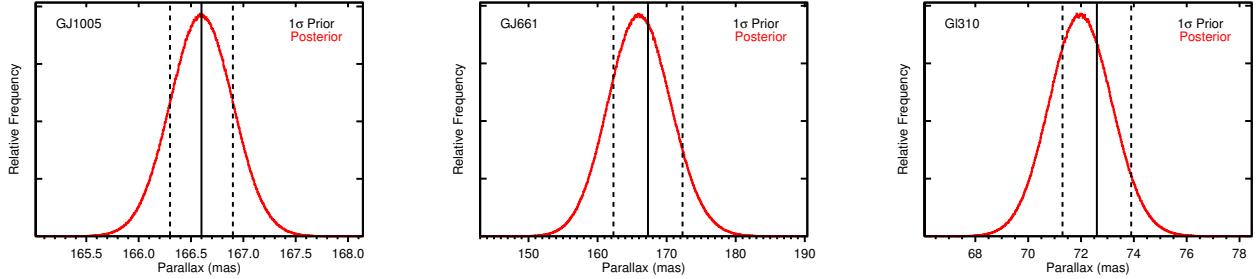
happen; for all systems the parallax posterior matched the input values (Figure 4).

For computational efficiency we assumed one-dimensional, Gaussian errors on  $a''^3/P^2$ , rather than including the orbital posteriors inside of our MCMC fit. Errors on  $a''$  and  $P$  from our orbits fits are correlated and sometimes non-Gaussian. However, the posteriors on  $a''^3/P^2$  were described well by a Gaussian (Figure 5). The only deviations were the systems with the worst parameters overall, which tended to have somewhat asymmetric posteriors. However, these systems had little influence over the final fit.

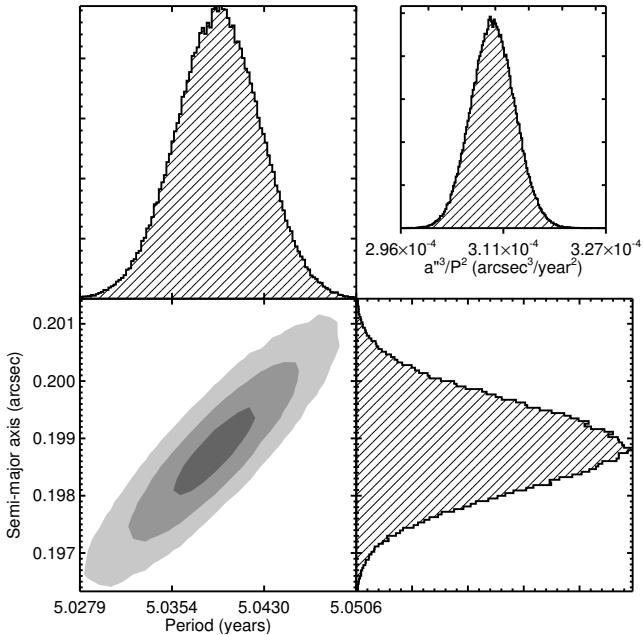
For main-sequence dwarfs at fixed metallicity, more massive stars should always be brighter. To force this on our relation, we required that the resulting fit have a negative derivative (higher  $M_{K_S}$  always gives lower mass) over the full range of input objects considered. We tested running without this constraint, and found similar results over most of the parameter range considered. The major difference was near the edges of the relation. With this constraint the fit would sometimes become double valued (two  $M_{K_S}$  values that yield the same mass) at or just beyond the input sample.

## 6.2. Results and Uncertainties

We show the output coefficient posteriors in Figure 6. The spread in the posteriors for individual coefficients is rather large, suggesting significant errors on each coefficient and a poor overall fit. However, the coefficients are



**Figure 4.** A comparison of example input  $1\sigma$  priors (black lines) to the output posteriors (red distributions). Gl 310 represents one of the largest differences between the prior and posterior, highlighting that the measured parallax prior dominates the MCMC exploration.



**Figure 5.** Example joint posterior on semi-major axis and orbital period (bottom left) for the system LHS 6167. Grey regions show 1, 2, and  $3\sigma$  (from darkest to lightest) of the points. The histograms above and to the right show the one-dimensional distributions of each parameter. The parameter fed into the  $M_{K_S} - M_*$  fit is  $a''^3/P^2$  (in arcsec $^3$  per year $^2$ ), which we show in the top right inset.

also highly correlated. So for a given value assigned to four of the five fit coefficients, there is a relatively narrow set of solutions for the final constant.

The final fit was tightly constrained over most of the sequence which we show in Figure 7 for individual masses, and combined masses in Figure 8 (a more realistic representation of how the fit was done). We provide the best fit (highest likelihood) coefficient values in Table 1. Since the gain from four to five coefficients was small (but significant), we include coefficients for both in Table 1,

but we adopt the five-parameter (4th order) solution for all analyses and figures.

To estimate the uncertainty the relation we computed the standard deviation in the derived masses for a fixed  $M_{K_S}$  across all MCMC steps, which we report in Table 2. In the best-constrained regions ( $5 \lesssim M_{K_S} \lesssim 7$ ,  $0.3 \lesssim M_* \lesssim 0.6$ ) the relation is good to 2-3%. However, the fit is significantly less-well constrained near the edges, where mass errors are 7-8%. Some of this is due to expected behavior of polynomials near the edge and a lack of data on the extreme ends. Stars at the high-mass end of our sample also have weaker orbital constraints overall.

The values in Table 2 can be used as an approximate reference for estimating errors on the output mass for a given  $M_{K_S}$  if combined with errors from the distance and  $K_S$  magnitude. However, this can be misleading, as the posterior is not Gaussian or symmetric over much of the mass range. It is better to sample the posterior at a given  $M_{K_S}$  to produce a posterior in mass. To facilitate use of our  $M_{K_S} - M_*$  and help provide realistic uncertainties, we include the fit posteriors and provide a simple code that provides output  $M_*$  posteriors<sup>9</sup> given an  $M_{K_S}$  and error.

The fit with the highest likelihood yields a  $\chi^2_\nu$  of 1.5 with 57 degrees of freedom (accounting for correlated errors). The larger than expected scatter is primarily driven by a handful of systems that agree with the best-fit to within  $\simeq 3\%$ , but also have small total mass uncertainties ( $\lesssim 1\%$ ). The MCMC explores the range of solutions, so if we factor in the scatter in the final  $M_{K_S} - M_*$  relation then the  $\chi^2_\nu$  value decreases to 0.98. As a result, using the fit posteriors provides uncertainties on  $M_*$  that encompass the larger scatter.

The larger than expected scatter could be due to systematically underestimated uncertainties, most likely in the input distances. As we noted with the *Gaia* DR2

<sup>9</sup> [https://github.com/awmann/M\\_-M\\_K-](https://github.com/awmann/M_-M_K-)

**Table 1.** Best-fit Coefficients for Equations 4 and 6

$a_0$	$a_1$	$a_2$	$a_3$	$a_4$	f	$\chi^2_\nu$
-0.663193	-0.203256	$8.202319 \times 10^{-3}$	$5.389355 \times 10^{-3}$	$-5.489318 \times 10^{-4}$	...	1.46
-0.642761	-0.201882	$-7.462048 \times 10^{-4}$	$4.741565 \times 10^{-3}$	...	0.058	1.58
-0.649004	-0.201198	$1.770185 \times 10^{-4}$	$4.701115 \times 10^{-3}$	...	...	1.96
-0.641361	-0.205263	$3.107211 \times 10^{-3}$	$5.301090 \times 10^{-3}$	$-2.199428 \times 10^{-4}$	0.068	1.42

NOTE—The first row is the preferred fit, as the  $f$  parameter is currently considered not significant. If [Fe/H] is considered the second row is preferred over the last.

parallaxes (Section 5.1), orbital motion can shift the centroid and hence impact the measured parallax. This is particularly true for systems with orbital periods near an alias of one year (the parallax signal’s period). Most of the systems included were not impacted by this because they either had have their parallaxes drawn from a nearby bound (and single) star, were calculated accounting for the presence of the binary, or have orbits far from the timescale of the parallactic signal. But a small impact on the sample could explain the higher scatter.

If the parallactic or orbital errors were underestimated this could bias our result towards the most precise points. We explored this possibility by redoing our MCMC analysis adding a free parameter that uniformly increases the parallax uncertainty of all systems. The final fit parameters and resulting relation were negligibly different from those derived without the added error term, and the only meaningful change was a reduction in the  $\chi^2_\nu$  of the best-fit solution. Because the change was small, we did not explore adding more sophisticated error terms to the adopted parallax, such as using the orbital fits to model the deviation in the unresolved binary centroid.

Another possibility for this scatter is stellar physics not well accounted for in Equation 4, such as metallicity.

### 6.3. The role of metallicity

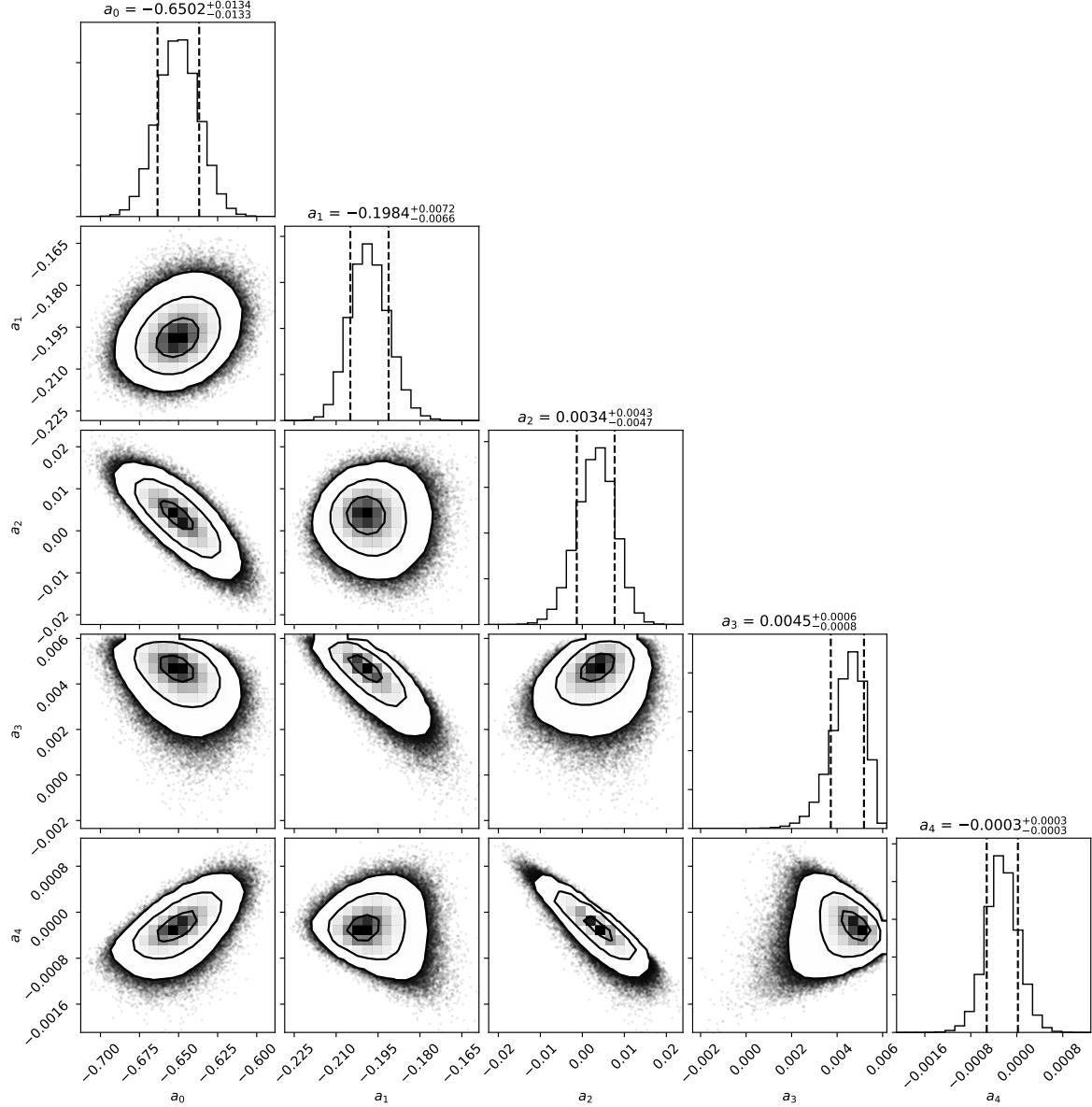
We explored the expected effects of [Fe/H] on the  $M_{K_S} - M_*$  relation using the Mesa Isochrones and Stellar Tracks (MIST, Dotter 2016; Choi et al. 2016). We also tested the Dartmouth Stellar Evolution Database (DSEP, Dotter et al. 2008), and found results to be broadly in agreement. We prefer MIST because their model atmospheres include a newer set of molecular lines (e.g., TiO) that are important in M dwarfs and may impact the synthetic photometry. The Lyon models (BHAC15, Baraffe et al. 2015) use a wrapper atmosphere in reasonable agreement with nearby stars with interferometrically determined  $T_{\text{eff}}$  values (Boyajian et al. 2012; Mann et al. 2013b), and go to lower masses than DSEP

and MIST. However, the BHAC15 models cover only solar metallicity.

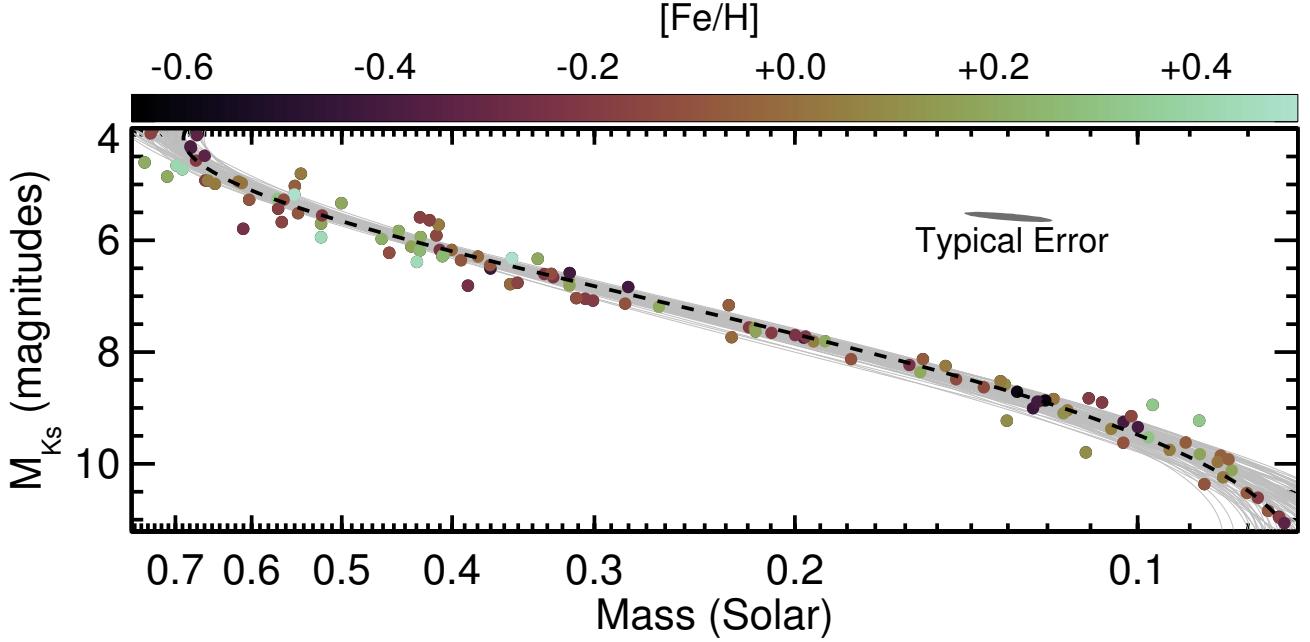
We show the expected  $M_{K_S}$  tracks from MIST for different metallicities in Figure 9 alongside our empirical determinations. MIST models do not extend below  $0.1M_\odot$ , so we restricted our comparison to above that mass. For this comparison we assumed a fixed age of 5 Gyr, although the choice of age from 1-10 Gyr makes a negligible difference for the mass range shown. Note that metal-rich stars are expected to be *less luminous* in  $M_{K_S}$  (and  $L_{\text{bol}}$ ) for a fixed  $M_*$ , whereas the opposite trend is seen for a fixed  $T_{\text{eff}}$  and most color selections.

