



# Close Binary Companions to APOGEE DR16 Stars: 20,000 Binary-star Systems Across the Color–Magnitude Diagram

Adrian M. Price-Whelan<sup>1</sup> , David W. Hogg<sup>1,2,3</sup> , Hans-Walter Rix<sup>3</sup> , Rachael L. Beaton<sup>4,5,26,27</sup> , Hannah M. Lewis<sup>6</sup> , David L. Nidever<sup>7,8</sup> , Andrés Almeida<sup>9</sup>, Carles Badenes<sup>10</sup> , Rodolfo Barba<sup>11</sup> , Timothy C. Beers<sup>12</sup> , Joleen K. Carlberg<sup>13</sup> , Nathan De Lee<sup>14,15</sup> , José G. Fernández-Trincado<sup>16</sup>, Peter M. Frinchaboy<sup>17</sup> , D. A. García-Hernández<sup>18,19</sup> , Paul J. Green<sup>20</sup> , Sten Hasselquist<sup>21,28</sup>, Penélope Longa-Peña<sup>22</sup>, Steven R. Majewski<sup>6</sup> , Christian Nitschelm<sup>22</sup>, Jennifer Sobeck<sup>23</sup>, Keivan G. Stassun<sup>15</sup> , Guy S. Stringfellow<sup>24</sup> , and Nicholas W. Troup<sup>25</sup>

<sup>1</sup> Center for Computational Astrophysics, Flatiron Institute, Simons Foundation, 162 Fifth Avenue, New York, NY 10010, USA; [aprice-whelan@flatironinstitute.org](mailto:aprice-whelan@flatironinstitute.org)

<sup>2</sup> Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA

<sup>3</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

<sup>4</sup> Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA

<sup>5</sup> The Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA

<sup>6</sup> Department of Astronomy, University of Virginia, Charlottesville, VA 22904-4325, USA

<sup>7</sup> Department of Physics, Montana State University, P.O. Box 173840, Bozeman, MT 59717-3840, USA

<sup>8</sup> NSF's National Optical-Infrared Astronomy Research Laboratory, 950 North Cherry Avenue, Tucson, AZ 85719, USA

<sup>9</sup> Instituto de Investigación Multidisciplinario en Ciencia y Tecnología, Universidad de La Serena, Benavente 980, La Serena, Chile

<sup>10</sup> Department of Physics and Astronomy, and Pittsburgh Particle Physics, Astrophysics and Cosmology Center (PITT PACC), University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260, USA

<sup>11</sup> Departamento de Astronomía, Facultad de Ciencias, Universidad de La Serena, Cisternas 1200, La Serena, Chile

<sup>12</sup> Department of Physics and JINA Center for the Evolution of the Elements, University of Notre Dame, Notre Dame, IN 46556, USA

<sup>13</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore MD 21218, USA

<sup>14</sup> Department of Physics, Geology, and Engineering Technology, Northern Kentucky University, Highland Heights, KY 41099, USA

<sup>15</sup> Department of Physics and Astronomy, Vanderbilt University, VU Station 1807, Nashville, TN 37235, USA

<sup>16</sup> Instituto de Astronomía y Ciencias Planetarias, Universidad de Atacama, Copayapu 485, Copiapó, Chile

<sup>17</sup> Department of Physics & Astronomy, Texas Christian University, Fort Worth, TX 76129, USA

<sup>18</sup> Instituto de Astrofísica de Canarias (IAC), E-38205 La Laguna, Tenerife, Spain

<sup>19</sup> Universidad de La Laguna (ULL), Departamento de Astrofísica, E-38206 La Laguna, Tenerife, Spain

<sup>20</sup> Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

<sup>21</sup> Department of Physics and Astronomy, University of Utah, 115 S. 1400 E., Salt Lake City, UT 84112, USA

<sup>22</sup> Centro de Astronomía (CITEVA), Universidad de Antofagasta, Avenida Angamos 601, Antofagasta 1270300, Chile