Above  $0.4M_\odot$  there is a slight trend for metal-rich stars in our sample to land below the median sequence, as expected from the models. However, many metal-poor stars also land below the sequence, and there is no obvious trend below  $0.4M_\odot$ . We made this comparison more quantitative by interpolating over the model grid to create a predicted mass for each system using  $M_{K_S}$  and [Fe/H]. The predicted system masses were statistically worse than those from our best-fit relation ( $\chi^2_\nu$  of 6.5 using the model grid versus 1.5 with Equation 4), and worse than simply assuming every star has [Fe/H]=-0.3 or [Fe/H]=0.0 ( $\chi^2_\nu$  of 3.9 and 4.1). Allowing a zero-point correction to all metallicities (i.e., assuming they are systematically off from the model scale) improves the result ( $\chi^2_\nu = 4.0$ ), but a metallicity-free solution is still preferred. This strongly suggests that the models are overestimating the effect of [Fe/H].

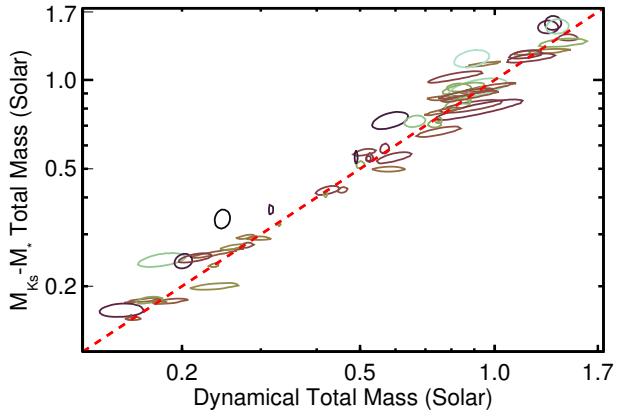
The residuals from our best-fit relation confirmed a weak (or no) effect on the derived  $M_*$  due to changes in [Fe/H], as we show in Figure 10. A Spearman rank test yielded no significant correlation between the residuals and [Fe/H]. We tried resampling the measurements using their uncertainties, and <1% of samples showed a significant correlation. We also repeated this test, but restricted to just the best-characterized systems (< 5% precision on mass) and still found no significant trend with [Fe/H].



**Figure 6.** Posterior projections for the  $a_i$  values derived from our MCMC fit to Equation 4. Contours denote the 1, 2, and 3 $\sigma$  confidence intervals, and the dashed lines in the histogram mark  $1\sigma$ . Figure was generated using corner.py (Foreman-Mackey 2016).

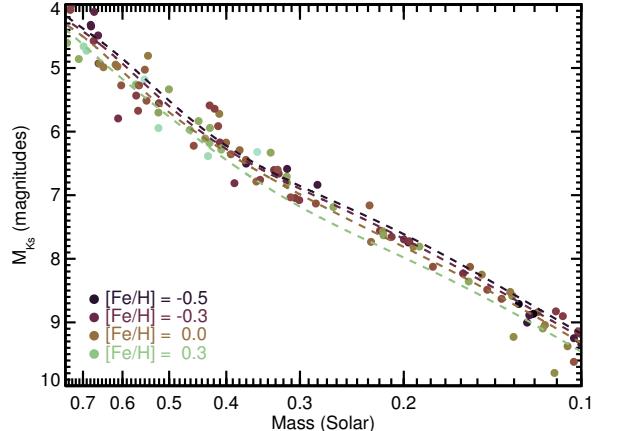


**Figure 7.** Absolute  $K_S$ -band magnitude as a function of mass for targets in our sample (circles). Points are color-coded by their estimated metallicity. We show a typical error ellipse in grey to highlight that errors on mass and  $M_{K_S}$  measurements are correlated, and usually parallel to the  $M_* - M_{K_S}$  relation (both depend on the parallax). The black dashed line indicates the best-fit (highest likelihood) from our MCMC analysis. To provide an estimate of the scatter in the relation as a function of mass, we show 100 randomly selected fits from the MCMC chain in grey. Note that our orbit fits only provide *total* mass; we used the mass ratios derived from the resulting  $M_{K_S} - M_*$  relation for the purposes of this figure.



**Figure 8.** Predicted total (system)  $M_*$  from the  $M_{K_S} - M_*$  relation as a function of the total mass determined from the orbit fit. Ellipses represent  $\simeq 1\sigma$  distribution of values for each point, accounting for parallax errors common to both the predicted and dynamical mass. Color-coding by  $[Fe/H]$  matches that of Figure 7.

Our sample is limited in its  $[Fe/H]$  range; 64% of the targets are  $-0.2 < [Fe/H] < +0.2$  and only one target has  $[Fe/H] < -0.5$ . It is possible that our best-fit relation



**Figure 9.**  $M_{K_S}$  as a function of  $M_*$  using MIST tracks of different metallicities (dashed lines) compared to empirical mass determinations (points). Color-coding by metallicity is the same for the points and lines, and matches the color scale of Figure 7. The plot cuts at  $0.1M_\odot$ , as the MIST models do not go below this limit.

masked any  $[Fe/H]$  term by shifting the fit to match the typical metallicity of stars at a given  $M_{K_S}$ . We explored

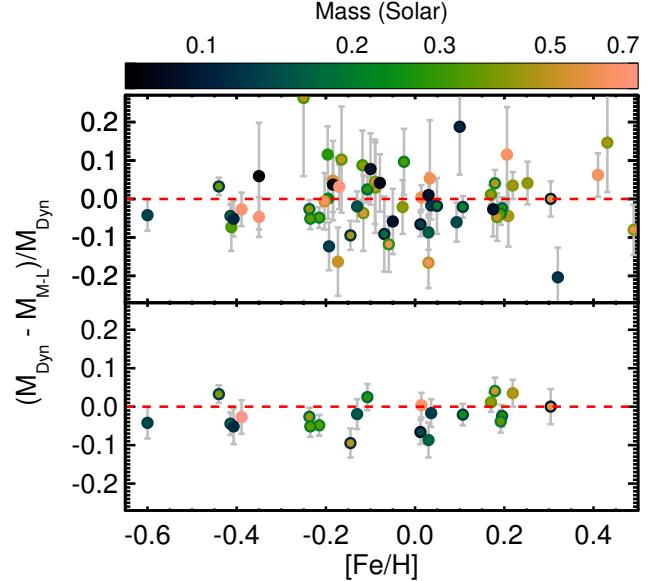
**Table 2.** Error in  $M_{K_S} - M_*$  Relation

$M_{K_S}$	$M_*$	$\sigma_{M_*}^a$	$\sigma_{M_*}^a$	SpT <sup>b</sup>
(mag)	$M_\odot$	$M_\odot$	%	
No [Fe/H] term, 5 $a_i$ terms				
4.0	0.712	0.024	0.034	K4.5
4.5	0.679	0.015	0.022	K6.5
5.0	0.611	0.013	0.021	M0.0
5.5	0.524	0.011	0.020	M1.5
6.0	0.434	0.009	0.020	M2.5
6.5	0.352	0.008	0.023	M3.0
7.0	0.281	0.008	0.028	M3.5
8.0	0.179	0.006	0.032	M4.5
8.5	0.144	0.005	0.032	M5.0
9.0	0.118	0.004	0.033	M6.0
9.5	0.100	0.004	0.038	M6.5
10.0	0.086	0.004	0.047	M7.5
10.5	0.077	0.005	0.060	M9.0
11.0	0.071	0.006	0.080	L1.0
[Fe/H] term, 4 $a_i$ terms				
4.0	0.731	0.028	0.038	K4.5
4.5	0.679	0.015	0.023	K6.5
5.0	0.605	0.011	0.018	M0.0
5.5	0.519	0.010	0.020	M1.5
6.0	0.434	0.009	0.021	M2.5
6.5	0.354	0.008	0.022	M3.0
7.0	0.285	0.006	0.022	M3.5
8.0	0.182	0.004	0.025	M4.5
8.5	0.147	0.004	0.028	M5.0
9.0	0.120	0.004	0.033	M6.0
9.5	0.100	0.004	0.038	M6.5
10.0	0.086	0.004	0.044	M7.5
10.5	0.076	0.004	0.054	M9.0
11.0	0.071	0.005	0.068	L1.0

NOTE—This table assumes  $M_{K_S}$  (and [Fe/H]) are known perfectly. Total errors on  $M_*$  should take into account errors in the measured parameters and the relation.

<sup>a</sup>The uncertainty in the resulting  $M_*$  at a given  $M_{K_S}$

<sup>b</sup>Spectral types are given for reference, but are extremely rough because of a significant dependence on metallicity. It is not recommended to use this table as a means to compute  $M_{K_S}$  or  $M_*$  from a spectral type or vice-versa.



**Figure 10.** Fractional difference between the orbital and predicted system mass as a function of metallicity of the system. The top panel contains all systems, while the bottom shows just those with <5% uncertainties on  $M_*$ . Points are color-coded by the masses of components, with the inner dot corresponding to the primary star's estimated mass, and the outer circle the companion's estimated mass.

this by fitting for a term in [Fe/H] of the form:

$$M_{tot} = (1 + f[\text{Fe}/\text{H}]) (10^{b_1} + 10^{b_2}), \quad (6)$$

where  $b_j$  is defined in Equation 5. This form assumes that a linear change in [Fe/H] corresponds to a fractional change in  $M_*$  (e.g.,  $f=0.1$  would correspond to a 10% change in derived  $M_*$  per dex for a fixed  $M_{K_S}$ ). This is generally consistent with the evolutionary models, although there is a change in the scale ( $f$ ) over the range of masses considered here.

For this analysis we excluded the two L dwarfs from the sample because of questionable assigned metallicities. As with our fit to Equation 4, we tested range of values for  $n$  (number of  $a_i$  coefficients). When including the [Fe/H] term, we found significant improvements in the fit up to  $n = 4$ , compared to  $n = 5$  without  $f$ . Our MCMC fitting method was otherwise identical to that outlined in Section 6. Since we were primarily interested in the relative effect of [Fe/H], and not necessarily the absolute scale, we assumed the smaller (relative) errors on [Fe/H] (see Section 5.2).

We show the output coefficient posteriors in Figure 11 and list the best-fit coefficients in Table 1 for both the third- ( $n = 4$ ) and fourth-order ( $n = 5$ ) fits. Although

the best-fit value of  $f$  is relatively small ( $\simeq 0.06$  per dex [Fe/H]), the spread is rather broad ( $\pm 0.06$  per dex), a consequence of the narrow range of [Fe/H] values compared to overall measurement uncertainties.

To compare to the models, we fit the MIST grid points in the same manner as the empirical dataset following Equation 6. In Figure 12 we show the posterior on  $f$  from the model grid compared to that from the dynamical masses. While the posteriors have some overlap, more than 96% of the difference between the two posteriors ( $f_{\text{model}} - f_{\text{dynamical}}$ ) is  $> 0$ , indicating (with marginal significance) that the models overpredict the effect of [Fe/H] on the  $M_{K_S} - M_*$  relation. It is unclear if (if true) this applies to the full mass range considered; there is some indication that the metallicity effect is stronger for higher-mass stars in the sample. The models suggest  $f$  is  $\simeq 30\%$  larger for  $[\text{Fe}/\text{H}] < -1.0$ . Thus a set of subdwarf binaries will provide more leverage to test over what mass range, if any, the discrepancy with the models is statistically significant.

#### 6.4. Comparison to previous relations

##### 6.4.1. Delfosse et al. (2000)

Delfosse et al. (2000) is one of the most commonly used  $M_{K_S} - M_*$  relations, covering  $0.1M_\odot < M_* < 0.6M_\odot$ . Like our work, the calibration was built primarily on astrometric binaries. Nearly all the targets in Delfosse et al. (2000) were included in our sample, with the exception of triple systems and eclipsing binaries, both of which we intentionally avoided because of the complexity of computing total masses in the first scenario and estimating  $M_{K_S}$  in the latter. Because of the sample overlap, consistency is expected. However, a comparison can be useful to see how past use of Delfosse et al. (2000) may change with the updated relation.

We show the comparison in Figure 13, including the points used in the Delfosse et al. (2000) calibration as well as the two fit lines. The relations are in agreement to better than 10% over the full mass range considered, and within 5% over most of the calibration region ( $0.15M_\odot \lesssim M_* \lesssim 0.5M_\odot$ ). Given errors often quoted for the Delfosse et al. (2000) relation (5-10%) the relations are extremely consistent.

##### 6.4.2. Mann et al. (2015)

Mann et al. (2015) built a catalog of 183 M dwarfs with precision  $T_{\text{eff}}$  and  $R_*$ , constructed by calibrating methods using empirical radii measurements from long-baseline optical interferometry (Boyajian et al. 2012) and precision bolometric fluxes (e.g., Mann & von Braun 2015; Boyajian et al. 2015). Masses were measured for these stars by interpolating the parameters onto an updated version

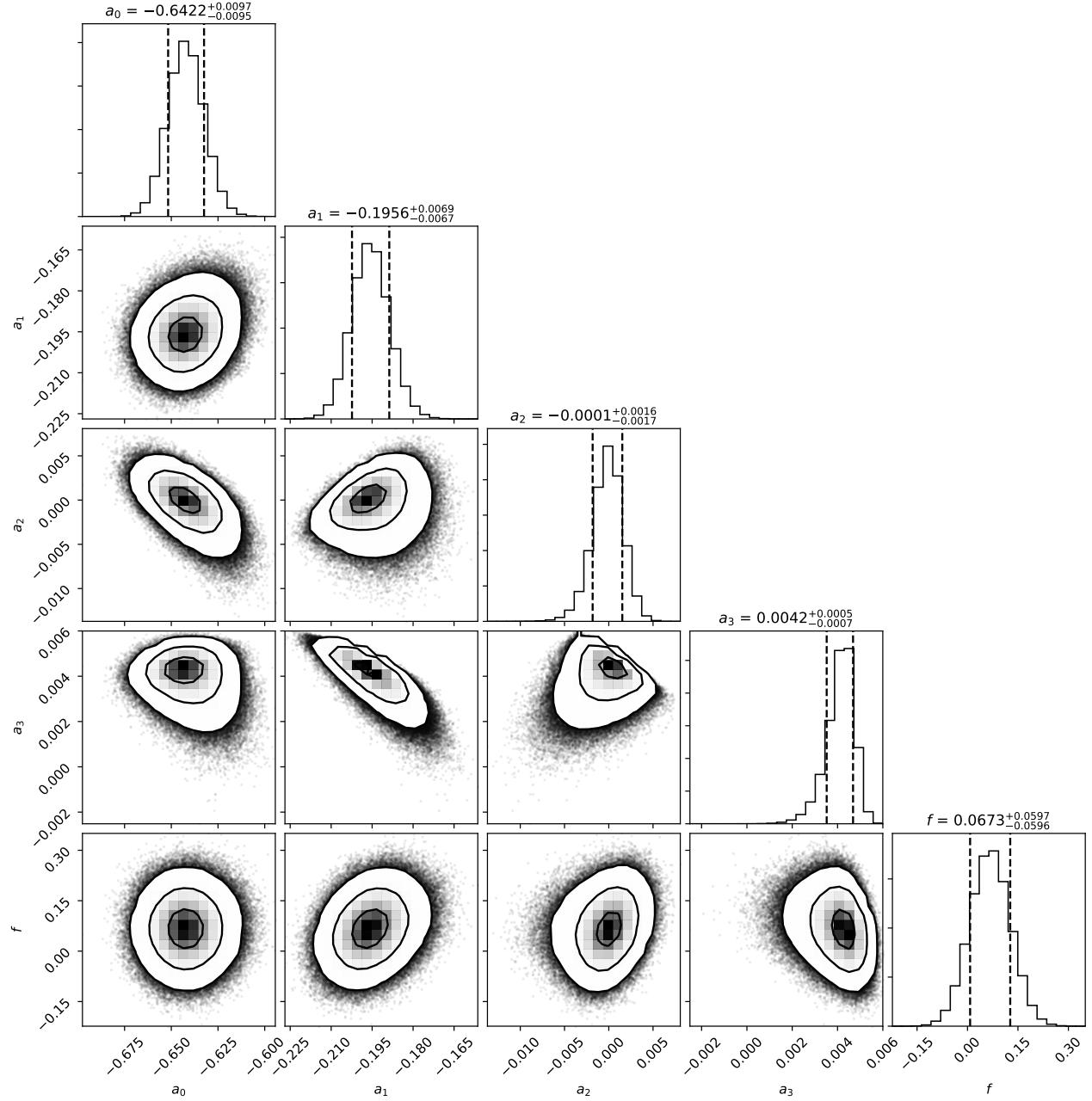
of the DSEP models as described in Feiden & Chaboyer (2013, 2014) and Muirhead et al. (2014). Although these masses were model-dependent, when combined with the more empirical radii, they precisely reproduced the mass-radius relation from low-mass eclipsing binaries. The consistency between these relations and the empirical determinations from eclipsing binaries suggested that the model-based masses were accurate to  $\simeq 3\%$  or better, and motivated the development of a  $M_{K_S} - M_*$  relation from the Mann et al. (2015) sample. A comparison to our relation can be seen in part as a test on the updated DSEP models, in addition to the results given in Mann et al. (2015).

We show our fit with uncertainties alongside Mann et al. (2015)'s in Figure 14. The two fits track each other extremely well, with a maximum divergence of  $\simeq 5\%$ . Given the quoted 2-3% uncertainties from Mann et al. (2015) and the 2-3% scatter in our relation, this difference is not significant. There is a hint of tension at the high-mass end, where our MCMC posterior is asymmetric, and at  $0.2-0.3M_\odot$  where the difference is the largest, but never does the offset exceed the estimated uncertainties.

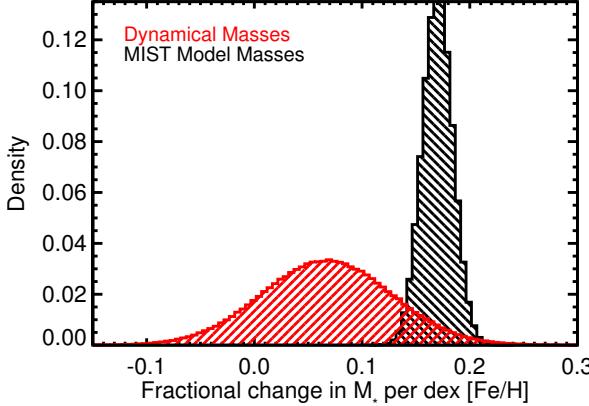
##### 6.4.3. Benedict et al. (2016)

Like our work, the Benedict et al. (2016) relation was also based primarily on masses derived from M dwarf astrometric binaries. The Benedict et al. (2016) uses absolute astrometry from *HST* fine guidance sensors and radial velocities for a subset of systems. In addition to the exquisite precision provided by *HST*, this combination yields individual masses, and independent constraints on the systems parallax. So although our sample is larger and contains most of the targets in Benedict et al. (2016), their analysis is not subject to many of the complications of our own. Differences between the two results could therefore shed light on complications with our use of total masses and perhaps reveal complications with our input parallaxes.

We compare our  $M_{K_S} - M_*$  relation Benedict et al. (2016)'s in Figure 15. The two relations are in excellent agreement for  $0.09M_\odot \lesssim M_* \lesssim 0.25M_\odot$ . Below this regime, the Benedict et al. (2016) fit is anchored by the one star, GJ1245C. They have only two other stars in this low-mass regime (GJ 2005B and C) and were poorly characterized compared to the rest of their sample. GJ1245AC is in our sample, but we use a parallax from Lindegren et al. (2018) on GJ 1245B for this system, which places it  $5.5\sigma$  more distant than the parallax adopted by Benedict et al. (2016). Our parallax-free orbital parameters for this system are in excellent agreement with Benedict et al. (2016), but the Lindegren



**Figure 11.** Same as Figure 6, but for the fit following Equation 6, i.e., including the [Fe/H] term,  $f$ .

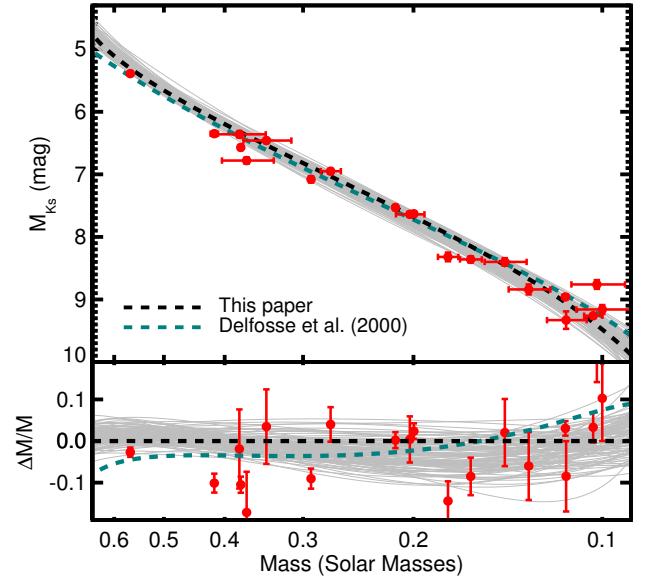


**Figure 12.** Comparison of the posterior on  $f$  (fractional change in  $M_*$  per dex in metallicity for a fixed  $M_{K_S}$ ; Equation 6) predicted by the MIST models (black) compared to that using our dynamical masses (red). There are an identical number of points in each posterior and the bin sizes are the same.

et al. (2018) parallax makes the final parameters more consistent with our overall relation. Since the Benedict et al. (2016) parallax accounted for the binarity of this system, the origin of difference between these two parallaxes unclear. If GJ 1245B is itself an unresolved binary, this could explain the discrepant parallax. If the Benedict et al. (2016) parallax is correct, we also need to explain why GJ 1245C is significantly more luminous than similar-mass stars.

Above  $0.3M_\odot$ , Benedict et al. (2016) yields masses as much as 10% higher than our own for a fixed  $M_{K_S}$ . Our fit agrees reasonably well with the binaries analyzed by Benedict et al. (2016) in this mass regime. The divergence is driven by points not analyzed by Benedict et al. (2016), but included in their fit from earlier literature sources.

Inspection of the literature data incorporated by Benedict et al. (2016) makes the origin of this difference clear; many of these are eclipsing binaries and have  $\Delta K$ -band magnitudes of mixed quality and/or lack parallaxes needed for a precise  $M_{K_S}$ . GU Boo, for example, has absolute magnitudes estimated from an optical eclipse depth combined with bolometric corrections (López-Morales & Ribas 2005), which are drawn from models that perform poorly on M dwarfs (Lejeune et al. 1998; Hauschildt et al. 1999). Similarly, for GJ 2069 AC (Cu Cnc) Benedict et al. (2016) adopted  $M_{K_S}$  from Ribas (2003) that disagree with the 2MASS  $K_S$  and *Gaia* DR2 parallax (for either AC or B) using any  $\Delta K_S$ . We conclude that the  $M_{K_S}$  determinations for these sys-



**Figure 13.** Absolute  $K_S$ -band magnitude as a function of mass for astrometric binaries analyzed by Delfosse et al. (2000) (red circles). The resulting  $M_* - M_{K_S}$  relation from Delfosse et al. (2000) is shown as a teal dashed line, while the best-fit relation from this paper is shown as a blacked dashed line (with error in grey as in Figure 7). The bottom panel shows the residual of the Delfosse et al. (2000) points compared to our relation, with the Benedict et al. (2016) relation in teal for reference. Errors in the residuals only reflect errors in the mass and  $M_{K_S}$ , and do not account for errors in our  $M_* - M_{K_S}$  relation.

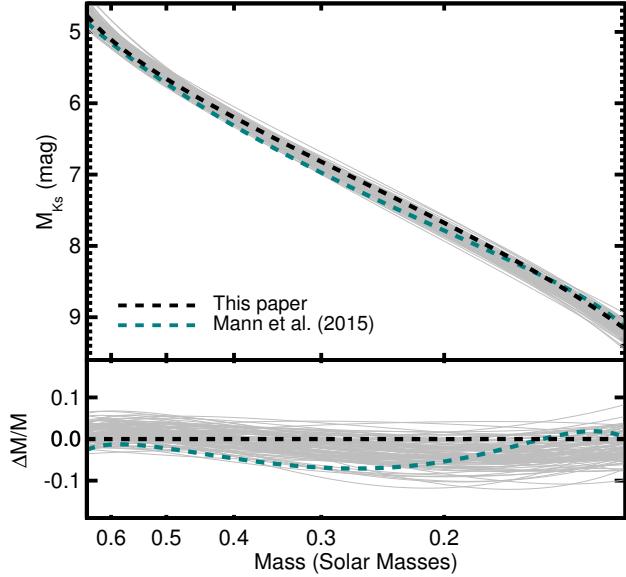
tems need to be revised before including them in future analyses of the  $M_{K_S} - M_*$  relation.

## 7. CONCLUSIONS & DISCUSSION

### 7.1. Summary

The mass-luminosity has proven to be a critical tool for estimating masses of cool stars for decades, and has applications from studying extrasolar planets to measuring the initial mass function and mass-to-light ratio in galaxies. The  $M_{K_S} - M_*$  relation has been particularly useful because  $M_{K_S}$  is easily measured if the stellar distance is known, and because it mitigates the effect of metallicity compared to other bands. With the arrival of *Gaia* parallaxes combined with existing 2MASS photometry, nearly all the early M-dwarfs out to  $\simeq 100$  pc and late M-dwarfs out to  $\simeq 50$  pc have precise ( $\lesssim 1\%$ ) parallaxes and  $M_{K_S}$  magnitudes, massively increasing the utility of this relation.

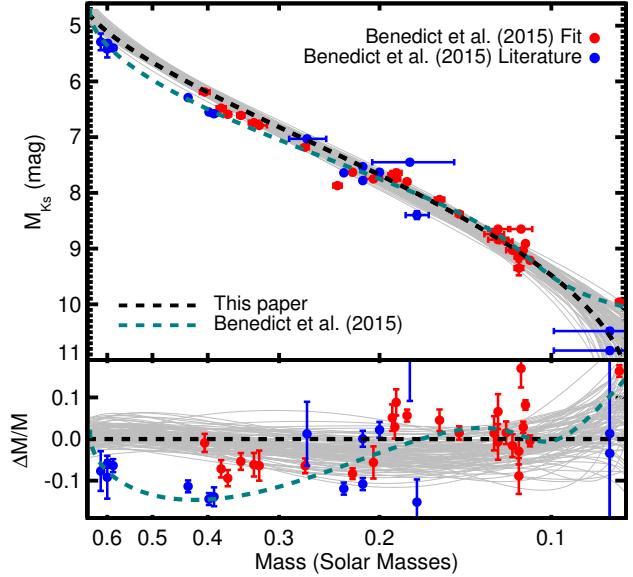
We endeavored to improve on existing  $M_{K_S} - M_*$  relations and explore the reliability of model predictions by expanding the sample of calibrators and exploring the



**Figure 14.** Comparison of  $M_* - M_{K_S}$  from Mann et al. (2015), shown as a teal dashed line, to that from this paper, which is shown in black, with 100 randomly selected realizations of the MCMC (as with Figures 7 and 15). Individual points from Mann et al. (2015) on which the calibration is based are not shown (for clarity), but they follow a tight sequence around the teal line. Only the range of masses covered by Mann et al. (2015) are shown.

role of [Fe/H]. As part of this effort, we have been monitoring a set of nearby late K and M dwarf visual binaries using adaptive optics imaging. We combined these data with similar data from Keck, CFHT, Gemini, and VLT archives, as well as astrometric measurements from the literature. The literature astrometry was particularly important to provide the baseline on long-period systems. We used these data and parallaxes from the literature or derived from MEarth astrometry to measure the total masses of a set of 52 binaries. We combined this with nine ultracool binaries from (Dupuy & Liu 2017), which were analyzed in a nearly identical way to the procedure used in this paper and provide significantly improved coverage near the hydrogen burning limit.

Following the formalism in Section 6, we used our dynamical total masses and resolved  $M_{K_S}$  magnitudes to fit for the empirical  $M_{K_S} - M_*$  relation. The result can provide masses precise to 2–3% from  $0.3 - 0.6M_\odot$  and  $< 5\%$  over the whole M dwarf sequence. The relation also covers almost an order of magnitude in  $M_*$ , from  $\simeq 0.7M_\odot$  down to the hydrogen burning limit ( $\simeq 0.075M_\odot$ ). The scatter around the best-fit is larger than expected from the mass uncertainties, especially when considering



**Figure 15.** Absolute  $K_S$ -band magnitude as a function of  $M_*$  for astrometric binaries analyzed by Benedict et al. (2016) (red circles) and those used in the Benedict et al. (2016) relation, but pulled from the literature (blue circles). The resulting  $M_* - M_{K_S}$  relation from Benedict et al. (2016) is shown as a teal dashed line, while the best-fit relation from this paper is as a blacked dashed line (with random samplings in grey as in Figure 7). The bottom panel shows the residual of the Benedict et al. (2016) points compared to our relation, with the Benedict et al. (2016) relation in teal for reference. Errors in the residuals only reflect errors in  $M_*$  and  $M_{K_S}$ , and do not account for errors in our  $M_* - M_{K_S}$  relation.

that errors in  $M_{K_S}$  and  $M_*$  are correlated. This might be due to underestimated errors in the parallaxes due to centroid motion of the binary, or intrinsic scatter from additional physics (e.g., metallicity and age).

Using empirically calibrated spectroscopic abundances, we explored the role of [Fe/H] on the  $M_{K_S} - M_*$  relation, comparing our results to expectations from evolutionary models as a guide. While the MIST models follow our empirical data relatively well, there is some evidence (96% chance) that the models are overestimating the effect of [Fe/H]. We consider this suggestive, but not conclusive at this time, as the narrow metallicity range of our sample limits our ability to explore this further.

We compared our relation to recent similar relations in the literature. Given quoted uncertainties, both the Delfosse et al. (2000) and Mann et al. (2015) relations follow ours well over most of the sequence. Since the Mann et al. (2015) masses are rooted in an updated version of the DSEP code, the level of agreement suggests

that these models perform extremely well in predicting the masses of old, main-sequence M dwarfs.

Our results agreed well with the systems analyzed by Benedict et al. (2016), but our relation diverges from Benedict et al. (2016) above  $\simeq 0.35M_{\odot}$ . We attribute this difference to literature points included in the Benedict et al. (2016) fit (mostly eclipsing binaries) with uncertain distances and  $\Delta K_S$  magnitudes. Increasing availability of *Gaia* parallaxes for these systems as well as ongoing efforts to measure their eclipses in range of wavelengths (e.g., Han et al. 2017) should significantly improve their utility for studying the mass-luminosity relation of M dwarfs.

### 7.2. Complicating factors

Our analysis made heavy use of literature astrometry (positions and separations). Many of these measurements had either no reported uncertainties, or underestimated uncertainties due to uncorrected systematics (e.g., image distortion). Our method for assigning or correcting errors assumed that all measurements have a common missing error term per data source. It is more likely that errors depend on the separation and contrast ratio, as well as quantities that were not consistently reported, like weather, setup, and observational strategy. For references with reported errors, these effects were likely be taken into account already (and were preserved by adding a term in quadrature with reported errors). However, for those with no errors, we may have overestimated errors for well-resolved binaries, and underestimated it for marginally or unresolved systems. Further, this technique assumed an uncorrelated error term. In the case of an erroneous pixel scale or imperfectly aligned instrument, all measurements from a common instrument may be off in the same direction. In practice, it is difficult to correct for these effects, especially in cases where multiple instruments or instrumental setups are used. The number of binaries with sufficiently well-characterized orbits was too small for deriving more robust error terms. So we adopted this as the best possible option to take full advantage of the long-baseline provided by the decades of data.

As a test of this method, we tried refitting a subset of 10 randomly selected binaries twice, first doubling the error term we added, then halving it. With the exception of one system that failed to converge with the doubled input errors, overall parameters and errors did not change significantly (although the final  $\chi^2_{\nu}$  values changed). The main reason for this is that our measurements are generally far more precise (with the exception of measurements from *HST*) and disproportionately weigh on the final solution. Further, in the

case of significantly underestimated errors, the MCMC usually explores a range of solutions to result in a similarly broad posterior as the one derived using larger input errors. The problem with this case is that MCMC will underweight more precise points (e.g., those from NIRC2/NaCo). Overall, many issues that might arise from a single source of data were mitigated because we used a wide range of data sources, each of which were subject to different systematics. We conclude this treatment of literature errors does not significantly impact our final result.

The presence of unresolved tight companions (triples or quadruples) could bias the overall relation. For the longest-period systems, a close-in (orbital period of  $< 10$  days) third star could be stable, and would not have been detected in our AO data. An unseen companion would drive the total mass of the system higher, but has much smaller effect on the total luminosity. So unresolved triples will sit *low* in Figure 7 (higher mass for a fixed  $M_{K_S}$ ). A few stars do sit seemingly well below the fit, but no more so than expected given their mass uncertainties. The closest systems have radial velocity data, which rules out the presence of all companions massive enough to impact our results. Such triples are relatively rare (Fischer & Marcy 1992; Raghavan et al. 2010), especially those with a companion low-mass enough to avoid perturbing the orbits (which would be detected as a high  $\chi^2_{\nu}$  value in our fits). We consider it unlikely that higher-order systems are significantly impacting our results.

We endeavored to select stars that land above the hydrogen burning limit, as stars below this range should not follow a strict mass-luminosity relation. Most of our objects near this boundary came from Dupuy & Liu (2017), and we adopted their classification as stellar or substellar. Although this classification has some uncertainty, and is model-dependent. It is possible that some objects near the low-mass edge of our sample are brown dwarfs. Inclusion of any such substellar (old) systems would bias the fit towards higher  $M_*$  for a given  $M_{K_S}$ . Objects just below this limit lose luminosity slowly, and still may follow the same relation on  $\simeq$ Gyr timescales, so we expect this bias, if present, to be small. Further, the excellent agreement between all points  $< 0.85M_{\odot}$  suggests all objects are obeying the same  $M_{K_S} - M_*$  relation (Figure 7).

### 7.3. Suggestions when using our $M_{K_S} - M_*$ relation

To help users interested in using  $M_{K_S}$  to compute a realistic  $M_*$  and  $\sigma_{M_*}$ , we provide a simple code to sample the fit posterior. Before using that code or

the provided MCMC posteriors, take note of the following suggestions:

- The fit behaves poorly near the edges of the calibration sample ( $M_{K_S} \gtrsim 11.0$  and  $M_{K_S} \lesssim 4.0$ ). The scatter in the MCMC posterior accounts for this, but restrict use to  $11.0 < M_{K_S} < 4.0$  ( $0.07M_\odot < M_* < 0.70M_\odot$ ), and a ‘safe’ range would be  $10.5 < M_{K_S} < 4.5$  ( $0.08M_\odot < M_* < 0.68M_\odot$ ).
- Our relation is only valid for main-sequence stars, and the role of youth and activity is untested. Based on the Lyon models (Baraffe et al. 2015), we advise restricting use to  $>100$  Myr above  $0.4M_\odot$ ,  $>300$  Myr to  $0.2M_\odot$ ,  $>500$  Myr to  $0.1M_\odot$ , and  $> 1$  Gyr below  $0.1M_\odot$ . A safer cut would be to only use this on stars  $>1$  Gyr, similar to the input calibration sample.
- The result is only tested over  $-0.60 < [\text{Fe}/\text{H}] < +0.45$ . We provide a fit that attempts to take into account changes due to  $[\text{Fe}/\text{H}]$ , but the effect is poorly constrained (fractional change of  $0.06 \pm 0.06$  in  $M_*$  per dex in  $[\text{Fe}/\text{H}]$ ). Assuming the target is within the metallicity range of M dwarfs in the Solar neighborhood ( $-0.05 \pm 0.18$ , Gaidos et al. 2014) then the  $M_*$  uncertainty introduced by the unknown  $[\text{Fe}/\text{H}]$  value is  $\simeq 1\%$ . However, when targeting more metal-poor populations like the thick disk or halo ( $[\text{Fe}/\text{H}] \lesssim -0.5$ ) or the effect on the derived  $M_*$  will be  $>5\%$ . We advise avoiding this relation for stars or populations with  $[\text{Fe}/\text{H}] < -0.5$ . The effect of more detailed abundances (e.g.,  $[\alpha/\text{Fe}]$ ) are completely untested.
- Objects just below the hydrogen burning limit age slowly, so this relation may give reasonable results. But the uncertainties assume the target is a star. Since the boundary likely depends on metallicity (Burrows et al. 2001), it is not possible to use a simple  $M_{K_S}$  cut even if the target is known to be old. So we do not recommend using this relation for  $M_{K_S} > 10.5$  without independent information that an object is stellar.

#### 7.4. Future directions

We intentionally selected targets that had  $\Delta K$  measurements, as  $M_{K_S}$  was known to give the tightest relation with  $M_*$  for M dwarfs. Unfortunately, there is no other band with contrast ratios for all systems considered here. Most of our monitoring was done with NIRC2/Keck, so many systems have a  $\Delta H$ , but only about 1/3 of the sample have measurements in an optical band. This limits the utility of the sample, as *Gaia G*, *BP*, and *RP*

are now widely available for early and mid-M dwarfs, and are measured with better precision than 2MASS  $K_S$ . The increased capabilities of new speckle cameras (e.g., Horch et al. 2009; Tokovinin et al. 2010) offer the opportunity to add optical contrasts. These can be converted to *Gaia* bandpasses given reasonable assumptions about the component spectra, and turned into a  $M_G - M_*$  relation that can be easily applied to millions of K and M dwarfs. Complementary optical data also provide colors for individual components, from which we can measure component  $T_{\text{eff}}$  and luminosity (e.g., Mann et al. 2015; Kraus et al. 2017).

Our sample was limited mostly to stars in the Solar neighborhood, and hence was heavily biased towards the narrow  $[\text{Fe}/\text{H}]$  distribution of nearby stars. We identified five additional  $[\text{Fe}/\text{H}] < -0.5$  systems including two subdwarf binaries. However all systems had short baselines in the literature compared to their expected orbital periods. It may take 1-2 more years to complete their orbits at the required precision for this kind of analysis. The availability of *Gaia* parallaxes will also help improve the precision of the known metal-poor systems and aid in the identification of new ones.

Although the sample studied here contained 122 stars, we only had half that many data points (61) constraining the fit, as the mass ratios were not known. This also can add complications to the fit at the bottom of the sequence, where the relation is steeper. Here, small changes in the shape can have a significant impact on the implied mass ratio for a given system, providing a means to generate many reasonable solutions to the same set of system masses. This can be seen in Figure 7 as a spread in the range of possible solutions  $M_{K_S} > 10$ . The solution to this problem is to add in radial velocity observations for amenable systems, which provides a direct constraint on the mass ratio. These data can also be used to simultaneously fit for the parallax, which may turn out to be crucial given existing problems with *Gaia* parallaxes of binaries (see Section 5.1).

We intentionally tried to avoid young stars when selecting the sample considered here. The ages of our input sample is not known, preventing any study into the effects of age on the  $M_{K_S} - M_*$ . However, our larger sample of binaries contains a range of known members of nearby young moving groups (YMGs) and the Hyades cluster. Many of these have complete or nearly complete orbital solutions, and will soon provide a powerful set of mass benchmarks with known ages. These systems also span ages from 10-650 Myr, offering the chance to both test pre-main-sequence models of M dwarfs (Montet et al. 2015; Rizzuto et al. 2016; Czekala et al. 2016; Nielsen et al. 2016) and explore the role of activity on M dwarf

parameters (e.g., Spada et al. 2013; Kesseli et al. 2018). The current sample can be included in such work when combined with optical activity indicators and rotation periods like those expected from the Transiting Exoplanet Survey Satellite (*TESS*, Ricker et al. 2014).

AWM was supported through Hubble Fellowship grant 51364 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555.

Based on observations obtained at the Gemini Observatory (acquired through the Gemini Observatory Archive), which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina), and Ministério da Ciência, Tecnologia e Inovação (Brazil). Observations taken from programs GN-2008B-Q-57, GN-2009B-Q-10, GN-2010B-Q-9, and GN-2011A-Q-26.

Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under ESO programmes 071.C-0388(A), 072.C-0570(A), 073.C-0155(A), 075.C-0521(A), 075.C-0733(A), 077.C-0783(A), 078.C-0441(A), 079.C-0216(A), 080.C-0424(A), 081.C-0430(A), 082.C-0518(A), 082.C-0518(B), 085.C-0867(B), 086.C-0515(A), 086.C-0515(B), 090.C-0448(A), 091.D-0804(A), 098.C-0597(A), 382.C-0324(A), and 382.D-0754(A).

This research has made use of the Keck Observatory Archive (KOA), which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExScI), under contract with the National Aeronautics and Space Administration.

Based on observations obtained at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council of Canada, the Institut National des Sciences de l’Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii.

This work was supported by a NASA Keck PI Data Award (award numbers 1554237, 1544189, 1535910, and 1521162), administered by the NASA Exoplanet Science Institute. Data presented herein were obtained at the W. M. Keck Observatory from telescope time allocated to the National Aeronautics and Space Administration through the agency’s scientific partnership with the California Institute of Technology and the University of California.

The Observatory was made possible by the generous financial support of the W. M. Keck Foundation.

The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

This work presents results from the European Space Agency (ESA) space mission Gaia. Gaia data are being processed by the Gaia Data Processing and Analysis Consortium (DPAC). Funding for the DPAC is provided by national institutions, in particular the institutions participating in the Gaia MultiLateral Agreement (MLA). The Gaia mission website is <https://www.cosmos.esa.int/gaia>. The Gaia archive website is <https://archives.esac.esa.int/gaia>.

*Software:* emcee, corner.py, mpfit, scipy, pyfits, astropy, python, spextools, IDL

*Facilities:* Keck:II (NIRC2), IRTF (SpeX), CFHT (PUEO, aobir), VLT:Antu (NaCo); Gemini:North (NIRI)

## REFERENCES

- Allard, F., Homeier, D., & Freytag, B. 2012, Royal Society of London Philosophical Transactions Series A, 370, 2765
- Arsenault, R., Salmon, D. A., Kerr, J. M., et al. 1994, in Proc. SPIE, Vol. 2201, Adaptive Optics in Astronomy, ed. M. A. Ealey & F. Merkle, 833–842
- Ballard, S., & Johnson, J. A. 2016, ApJ, 816, 66
- Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, A&A, 577, A42
- Bartlett, J. L., Lurie, J. C., Riedel, A., et al. 2017, AJ, 154, doi:10.3847/1538-3881/aa8457
- Benedict, G. F., Henry, T. J., Franz, O. G., et al. 2016, AJ, 152, 141
- Bochanski, J. J., Munn, J. A., Hawley, S. L., et al. 2007, AJ, 134, 2418
- Bonfils, X., Delfosse, X., Udry, S., et al. 2005, A&A, 442, 635
- Bowler, B. P., Liu, M. C., Shkolnik, E. L., et al. 2012, ApJ, 753, doi:10.1088/0004-637X/753/2/142
- Boyajian, T., von Braun, K., Feiden, G. A., et al. 2015, MNRAS, 447, 846
- Boyajian, T. S., von Braun, K., van Belle, G., et al. 2012, ApJ, 757, 112
- Brewer, J. M., Fischer, D. A., Basu, S., Valenti, J. A., & Piskunov, N. 2015, ApJ, 805, 126
- Brewer, J. M., Fischer, D. A., Valenti, J. A., & Piskunov, N. 2016, ApJS, 225, 32
- Burrows, A., Hubbard, W. B., Lunine, J. I., & Liebert, J. 2001, Reviews of Modern Physics, 73, 719
- Casagrande, L., Schönrich, R., Asplund, M., et al. 2011, A&A, 530, A138
- Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
- Cohen, M., Wheaton, W. A., & Megeath, S. T. 2003, AJ, 126, 1090
- Conroy, C., & van Dokkum, P. 2012, ApJ, 747, 69
- Covey, K. R., Hawley, S. L., Bochanski, J. J., et al. 2008, AJ, 136, 1778
- Cushing, M. C., Vacca, W. D., & Rayner, J. T. 2004, PASP, 116, 362
- Czekala, I., Andrews, S. M., Torres, G., et al. 2016, ApJ, 818, doi:10.3847/0004-637X/818/2/156
- Delfosse, X., Forveille, T., Ségransan, D., et al. 2000, A&A, 364, 217
- Diolaiti, E., Bendinelli, O., Bonaccini, D., et al. 2000, A&AS, 147, 335
- Dittmann, J. A., Irwin, J. M., Charbonneau, D., & Berta-Thompson, Z. K. 2014, ApJ, 784, 156
- Dittmann, J. A., Irwin, J. M., Charbonneau, D., et al. 2017, Nature, 544, 333
- Dotter, A. 2016, ApJS, 222, 8
- Dotter, A., Chaboyer, B., Jevremović, D., et al. 2008, ApJS, 178, 89
- Doyon, R., Nadeau, D., Vallee, P., et al. 1998, in Proc. SPIE, Vol. 3354, Infrared Astronomical Instrumentation, ed. A. M. Fowler, 760–768
- Dressing, C. D., & Charbonneau, D. 2013, ApJ, 767, 95
- Dressing, C. D., Newton, E. R., Schlieder, J. E., et al. 2017, ApJ, 836, 167
- Dupuy, T. J., Kratter, K. M., Kraus, A. L., et al. 2016, ApJ, 817, 80
- Dupuy, T. J., & Liu, M. C. 2012, ApJS, 201, 19
- . 2017, ApJS, 231, 15
- Dupuy, T. J., Liu, M. C., & Bowler, B. P. 2009, ApJ, 706, 328
- Dupuy, T. J., Liu, M. C., Bowler, B. P., et al. 2010, ApJ, 721, 1725
- Evans, T. M., Ireland, M. J., Kraus, A. L., et al. 2012, ApJ, 744, 120
- Feiden, G. A., & Chaboyer, B. 2012, ApJ, 757, 42
- . 2013, ApJ, 779, 183
- . 2014, A&A, 571, A70
- Ferguson, D., Gardner, S., & Yanny, B. 2017, ApJ, 843, 141
- Finch, C. T., & Zacharias, N. 2016, AJ, 151, doi:10.3847/0004-6256/151/6/160
- Fischer, D. A., & Marcy, G. W. 1992, ApJ, 396, 178
- Foreman-Mackey, D. 2016, The Journal of Open Source Software, 24, doi:10.21105/joss.00024
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
- Gagné, J., Lafrenière, D., Doyon, R., Malo, L., & Artigau, É. 2014, ApJ, 783, 121
- . 2015, ApJ, 798, 73
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, ArXiv e-prints, arXiv:1804.09365
- . 2016, A&A, 595, A2
- Gaidos, E., Mann, A. W., Kraus, A. L., & Ireland, M. 2016, MNRAS, 457, 2877
- Gaidos, E., Mann, A. W., Lépine, S., et al. 2014, MNRAS, 443, 2561
- Goldin, A., & Makarov, V. V. 2006, The Astrophysical Journal Supplement Series, 166, 341
- Goodman, J., & Weare, J. 2010, Commun. Appl. Math. Comput. Sci., 5, 65
- Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, Vol. 5, No. 1, p. 65-80, 2010, 5, 65
- Han, E., Muirhead, P. S., Swift, J. J., et al. 2017, AJ, 154, doi:10.3847/1538-3881/aa803c

- Hansen, C. J., Kawaler, S. D., & Trimble, V. 2004, Stellar interiors : physical principles, structure, and evolution (Springer)
- Hartkopf, W. I., McAlister, H. A., & Mason, B. D. 2001, AJ, 122, 3480
- Hauschildt, P. H., Allard, F., & Baron, E. 1999, ApJ, 512, 377
- Hejazi, N., De Robertis, M. M., & Dawson, P. C. 2015, AJ, 149, 140
- Henry, T. J., Kirkpatrick, J. D., & Simons, D. A. 1994, AJ, 108, 1437
- Henry, T. J., & McCarthy, Jr., D. W. 1993, AJ, 106, 773
- Hinkel, N. R., Timmes, F. X., Young, P. A., Pagano, M. D., & Turnbull, M. C. 2014, AJ, 148, 54
- Hinkel, N. R., Young, P. A., Pagano, M. D., et al. 2016, ApJS, 226, 4
- Hodapp, K. W., Jensen, J. B., Irwin, E. M., et al. 2003, PASP, 115, 1388
- Horch, E. P., Veillette, D. R., Baena Gallé, R., et al. 2009, AJ, 137, 5057
- Janson, M., Bergfors, C., Brandner, W., et al. 2014, ApJ, 789, 102
- Janson, M., Hormuth, F., Bergfors, C., et al. 2012, ApJ, 754, 44
- Johnson, J. A., & Apps, K. 2009, ApJ, 699, 933
- Jurić, M., Ivezić, Ž., Brooks, A., et al. 2008, ApJ, 673, 864
- Kesseli, A. Y., Muirhead, P. S., Mann, A. W., & Mace, G. 2018, ArXiv e-prints, arXiv:1804.04133
- Kraus, A. L., Cody, A. M., Covey, K. R., et al. 2015, ApJ, 807, 3
- Kraus, A. L., Ireland, M. J., Huber, D., Mann, A. W., & Dupuy, T. J. 2016, AJ, 152, 8
- Kraus, A. L., Ireland, M. J., Martinache, F., & Lloyd, J. P. 2008, ApJ, 679, 762
- Kraus, A. L., Shkolnik, E. L., Allers, K. N., & Liu, M. C. 2014, AJ, 147, 146
- Kraus, A. L., Douglas, S. T., Mann, A. W., et al. 2017, ApJ, 845, 72
- Law, N. M., Dhital, S., Kraus, A., Stassun, K. G., & West, A. A. 2010, ApJ, 720, 1727
- Law, N. M., Hodgkin, S. T., & Mackay, C. D. 2008, MNRAS, 384, 150
- Lee, J., & Song, I. 2018, MNRAS, 475, 2955
- Lejeune, T., Cuisinier, F., & Buser, R. 1998, Astronomy and Astrophysics Supplement Series, 130, 65
- Lépine, S., Hilton, E. J., Mann, A. W., et al. 2013, AJ, 145, 102
- Lindegren, L., Hernandez, J., Bombrun, A., et al. 2018, ArXiv e-prints, arXiv:1804.09366
- Liu, M. C., Dupuy, T. J., & Ireland, M. J. 2008, ApJ, 689, 436
- Lloyd, J. P., Martinache, F., Ireland, M. J., et al. 2006, ApJL, 650, L131
- López-Morales, M., & Ribas, I. 2005, ApJ, 631, 1120
- Lu, J. R., Ghez, A. M., Yelda, S., et al. 2010, in Proceedings of the SPIE, Volume 7736, id. 77361I (2010)., Vol. 7736
- Mace, G. N., Mann, A. W., Skiff, B. A., et al. 2018, ApJ, 854, doi:10.3847/1538-4357/aaa8dd
- Malo, L., Artigau, É., Doyon, R., et al. 2014, ApJ, 788, 81
- Mann, A. W., Brewer, J. M., Gaidos, E., Lépine, S., & Hilton, E. J. 2013a, AJ, 145, 52
- Mann, A. W., Deacon, N. R., Gaidos, E., et al. 2014, AJ, 147, 160
- Mann, A. W., Feiden, G. A., Gaidos, E., Boyajian, T., & von Braun, K. 2015, ApJ, 804, 64
- Mann, A. W., Gaidos, E., & Ansdell, M. 2013b, ApJ, 779, 188
- Mann, A. W., Gaidos, E., Lépine, S., & Hilton, E. J. 2012, ApJ, 753, 90
- Mann, A. W., & von Braun, K. 2015, PASP, 127, 102
- Mann, A. W., Newton, E. R., Rizzuto, A. C., et al. 2016, AJ, 152, 61
- Mann, A. W., Vanderburg, A., Rizzuto, A. C., et al. 2018, AJ, 155, 4
- Markwardt, C. B. 2009, in Astronomical Society of the Pacific Conference Series, Vol. 411, Astronomical Data Analysis Software and Systems XVIII, ed. D. A. Bohlender, D. Durand, & P. Dowler, 251
- Martinez, A. O., Crossfield, I. J. M., Schlieder, J. E., et al. 2017, ApJ, 837, 72
- Masciadri, E., Brandner, W., Bouy, H., et al. 2003, A&A, 411, 157
- McConnell, N. J., Lu, J. R., & Mann, A. W. 2016, ApJ, 821, 39
- Montet, B. T., Bowler, B. P., Shkolnik, E. L., et al. 2015, ApJL, 813, L11
- Muirhead, P. S., Hamren, K., Schlawin, E., et al. 2012a, ApJL, 750, L37
- Muirhead, P. S., Johnson, J. A., Apps, K., et al. 2012b, ApJ, 747, 144
- Muirhead, P. S., Becker, J., Feiden, G. A., et al. 2014, ApJS, 213, 5
- Muirhead, P. S., Mann, A. W., Vanderburg, A., et al. 2015, ApJ, 801, 18
- Mulders, G. D., Pascucci, I., & Apai, D. 2015, ApJ, 798, 112
- Neuhäuser, R., Guenther, E. W., Wuchterl, G., et al. 2005, A&A, 435, L13
- Neves, V., Bonfils, X., Santos, N. C., et al. 2014, A&A, 568, A121

- . 2012, A&A, 538, A25
- Newton, E. R., Charbonneau, D., Irwin, J., et al. 2014, AJ, 147, 20
- Nielsen, E. L., De Rosa, R. J., Wang, J., et al. 2016, AJ, 152, 175
- Nutzman, P., & Charbonneau, D. 2008, PASP, 120, 317
- Park, C., Jaffe, D. T., Yuk, I.-S., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1
- Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, Proceedings of the National Academy of Science, 110, 19273
- Raghavan, D., McAlister, H. A., Henry, T. J., et al. 2010, ApJS, 190, 1
- Rayner, J., Bond, T., Bonnet, M., et al. 2012, in Proc. SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84462C
- Rayner, J. T., Toomey, D. W., Onaka, P. M., et al. 2003, PASP, 115, 362
- Reid, I. N., Cruz, K. L., Allen, P., et al. 2004, AJ, 128, 463
- Ribas, I. 2003, A&A, 398, 239
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9143, 20
- Riedel, A. R., Blunt, S. C., Lambrides, E. L., et al. 2017, AJ, 153, 95
- Rizzuto, A. C., Ireland, M. J., Dupuy, T. J., & Kraus, A. L. 2016, ApJ, 817, 164
- Rizzuto, A. C., Mann, A. W., Vanderburg, A., Kraus, A. L., & Covey, K. R. 2017, AJ, 154, 224
- Rojas-Ayala, B., Covey, K. R., Muirhead, P. S., & Lloyd, J. P. 2010, ApJL, 720, L113
- . 2012, ApJ, 748, 93
- Schlieder, J. E., Lépine, S., Rice, E., et al. 2012, AJ, 143, 114
- Ségransan, D., Delfosse, X., Forveille, T., et al. 2000, A&A, 364, 665
- Seifahrt, A., Röll, T., Neuhauser, R., et al. 2008, A&A, 484, 429
- Service, M., Lu, J. R., Campbell, R., et al. 2016, PASP, 128, 095004
- Shields, A. L., Ballard, S., & Johnson, J. A. 2016, PhR, 663, 1
- Shkolnik, E. L., Allers, K. N., Kraus, A. L., Liu, M. C., & Flagg, L. 2017, AJ, 154, doi:10.3847/1538-3881/aa77fa
- Shkolnik, E. L., Anglada-Escudé, G., Liu, M. C., et al. 2012, ApJ, 758, 56
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
- Söderhjelm, S. 1999, A&A, 341, 121
- Somers, G., & Stassun, K. G. 2017, AJ, 153, 101
- Spada, F., Demarque, P., Kim, Y.-C., & Sills, A. 2013, ApJ, 776, 87
- Spinello, C., Barnabè, M., Koopmans, L. V. E., & Trager, S. C. 2015, MNRAS, 452, L21
- Stapelfeldt, K. R., Ménard, F., Watson, A. M., et al. 2003, ApJ, 589, 410
- Stassun, K. G., Kratter, K. M., Scholz, A., & Dupuy, T. J. 2012, ApJ, 756, 47
- Tarter, J. C., Backus, P. R., Mancinelli, R. L., et al. 2007, Astrobiology, 7, 30
- Terrien, R. C., Mahadevan, S., Bender, C. F., et al. 2012, ApJL, 747, L38
- Terrien, R. C., Mahadevan, S., Deshpande, R., & Bender, C. F. 2015, ApJS, 220, 16
- Teske, J. K., Ghezzi, L., Cunha, K., et al. 2015, ApJL, 801, L10
- Tokovinin, A. 2018, ApJS, 235, 6
- Tokovinin, A., Mason, B. D., & Hartkopf, W. I. 2010, AJ, 139, 743
- Tokovinin, A. A., & Smekhov, M. G. 2002, A&A, 382, 118
- Torres, G., Andersen, J., & Giménez, A. 2010, A&A Rv, 18, 67
- Vacca, W. D., Cushing, M. C., & Rayner, J. T. 2003, PASP, 115, 389
- van Altena, W. F., Lee, J. T., & Hoffleit, E. D. 1995, The general catalogue of trigonometric [stellar] parallaxes (Yale University Observatory)
- van Dokkum, P. G. 2001, PASP, 113, 1420
- Van Grootel, V., Fernandes, C. S., Gillon, M., et al. 2018, ApJ, 853, doi:10.3847/1538-4357/aaa023
- van Leeuwen, F. 2007, A&A, 474, 653
- van Leeuwen, F., & Fantino, E. 2005, A&A, 439, 791
- Veyette, M. J., Muirhead, P. S., Mann, A. W., & Allard, F. 2016, ApJ, 828, 95
- Ward-Duong, K., Patience, J., De Rosa, R. J., et al. 2015, MNRAS, 449, 2618
- Wheeler, J. C. 2012, ApJ, 758, 123
- Winters, J. G., Hambly, N. C., Jao, W.-C., et al. 2015, AJ, 149, 5
- Woolf, V. M., & West, A. A. 2012, MNRAS, 422, 1489
- Yelda, S., Lu, J. R., Ghez, A. M., et al. 2010, ApJ, 725, 331
- Yi, Z.-P., Luo, A.-L., Zhao, J.-K., et al. 2015, Research in Astronomy and Astrophysics, 15, 860
- Yuk, I.-S., Jaffe, D. T., Barnes, S., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

**Table 3.** Binary Sample

Name	Comp	R.A.	Dec	$K_S$ (mag)	$\Delta K_S$ (mag)	System $M_*$ $M_\odot$	[Fe/H] <sup>a</sup>	Plx	Plx Ref
Systems analyzed in this paper									
GJ 1005	AB	00:15:28	-16:08:01	6.310±0.018 <sup>b</sup>	2.501±0.040	0.319±0.002	-0.41	166.60±0.30	3
GJ 2005	BC	00:24:44	-27:08:24	8.241±0.030	0.931±0.016	0.157±0.001	-0.08	128.5±1.5	3
GJ 22	AC	00:32:29	+67:14:08	6.037±0.023	1.493±0.018	0.569±0.011	-0.24	99.20±0.60	3
Gl 54	AB	01:10:22	-67:26:42	5.132±0.024	0.254±0.016	0.750±0.010	+0.17	126.90±0.40	3
GJ 1038	AB	01:25:01	-32:51:04	6.207±0.021	0.994±0.017	1.278±0.285	+0.03	39.8±1.6	2
GJ 65	AB	01:39:01	-17:57:02	5.343±0.021	0.068±0.015	0.237±0.005	+0.04	373.7±2.7	5
Gl 84	AB	02:05:04	-17:36:52	5.662±0.020	0.108±0.016	0.514±0.027	-0.14	109.6±1.9	2,5
2M0213+36	AB	02:13:20	+36:48:50	8.518±0.018	3.232±0.016	0.245±0.036	-0.07	74.6±3.5	6
HD 15285	AB	02:27:45	+04:25:55	5.113±0.018	1.171±0.016	1.389±0.079	+0.41	58.3±1.1	2
GJ 99	AB	02:28:46	+32:15:34	6.062±0.018	1.083±0.016	1.456±0.188	+0.21	38.2±1.6	2
Gl 125	AB	03:09:30	+45:43:58	5.840±0.016	0.098±0.016	0.905±0.082	+0.49	63.5±1.9	2
Gl 190	AB	05:08:35	-18:10:18	5.314±0.018	0.287±0.016	0.835±0.054	+0.25	107.9±2.1	2
GJ 1081	AB	05:33:19	+44:48:58	7.335±0.027	0.033±0.017	0.580±0.043	-0.03	65.20±0.40	3
GJ 234	AB	06:29:23	-02:48:49	5.421±0.018 <sup>b</sup>	0.697±0.036	0.333±0.002	+0.11	241.00±0.40	3
GJ 3412	AB	06:54:04	+60:52:18	6.371±0.018 <sup>b</sup>	4.269±0.017	0.528±0.007	-0.21	95.60±0.30	3
GJ 3421	AB	07:03:55	+52:42:06	7.776±0.021	0.001±0.018	0.248±0.009	-0.60	112.6±1.4	3
Gl 263	AB	07:04:17	-10:30:31	6.427±0.017	1.100±0.016	0.951±0.144	+0.43	62.0±2.9	2,5
Gl 277	AaAb	07:31:57	+36:13:09	5.927±0.017	0.811±0.016	0.730±0.028	+0.18	83.327±0.069	12
GJ 3454	AB	07:36:25	+07:04:43	7.282±0.020	1.086±0.016	0.280±0.015	+0.03	117.1±2.1	1
Gl 301	AB	08:13:08	-13:55:01	5.815±0.026	2.134±0.017	1.227±0.146	-0.18	52.1±2.0	5,8
Gl 310	AB	08:36:25	+67:17:42	5.580±0.015	0.506±0.018	0.776±0.067	-0.06	72.6±1.3	2
Gl 330	AB	08:57:04	+11:38:49	6.486±0.023	0.103±0.017	0.913±0.146	-0.17	58.8±2.7	2
LHS 6167	AB	09:15:36	-10:35:47	7.733±0.017	1.140±0.016	0.281±0.009	-0.13	103.3±1.0	10
GJ 340	AB	09:17:53	+28:33:38	4.767±0.017	0.251±0.020	1.357±0.055	-0.39	57.92±0.76	2
GJ 352	AB	09:31:19	-13:29:19	5.511±0.020	0.161±0.019	0.780±0.075	-0.03	98.5±3.1	5,9
Gl 381	AB	10:12:04	-02:41:04	6.193±0.026	0.029±0.018	0.765±0.083	-0.12	83.5±2.8	2,5
GJ 416	AB	11:11:33	-14:59:28	5.874±0.017	0.995±0.046	1.192±0.103	-0.20	46.4±1.3	2
Gl 469	AB	12:28:57	+08:25:31	6.956±0.026	0.913±0.016	0.504±0.010	+0.19	76.40±0.50	3
Gl 473	AB	12:33:17	+09:01:15	6.042±0.020	1.018±0.016	0.262±0.016	+0.09	227.9±4.6	5
Gl 494	AB	13:00:46	+12:22:32	5.578±0.016	0.119±0.018	0.667±0.036	+0.30	85.5±1.5	2
GJ 570	BC	14:57:26	-21:24:40	3.855±0.045 <sup>b</sup>	0.995±0.046	0.833±0.053	+0.18	171.22±0.94	2
GJ 600	AB	15:52:08	+10:52:28	5.819±0.023	1.120±0.016	1.228±0.032	+0.01	47.65±0.30	11
GJ 623	AB	16:24:09	+48:21:10	5.915±0.023	0.024±0.020	0.494±0.004	-0.44	125.00±0.30	3
GJ 1210	AB	17:07:40	+07:22:06	8.419±0.025	1.526±0.016	0.218±0.019	-0.19	86.8±2.5	1
Gl 660	AB	17:11:52	-01:51:06	6.656±0.021	2.060±0.016	0.607±0.053	-0.20	85.2±2.5	5,9
GJ 661	AB	17:12:07	+45:39:58	4.834±0.023	0.320±0.016	0.596±0.054	-0.41	167.3±5.0	2
GJ 4024	AB	17:43:00	+05:47:21	6.631±0.018	1.124±0.016	0.916±0.144	-0.09	50.1±2.5	2

*Table 3 continued*

**Table 3** (*continued*)

Name	Comp	R.A.	Dec	$K_S$	$\Delta K_S$	System $M_*$	[Fe/H] <sup>a</sup>	Plx	Plx Ref
				(mag)	(mag)	$M_\odot$		mas	
Gl 695	BC	17:46:25	+27:43:00	4.953±0.016	0.949±0.018	0.868±0.023	+0.22	120.33±0.16	2
Gl 747	AB	19:07:42	+32:32:41	6.416±0.020	1.145±0.016	0.429±0.027	-0.20	122.3±2.5	5
GJ 748	AB	19:12:14	+02:53:11	6.342±0.018 <sup>b</sup>	1.565±0.018	0.526±0.007	-0.23	98.40±0.30	3
HIP 95995	AB	19:31:08	+58:35:13	4.463±0.015	1.035±0.024	1.468±0.067	-0.17	58.37±0.54	2
HIP 96656	AB	19:39:05	+76:25:18	5.898±0.016	1.327±0.018	1.330±0.068	-0.35	33.67±0.53	2
GJ 1245	AC	19:53:54	+44:24:53	6.854±0.016	0.058±0.016	0.204±0.001	+0.01	214.529±0.082	12
Gl 791.2	AB	20:29:48	+09:41:20	7.307±0.024	1.931±0.016	0.291±0.023	+0.05	113.40±0.20	3
Gl 804	AB	20:44:21	+19:44:59	6.553±0.016	0.272±0.016	0.996±0.138	-0.09	47.4±2.1	2
Gl 831	AB	21:31:18	-09:47:26	6.379±0.020	0.075±0.017	0.418±0.003	+0.20	125.30±0.30	3
GJ 4210	AB	21:39:54	+27:36:43	7.376±0.024	0.289±0.049	0.998±0.237	-0.25	40.9±3.0	2
Gl 844	AB	22:01:49	+16:28:02	6.180±0.021	0.281±0.060	0.871±0.096	+0.21	61.8±2.2	2
HD 239960	AB	22:27:59	+57:41:45	4.777±0.029	0.067±0.016	0.460±0.011	-0.11	249.9±1.9	2
GJ 4287	AB	22:37:29	+39:22:51	5.872±0.027	0.094±0.003	0.954±0.091	+0.03	51.2±1.6	2
Gl 893.4	AB	23:16:39	+19:37:17	7.303±0.017	0.566±0.029	0.845±0.121	-0.17	32.5±1.5	1
Gl 913	AB	23:58:43	+46:43:44	5.831±0.020	0.490±0.020	0.890±0.121	-0.12	68.0±3.0	7
Systems from Dupuy & Liu (2017)									
LP415-20	AB	04:21:49	+19:29:08	11.668±0.020	0.357±0.025	0.240±0.028	+0.10 <sup>c</sup>	25.76±0.77	4
LHS1901	AB	07:11:11	+43:29:58	9.126±0.018	0.113±0.024	0.203±0.009	-0.41	76.4±1.1	4
2M0746+20	AB	07:46:42	+20:00:32	10.468±0.022	0.743±0.075	0.153±0.002	-0.18	81.24±0.25	4
2M1017+13	AB	10:17:07	+13:08:39	12.710±0.023	0.044±0.016	0.149±0.017	-0.35 <sup>d</sup>	32.2±1.2	4
2M1047+40	AB	10:47:13	+40:26:49	11.254±0.018	0.441±0.020	0.171±0.012	+0.17	39.02±0.90	4
LSJ1735+26	AB	17:35:12	+26:34:47	10.157±0.020	0.980±0.017	0.178±0.008	+0.03	64.82±0.95	4
2M1847+55	AB	18:47:03	+55:22:43	10.901±0.020	0.371±0.037	0.185±0.023	+0.32	30.5±1.2	4
2M2140+16	AB	21:40:29	+16:25:18	11.826±0.031	0.080±0.121	0.190±0.017	-0.10 <sup>c</sup>	29.52±0.88	4
2M2206-20	AB	22:06:22	-20:47:05	11.315±0.027	0.143±0.016	0.168±0.017	-0.05 <sup>c</sup>	36.6±1.2	4

NOTE—Parallax references: 1 = This work (MEarth), 2 = van Leeuwen & Fantino (2005), 3 = Benedict et al. (2016), 4 = Dupuy & Liu (2017), 5 = van Altena et al. (1995), 6 = Finch & Zacharias (2016), 7 = Goldin & Makarov (2006), 8 = Winters et al. (2015), 9 = Söderhjelm (1999), 10 = Bartlett et al. (2017), 11 = Gaia Collaboration et al. (2016), 12 = companion to star in Lindegren et al. (2018).

<sup>a</sup>Errors on [Fe/H] are limited primarily by the calibration (Mann et al. 2013a, 2014) or the overall [Fe/H] scale, and are 0.08 dex for all targets unless otherwise noted.

<sup>b</sup>Synthetic  $K_S$  magnitudes derived from spectra. All other  $K_S$  magnitudes are from 2MASS.

<sup>c</sup>Abundance derived from lower-resolution IRTF spectrum,  $\sigma$ [Fe/H] estimated to be 0.12 dex.

<sup>d</sup>L dwarfs are beyond the calibration range of Mann et al. (2014); [Fe/H] values should be used with caution.

Table 4. Orbital Parameters

Name	$P$ (years)	$a''$ (mas)	$e$	$i$ (degrees)	$\omega$ (degrees)	$\Omega$ (degrees)	$T_0$ MJD	$\chi^2$	dof
GJ 1005	4.55725 <sup>+0.00077</sup> <sub>-0.00077</sub>	312.90 <sup>+0.50</sup> <sub>-0.50</sub>	0.36123 <sup>+0.00097</sup> <sub>-0.00097</sub>	143.89 <sup>+0.25</sup> <sub>-0.25</sub>	345.21 <sup>+0.63</sup> <sub>-0.62</sub>	61.20 <sup>+0.41</sup> <sub>-0.40</sub>	58172.8 <sup>+1.9</sup> <sub>-1.9</sub>	88.5	71
GJ 2005	17.280 <sup>+0.010</sup> <sub>-0.010</sub>	463.34 <sup>+0.54</sup> <sub>-0.54</sub>	0.03029 <sup>+0.00090</sup> <sub>-0.00087</sub>	62.829 <sup>+0.049</sup> <sub>-0.049</sub>	141.5 <sup>+2.1</sup> <sub>-2.0</sub>	11.890 <sup>+0.079</sup> <sub>-0.079</sub>	59124 <sup>+38</sup> <sub>-36</sub>	90.9	79
GJ 22	15.4314 <sup>+0.0064</sup> <sub>-0.0063</sub>	509.50 <sup>+0.76</sup> <sub>-0.76</sub>	0.1577 <sup>+0.014</sup> <sub>-0.014</sub>	44.20 <sup>+0.16</sup> <sub>-0.16</sub>	104.75 <sup>+0.54</sup> <sub>-0.53</sub>	176.97 <sup>+0.21</sup> <sub>-0.21</sub>	57447.0 <sup>+5.4</sup> <sub>-5.4</sub>	84.8	75
Gl 54	1.14434 <sup>+0.00021</sup> <sub>-0.00021</sub>	126.14 <sup>+0.38</sup> <sub>-0.38</sub>	0.11712 <sup>+0.0023</sup> <sub>-0.0024</sub>	125.38 <sup>+0.35</sup> <sub>-0.34</sub>	47.06 <sup>+0.90</sup> <sub>-0.91</sub>	91.99 <sup>+0.39</sup> <sub>-0.39</sub>	55615.89 <sup>+0.87</sup> <sub>-0.88</sub>	44.6	39
GJ 1038	5.63 <sup>+1.42</sup> <sub>-0.90</sub>	135.8 <sup>+18.7</sup> <sub>-7.1</sub>	0.59 <sup>+0.16</sup> <sub>-0.15</sub>	72.1 <sup>+2.6</sup> <sub>-3.2</sub>	180 <sup>+19</sup> <sub>-22</sub>	104.9 <sup>+2.6</sup> <sub>-5.4</sub>	56634 <sup>+1712</sup> <sub>-271</sub>	0.8	7
GJ 65	26.368 <sup>+0.021</sup> <sub>-0.021</sub>	2049.4 <sup>+3.3</sup> <sub>-3.3</sub>	0.6204 <sup>+0.0021</sup> <sub>-0.0021</sub>	128.13 <sup>+0.15</sup> <sub>-0.15</sub>	283.350 <sup>+0.097</sup> <sub>-0.096</sub>	146.30 <sup>+0.15</sup> <sub>-0.15</sub>	70227 <sup>+15</sup> <sub>-15</sub>	78.6	97
Gl 84	13.392 <sup>+0.028</sup> <sub>-0.027</sub>	495.2 <sup>+1.1</sup> <sub>-1.1</sub>	0.3863 <sup>+0.0073</sup> <sub>-0.0072</sub>	91.797 <sup>+0.058</sup> <sub>-0.058</sub>	245.30 <sup>+0.44</sup> <sub>-0.43</sub>	102.991 <sup>+0.023</sup> <sub>-0.023</sub>	61579 <sup>+11</sup> <sub>-10</sub>	26.7	21
2M0213+36	6.441 <sup>+0.070</sup> <sub>-0.073</sub>	161.6 <sup>+1.3</sup> <sub>-1.3</sub>	0.4240 <sup>+0.0040</sup> <sub>-0.0041</sub>	115.37 <sup>+0.24</sup> <sub>-0.24</sub>	207.42 <sup>+0.75</sup> <sub>-0.74</sub>	83.82 <sup>+0.33</sup> <sub>-0.33</sub>	57603.5 <sup>+3.0</sup> <sub>-2.9</sub>	22.4	11
HD 15285	25.126 <sup>+0.030</sup> <sub>-0.030</sub>	558.01 <sup>+0.75</sup> <sub>-0.75</sub>	0.2327 <sup>+0.0013</sup> <sub>-0.0013</sub>	73.374 <sup>+0.048</sup> <sub>-0.048</sub>	231.89 <sup>+0.24</sup> <sub>-0.24</sub>	109.054 <sup>+0.025</sup> <sub>-0.025</sub>	65660 <sup>+10</sup> <sub>-10</sub>	75.5	87
GJ 99	24.023 <sup>+0.089</sup> <sub>-0.088</sub>	360.54 <sup>+0.74</sup> <sub>-0.72</sub>	0.2084 <sup>+0.0028</sup> <sub>-0.0026</sub>	84.618 <sup>+0.051</sup> <sub>-0.051</sub>	152.9 <sup>+2.0</sup> <sub>-2.0</sub>	98.863 <sup>+0.090</sup> <sub>-0.091</sub>	65119 <sup>+36</sup> <sub>-35</sub>	21.3	23
Gl 125	25.67 <sup>+0.20</sup> <sub>-0.19</sub>	534.5 <sup>+2.3</sup> <sub>-2.3</sub>	0.2271 <sup>+0.0044</sup> <sub>-0.0044</sub>	97.186 <sup>+0.026</sup> <sub>-0.026</sub>	181.38 <sup>+0.35</sup> <sub>-0.36</sub>	13.732 <sup>+0.082</sup> <sub>-0.081</sub>	64226 <sup>+74</sup> <sub>-73</sub>	14.1	23
Gl 190	0.96384 <sup>+0.00025</sup> <sub>-0.00025</sub>	99.14 <sup>+0.86</sup> <sub>-0.86</sub>	0.2439 <sup>+0.0094</sup> <sub>-0.0091</sub>	92.97 <sup>+0.25</sup> <sub>-0.25</sub>	186.5 <sup>+5.8</sup> <sub>-5.8</sub>	40.42 <sup>+0.17</sup> <sub>-0.17</sub>	55718.3 <sup>+6.4</sup> <sub>-6.4</sub>	33.6	29
GJ 1081	11.622 <sup>+0.033</sup> <sub>-0.033</sub>	279.0 <sup>+7.4</sup> <sub>-6.7</sub>	0.8648 <sup>+0.0066</sup> <sub>-0.0064</sub>	97.06 <sup>+0.51</sup> <sub>-0.50</sub>	231.9 <sup>+1.5</sup> <sub>-1.5</sub>	51.23 <sup>+0.25</sup> <sub>-0.25</sub>	61481 <sup>+29</sup> <sub>-29</sub>	14.8	19
GJ 234	16.5798 <sup>+0.0027</sup> <sub>-0.0027</sub>	1086.28 <sup>+0.26</sup> <sub>-0.26</sub>	0.38236 <sup>+0.00013</sup> <sub>-0.00013</sub>	52.919 <sup>+0.015</sup> <sub>-0.015</sub>	220.942 <sup>+0.020</sup> <sub>-0.020</sub>	30.384 <sup>+0.030</sup> <sub>-0.030</sub>	63398.76 <sup>+0.93</sup> <sub>-0.93</sub>	114.4	109
GJ 3412	13.5946 <sup>+0.0058</sup> <sub>-0.0058</sub>	440.1 <sup>+1.4</sup> <sub>-1.4</sub>	0.4768 <sup>+0.0022</sup> <sub>-0.0022</sub>	109.80 <sup>+0.12</sup> <sub>-0.12</sub>	58.47 <sup>+0.21</sup> <sub>-0.21</sub>	107.16 <sup>+0.18</sup> <sub>-0.18</sub>	59670.5 <sup>+2.8</sup> <sub>-2.8</sub>	40.9	51
GJ 3421	3.2857 <sup>+0.0011</sup> <sub>-0.0011</sub>	156.39 <sup>+0.25</sup> <sub>-0.25</sub>	0.2245 <sup>+0.0035</sup> <sub>-0.0034</sub>	131.62 <sup>+0.41</sup> <sub>-0.41</sub>	73.72 <sup>+0.36</sup> <sub>-0.36</sub>	174.19 <sup>+0.60</sup> <sub>-0.62</sub>	56133.3 <sup>+3.1</sup> <sub>-3.1</sub>	29.3	35
Gl 263	3.6205 <sup>+0.0021</sup> <sub>-0.0021</sub>	143.8 <sup>+2.1</sup> <sub>-2.0</sub>	0.7158 <sup>+0.0065</sup> <sub>-0.0064</sub>	103.28 <sup>+0.36</sup> <sub>-0.36</sub>	287.52 <sup>+0.51</sup> <sub>-0.51</sub>	81.04 <sup>+0.24</sup> <sub>-0.24</sub>	57093.6 <sup>+3.4</sup> <sub>-3.3</sub>	21.7	17
Gl 277	53.0 <sup>+9.2</sup> <sub>-7.2</sub>	1058 <sup>+11.3</sup> <sub>-84</sub>	0.48 <sup>+0.14</sup> <sub>-0.12</sub>	93.53 <sup>+0.48</sup> <sub>-0.34</sub>	22 <sup>+332</sup> <sub>-17</sub>	10.22 <sup>+0.26</sup> <sub>-0.44</sub>	71033 <sup>+12032</sup> <sub>-3397</sub>	18.4	13
GJ 3454	23.767 <sup>+0.045</sup> <sub>-0.045</sub>	632.90 <sup>+0.96</sup> <sub>-0.95</sub>	0.58648 <sup>+0.00061</sup> <sub>-0.00062</sub>	11.7 <sup>+1.1</sup> <sub>-1.2</sub>	66.8 <sup>+5.8</sup> <sub>-5.5</sub>	77.3 <sup>+5.2</sup> <sub>-5.4</sub>	57467.0 <sup>+2.6</sup> <sub>-2.6</sub>	14.7	19
Gl 301	62.2 <sup>+1.8</sup> <sub>-1.7</sub>	875 <sup>+10</sup> <sub>-10</sub>	0.67778 <sup>+0.0049</sup> <sub>-0.0048</sub>	52.31 <sup>+0.86</sup> <sub>-0.87</sub>	167.4 <sup>+1.1</sup> <sub>-1.1</sub>	142.0 <sup>+1.1</sup> <sub>-1.1</sub>	73905 <sup>+675</sup> <sub>-630</sub>	22.5	13
Gl 310	23.55 <sup>+0.42</sup> <sub>-0.38</sub>	548.3 <sup>+9.0</sup> <sub>-8.3</sub>	0.692 <sup>+0.012</sup> <sub>-0.012</sub>	122.50 <sup>+0.92</sup> <sub>-0.93</sub>	246.49 <sup>+0.71</sup> <sub>-0.73</sub>	49.95 <sup>+0.77</sup> <sub>-0.74</sub>	67038 <sup>+157</sup> <sub>-145</sub>	9.4	9
Gl 330	32.69 <sup>+0.42</sup> <sub>-0.40</sub>	582 <sup>+14</sup> <sub>-12</sub>	0.8301 <sup>+0.0071</sup> <sub>-0.0070</sub>	105.78 <sup>+0.51</sup> <sub>-0.49</sub>	309.0 <sup>+1.5</sup> <sub>-1.5</sub>	38.63 <sup>+0.50</sup> <sub>-0.51</sub>	76604 <sup>+309</sup> <sub>-301</sub>	23.0	17
Gl 340	34.23 <sup>+0.13</sup> <sub>-0.13</sub>	676.2 <sup>+1.2</sup> <sub>-1.3</sub>	0.3100 <sup>+0.0035</sup> <sub>-0.0035</sub>	77.02 <sup>+0.30</sup> <sub>-0.31</sub>	312.05 <sup>+0.79</sup> <sub>-0.78</sub>	25.23 <sup>+0.25</sup> <sub>-0.25</sub>	69557 <sup>+54</sup> <sub>-53</sub>	77.1	67
GJ 352	18.480 <sup>+0.026</sup> <sub>-0.026</sub>	633.7 <sup>+2.6</sup> <sub>-2.5</sub>	0.3322 <sup>+0.0026</sup> <sub>-0.0026</sub>	143.07 <sup>+0.72</sup> <sub>-0.70</sub>	288.37 <sup>+0.77</sup> <sub>-0.75</sub>	51.57 <sup>+0.45</sup> <sub>-0.46</sub>	65844 <sup>+23</sup> <sub>-23</sub>	24.0	15
Gl 381	7.5711 <sup>+0.0057</sup> <sub>-0.0058</sub>	294.4 <sup>+3.9</sup> <sub>-3.5</sub>	0.75328 <sup>+0.0066</sup> <sub>-0.0069</sub>	92.996 <sup>+0.079</sup> <sub>-0.082</sub>	273.95 <sup>+0.12</sup> <sub>-0.13</sub>	68.546 <sup>+0.052</sup> <sub>-0.051</sub>	57698.9 <sup>+4.3</sup> <sub>-4.5</sub>	33.8	33
GJ 416	7.2716 <sup>+0.00778</sup> <sub>-0.00778</sub>	184.46 <sup>+0.66</sup> <sub>-0.59</sub>	0.4571 <sup>+0.0037</sup> <sub>-0.0037</sub>	87.25 <sup>+0.24</sup> <sub>-0.24</sub>	154.3 <sup>+1.1</sup> <sub>-1.2</sub>	89.08 <sup>+0.11</sup> <sub>-0.11</sub>	58731 <sup>+11</sup> <sub>-11</sub>	34.4	23
Gl 469	11.5617 <sup>+0.0051</sup> <sub>-0.0051</sub>	310.89 <sup>+0.62</sup> <sub>-0.61</sub>	0.29986 <sup>+0.0030</sup> <sub>-0.0030</sub>	108.083 <sup>+0.063</sup> <sub>-0.063</sub>	268.67 <sup>+0.12</sup> <sub>-0.12</sub>	9.84 <sup>+0.12</sup> <sub>-0.12</sub>	58771.7 <sup>+3.1</sup> <sub>-3.1</sub>	37.2	37
Gl 473	15.851 <sup>+0.017</sup> <sub>-0.017</sub>	919.9 <sup>+11</sup> <sub>-11</sub>	0.2976 <sup>+0.0010</sup> <sub>-0.0010</sub>	103.05 <sup>+0.12</sup> <sub>-0.12</sub>	349.88 <sup>+0.65</sup> <sub>-0.65</sub>	143.12 <sup>+0.18</sup> <sub>-0.18</sub>	66130 <sup>+14</sup> <sub>-14</sub>	50.0	49

Table 4 continued

Table 4 (continued)

Name	<i>P</i> (years)	<i>a''</i> (mas)	<i>e</i>	<i>i</i> (degrees)	$\omega$ (degrees)	$\Omega$ (degrees)	<i>T</i> <sub>0</sub> MJD	$\chi^2$	dof
GJ 494	13.713 <sup>+0.033</sup> <sub>-0.033</sub>	428.09 <sup>+0.40</sup> <sub>-0.41</sub>	0.24260 <sup>+0.00083</sup> <sub>-0.00083</sub>	130.66 <sup>+0.19</sup> <sub>-0.19</sub>	158.70 <sup>+0.57</sup> <sub>-0.57</sub>	56.20 <sup>+0.16</sup> <sub>-0.16</sub>	59362 <sup>+17</sup> <sub>-17</sub>	33.1	23
GJ 570	0.84567 <sup>+0.00035</sup> <sub>-0.00035</sub>	144.1 <sup>+3.1</sup> <sub>-2.7</sub>	0.752 <sup>+0.15</sup> <sub>-0.14</sub>	107.43 <sup>+0.70</sup> <sub>-0.69</sub>	310.5 <sup>+1.8</sup> <sub>-1.8</sub>	16.30 <sup>+0.72</sup> <sub>-0.74</sub>	55524.63 <sup>+0.88</sup> <sub>-0.93</sub>	26.1	37
GJ 600	2.7808 <sup>+0.0012</sup> <sub>-0.0012</sub>	100.91 <sup>+0.58</sup> <sub>-0.57</sub>	0.3396 <sup>+0.0040</sup> <sub>-0.0039</sub>	34.49 <sup>+0.83</sup> <sub>-0.85</sub>	151.7 <sup>+1.4</sup> <sub>-1.4</sub>	115.8 <sup>+1.3</sup> <sub>-1.2</sub>	55896.6 <sup>+2.3</sup> <sub>-2.4</sub>	18.2	27
GJ 623	3.7372 <sup>+0.0020</sup> <sub>-0.0020</sub>	237.93 <sup>+0.47</sup> <sub>-0.46</sub>	0.6256 <sup>+0.0020</sup> <sub>-0.0020</sub>	153.40 <sup>+0.40</sup> <sub>-0.39</sub>	247.88 <sup>+0.63</sup> <sub>-0.64</sub>	98.69 <sup>+0.74</sup> <sub>-0.74</sub>	56774.6 <sup>+1.5</sup> <sub>-1.5</sub>	29.2	39
GJ 1210	14.298 <sup>+0.015</sup> <sub>-0.015</sub>	307.70 <sup>+0.24</sup> <sub>-0.23</sub>	0.48392 <sup>+0.00032</sup> <sub>-0.00032</sub>	112.351 <sup>+0.046</sup> <sub>-0.045</sub>	23.03 <sup>+0.19</sup> <sub>-0.19</sub>	60.598 <sup>+0.070</sup> <sub>-0.069</sub>	59161.7 <sup>+5.8</sup> <sub>-5.7</sub>	12.0	19
GJ 660	34.537 <sup>+0.082</sup> <sub>-0.082</sub>	765.1 <sup>+1.5</sup> <sub>-1.4</sub>	0.1990 <sup>+0.0054</sup> <sub>-0.0054</sub>	18.82 <sup>+0.69</sup> <sub>-0.68</sub>	218.1 <sup>+4.8</sup> <sub>-5.1</sub>	144.1 <sup>+6.0</sup> <sub>-5.7</sub>	72411 <sup>+68</sup> <sub>-65</sub>	9.6	9
GJ 661	12.9586 <sup>+0.0058</sup> <sub>-0.0059</sub>	776.6 <sup>+2.2</sup> <sub>-2.2</sub>	0.7508 <sup>+0.0029</sup> <sub>-0.0029</sub>	146.63 <sup>+0.36</sup> <sub>-0.36</sub>	98.3 <sup>+1.0</sup> <sub>-1.0</sub>	159.3 <sup>+1.2</sup> <sub>-1.2</sub>	57731.3 <sup>+2.5</sup> <sub>-2.5</sub>	44.9	67
GJ 4024	20.26 <sup>+0.34</sup> <sub>-0.30</sub>	361.9 <sup>+5.0</sup> <sub>-4.8</sub>	0.533 <sup>+0.013</sup> <sub>-0.013</sub>	104.70 <sup>+0.28</sup> <sub>-0.27</sub>	232.9 <sup>+1.5</sup> <sub>-1.5</sub>	129.06 <sup>+0.33</sup> <sub>-0.32</sub>	61648 <sup>+119</sup> <sub>-107</sub>	3.0	9
GJ 695	43.23 <sup>+0.28</sup> <sub>-0.27</sub>	1414 <sup>+10</sup> <sub>-10</sub>	0.1806 <sup>+0.0039</sup> <sub>-0.0038</sub>	66.08 <sup>+0.20</sup> <sub>-0.20</sub>	172.05 <sup>+0.54</sup> <sub>-0.52</sub>	60.93 <sup>+0.84</sup> <sub>-0.83</sub>	70376 <sup>+110</sup> <sub>-106</sub>	4.5	21
GJ 747	5.75030 <sup>+0.00079</sup> <sub>-0.00079</sub>	296.11 <sup>+0.69</sup> <sub>-0.68</sub>	0.2573 <sup>+0.0011</sup> <sub>-0.0010</sub>	77.623 <sup>+0.063</sup> <sub>-0.063</sub>	330.30 <sup>+0.72</sup> <sub>-0.72</sub>	84.639 <sup>+0.069</sup> <sub>-0.068</sub>	58827.9 <sup>+4.4</sup> <sub>-4.4</sub>	86.7	75
GJ 748	2.47604 <sup>+0.00028</sup> <sub>-0.00028</sub>	145.38 <sup>+0.46</sup> <sub>-0.46</sub>	0.4590 <sup>+0.0017</sup> <sub>-0.0017</sub>	131.77 <sup>+0.39</sup> <sub>-0.39</sub>	23.58 <sup>+0.60</sup> <sub>-0.61</sub>	176.83 <sup>+0.37</sup> <sub>-0.37</sub>	55457.06 <sup>+0.80</sup> <sub>-0.80</sub>	63.3	71
HIP95995	1.35257 <sup>+0.00046</sup> <sub>-0.00046</sub>	81.14 <sup>+0.98</sup> <sub>-0.97</sub>	0.4037 <sup>+0.0075</sup> <sub>-0.0073</sub>	148.6 <sup>+2.3</sup> <sub>-2.1</sub>	2.5 <sup>+3.3</sup> <sub>-3.6</sub>	68.2 <sup>+2.8</sup> <sub>-3.1</sub>	55558.1 <sup>+491.1</sup> <sub>-2.4</sub>	33.0	59
HIP96656	11.914 <sup>+0.026</sup> <sub>-0.026</sub>	193.2 <sup>+1.1</sup> <sub>-1.1</sub>	0.2419 <sup>+0.0088</sup> <sub>-0.0088</sub>	80.64 <sup>+0.27</sup> <sub>-0.27</sub>	70.33 <sup>+0.75</sup> <sub>-0.79</sub>	112.92 <sup>+0.42</sup> <sub>-0.42</sub>	57753 <sup>+16</sup> <sub>-15</sub>	25.4	33
GJ 1245	16.881 <sup>+0.016</sup> <sub>-0.016</sub>	830.54 <sup>+0.26</sup> <sub>-0.26</sub>	0.34251 <sup>+0.00083</sup> <sub>-0.00084</sub>	135.29 <sup>+0.14</sup> <sub>-0.14</sub>	217.16 <sup>+0.28</sup> <sub>-0.29</sub>	80.65 <sup>+0.20</sup> <sub>-0.20</sub>	63861.0 <sup>+7.1</sup> <sub>-7.1</sub>	28.9	49
GJ 791.2	1.47118 <sup>+0.00039</sup> <sub>-0.00042</sub>	97.2 <sup>+3.2</sup> <sub>-1.4</sub>	0.6302 <sup>+0.0084</sup> <sub>-0.0119</sub>	163.0 <sup>+9.0</sup> <sub>-11.1</sub>	87 <sup>+136</sup> <sub>-54</sub>	272 <sup>+34</sup> <sub>-153</sub>	55846 <sup>+11</sup> <sub>-528</sub>	53.5	37
GJ 804	14.709 <sup>+0.018</sup> <sub>-0.018</sub>	284.1 <sup>+2.3</sup> <sub>-2.2</sub>	0.6012 <sup>+0.0055</sup> <sub>-0.0054</sub>	129.51 <sup>+0.53</sup> <sub>-0.55</sub>	246.76 <sup>+0.28</sup> <sub>-0.27</sub>	79.3 <sup>+1.0</sup> <sub>-1.0</sub>	63914 <sup>+13</sup> <sub>-13</sub>	35.0	27
GJ 831	1.93195 <sup>+0.00013</sup> <sub>-0.00013</sub>	145.39 <sup>+0.16</sup> <sub>-0.16</sub>	0.38947 <sup>+0.00090</sup> <sub>-0.00090</sub>	49.61 <sup>+0.15</sup> <sub>-0.15</sub>	190.05 <sup>+0.42</sup> <sub>-0.42</sub>	144.28 <sup>+0.21</sup> <sub>-0.21</sub>	56100.40 <sup>+0.54</sup> <sub>-0.55</sub>	81.9	79
GJ 4210	12.776 <sup>+0.021</sup> <sub>-0.021</sub>	223.4 <sup>+1.3</sup> <sub>-1.2</sub>	0.3047 <sup>+0.0086</sup> <sub>-0.0086</sub>	69.95 <sup>+0.21</sup> <sub>-0.21</sub>	106.11 <sup>+0.97</sup> <sub>-0.97</sub>	149.79 <sup>+0.22</sup> <sub>-0.22</sub>	58699 <sup>+16</sup> <sub>-16</sub>	32.6	33
GJ 844	11.812 <sup>+0.045</sup> <sub>-0.044</sub>	306.7 <sup>+2.0</sup> <sub>-2.0</sub>	0.4850 <sup>+0.0052</sup> <sub>-0.0052</sub>	38.10 <sup>+0.78</sup> <sub>-0.81</sub>	267.31 <sup>+0.68</sup> <sub>-0.71</sub>	151.53 <sup>+0.76</sup> <sub>-0.77</sub>	61346 <sup>+16</sup> <sub>-16</sub>	4.5	3
HD 239960	45.35 <sup>+0.29</sup> <sub>-0.29</sub>	2453 <sup>+10</sup> <sub>-10</sub>	0.4147 <sup>+0.0021</sup> <sub>-0.0020</sub>	165.63 <sup>+0.51</sup> <sub>-0.51</sub>	209.2 <sup>+2.4</sup> <sub>-2.5</sub>	153.5 <sup>+2.7</sup> <sub>-2.9</sub>	73509 <sup>+100</sup> <sub>-97</sub>	50.3	49
GJ 4287	17.237 <sup>+0.018</sup> <sub>-0.018</sub>	336.4 <sup>+1.3</sup> <sub>-1.3</sub>	0.2381 <sup>+0.0033</sup> <sub>-0.0033</sub>	55.37 <sup>+0.29</sup> <sub>-0.29</sub>	116.65 <sup>+0.76</sup> <sub>-0.75</sub>	69.59 <sup>+0.20</sup> <sub>-0.19</sub>	61003 <sup>+22</sup> <sub>-22</sub>	12.2	37
GJ 893.4	19.165 <sup>+0.018</sup> <sub>-0.018</sub>	220.10 <sup>+0.52</sup> <sub>-0.52</sub>	0.4400 <sup>+0.0034</sup> <sub>-0.0034</sub>	116.63 <sup>+0.18</sup> <sub>-0.18</sub>	322.35 <sup>+0.63</sup> <sub>-0.62</sub>	28.97 <sup>+0.23</sup> <sub>-0.23</sub>	65451 <sup>+13</sup> <sub>-13</sub>	38.8	23
GJ 913	2.3912 <sup>+0.0012</sup> <sub>-0.0012</sub>	116.95 <sup>+0.28</sup> <sub>-0.27</sub>	0.5143 <sup>+0.0020</sup> <sub>-0.0020</sub>	114.19 <sup>+0.20</sup> <sub>-0.20</sub>	306.62 <sup>+0.24</sup> <sub>-0.23</sub>	112.29 <sup>+0.23</sup> <sub>-0.23</sub>	56666.07 <sup>+0.87</sup> <sub>-0.88</sub>	3.6	5

## APPENDIX

### A. CONVERTING OBSERVED $\Delta K_X$ TO 2MASS $\Delta K_S$ FOR M DWARFS

To place all  $K$ -band magnitudes on the 2MASS system, we derived a relation between  $\Delta K_X$  and  $\Delta K_S$  as a function of  $\Delta K_X$ , where  $X$  denotes the particular filter (e.g.,  $K_{\text{continuum}}$ ,  $K'$ ) used for the AO observations. For photometry, we only used observations taken with a filter somewhere in the  $K$ -band (all wavelengths are used for astrometry), which included observations with the narrow Bracket-Gamma filter.

To derive a conversion between contrasts, we used the 183 absolutely-flux calibrated spectra of nearby single stars from Mann et al. (2015), which cover a similar range of  $T_{\text{eff}}$  and  $M_*$  as the sample considered here. These spectra are mostly empirical; models are only used to fill in gaps in the spectrum, none of which land in the regions covered by the filters considered here.

First we randomly sampled two stars from the sample and scaled the absolute level of each spectrum by the star's distance (effectively placing both at 1 pc). We convolved each of the two stars with the relevant filter profiles for KeckII/NIRC2<sup>10</sup>, CFHT/KIR<sup>11</sup>, Gemini/NIRI<sup>12</sup>, and VLT/NaCo<sup>13</sup>, and integrate over the wavelength range of the filters to compute the flux ( $F$ ) in a given band. We used a filter profile for 2MASS  $K_S$  from Cohen et al. (2003). The  $\Delta K_X$  value for the given pair was then computed as  $2.5 \log_{10}(F_1/F_2)$ .

We repeated this process with 5000 unique combinations of the 183 stars for 12 different filter/instrument combinations. We then computed a best-fit line to  $\Delta K_S - \Delta K_X$  as a function of  $\Delta K_X$ , for each filter and instrument combination. We show four examples in Figure 16. For the majority of the filters considered, the trend is insignificant compared to errors in the underlying spectra and absolute calibration ( $\simeq 0.02$  mag). We did not apply a correction in these cases.

We did not see a significant difference in any derived correction based on the metallicity of the component stars, as expected based on how [Fe/H] changes  $K$ -band flux levels (Figure 3). Thus we did not attempt to include [Fe/H] in these relations. We also found no significant effect as a function of the mass of the primary. However, this was difficult to test due to limitations of the input sample. The Mann et al. (2015) spectral sample covers K7 to M7, but is poorly populated on the extreme ends. While we can make a wide range of combinations of low-contrast systems (M0+M0 to M6+M6), we have limited options for high-contrast (where primary can only be  $\sim$  K7-M1) systems where this effect is largest.

### B. ORBITS OF BINARIES

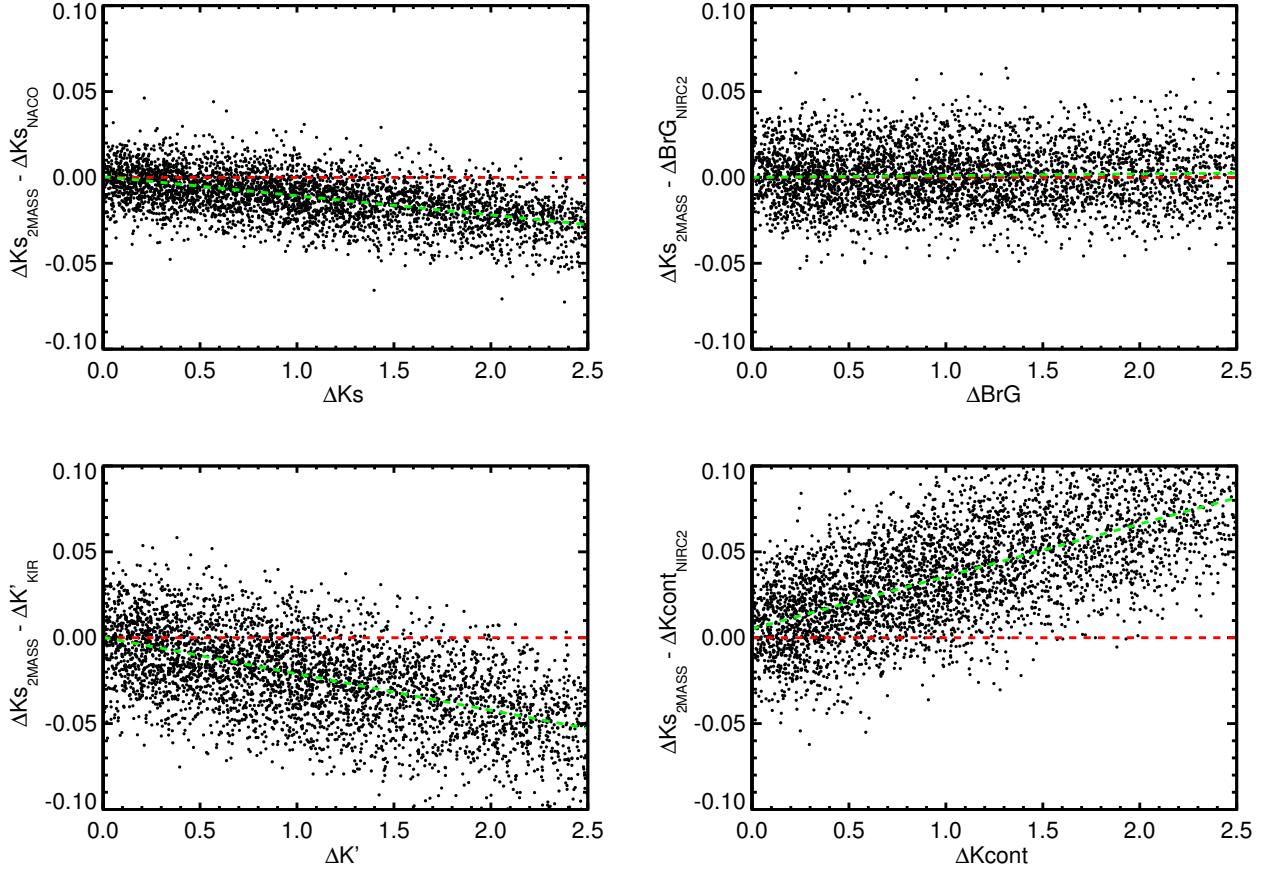
Here we show diagnostic plots of each of the binaries analyzed in this paper. The three binaries shown in Figure 1 are not repeated here. [FOR FILE SIZE REASONS I COMMENTED OUT THE ACTUAL FIGURES FOR NOW]

<sup>10</sup> NIRC2 Filters

<sup>11</sup> KIR Filters

<sup>12</sup> NIRI Filters

<sup>13</sup> NACO Filters



**Figure 16.** Difference between 2MASS  $\Delta K_S$  and four example  $\Delta$  magnitudes measured from our AO imaging, built from a grid of absolutely-calibrated spectra and the filter profiles provided for each filter. No corrections are applied for the NIRC2 Bracket-Gamma (BrG) and  $K_S$  (K-short) filter, as the trend is not significant compared to potential systematic issues in the calibration of the underlying spectra.