<sup>23</sup> Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195, USA

<sup>24</sup> Center for Astrophysics and Space Astronomy, Department of Astrophysical and Planetary Sciences, University of Colorado, 389 UCB, Boulder, CO 80309-0389, USA

<sup>25</sup> Department of Physics, Salisbury University, Salisbury, MD 21801, USA

Received 2020 January 24; revised 2020 April 10; accepted 2020 April 13; published 2020 May 19

## Abstract

Many problems in contemporary astrophysics—from understanding the formation of black holes to untangling the chemical evolution of galaxies—rely on knowledge about binary stars. This, in turn, depends on the discovery and characterization of binary companions for large numbers of different kinds of stars in different chemical and dynamical environments. Current stellar spectroscopic surveys observe hundreds of thousands to millions of stars with (typically) few observational epochs, which allows for binary discovery but makes orbital characterization challenging. We use a custom Monte Carlo sampler (*The Joker*) to perform discovery and characterization of binary systems through radial velocities, in the regime of sparse, noisy, and poorly sampled multi-epoch data. We use it to generate posterior samplings in Keplerian parameters for 232,495 sources released in APOGEE Data Release 16. Our final catalog contains 19,635 high-confidence close-binary ( $P \lesssim$  few years,  $a \lesssim$  few au) systems that show interesting relationships between binary occurrence rate and location in the color–magnitude diagram. We find notable faint companions at high masses (black hole candidates), at low masses (substellar candidates), and at very close separations (mass-transfer candidates). We also use the posterior samplings in a (toy) hierarchical inference to measure the long-period binary-star eccentricity distribution. We release the full set of posterior samplings for the entire parent sample of 232,495 stars. This set of samplings involves no heuristic “discovery” threshold and therefore can be used for myriad statistical purposes, including hierarchical inferences about binary-star populations and subthreshold searches.

*Unified Astronomy Thesaurus concepts:* Binary stars (154); Close binary stars (254); Radial velocity (1332); Spectroscopy (1558); Astronomy data analysis (1858); Bayesian statistics (1900); Astrostatistics (1882); Surveys (1671)

## 1. Introduction

Binary-star systems provide key context and constraints for nearly all subfields in astrophysics (e.g., Price-Whelan et al. 2019; Rix et al. 2019). For two concrete examples, a

<sup>26</sup> Hubble Fellow.

<sup>27</sup> Carnegie-Princeton Fellow.

<sup>28</sup> NSF Astronomy and Astrophysics Postdoctoral Fellow.

measurement of the occurrence rate of stellar-mass black holes in the Milky Way would enable new constraints on binary black hole formation channels to explain merger events observed by LIGO (Abbott et al. 2016, 2019), and interpretation of spectroscopic observations of high-redshift galaxies and their stellar populations depends on understanding the impacts of binary-star evolution on stellar population parameters (e.g., Eldridge et al. 2017). One common need for all applications is improved constraints on the population properties (e.g., period, eccentricity, mass ratio distributions, and occurrence rates) of stellar multiplets and their variations with stellar type, chemistry, and dynamical environment, especially at the extrema of these stellar characteristics. This has only been comprehensively done for a sample of a few hundred stars in the solar neighborhood (Raghavan et al. 2010), for specific stellar types (e.g., Moe & Di Stefano 2017), or with imprecise statistics using large samples of stars (Badenes et al. 2018).

While this problem spans a huge range in timescales (from hours to millennia), current or near-future stellar surveys such as Gaia, APOGEE, LAMOST, and SDSS-V (Zhao et al. 2012; Gaia Collaboration et al. 2016, 2018; Kollmeier et al. 2017; Majewski et al. 2017) have the capacity to deliver samples of binary stars and stellar companions orders of magnitude larger than are presently known, throughout all stages of stellar evolution. Even with existing data sets, we now have sufficiently large sample sizes and the robust tools needed to perform hierarchical inferences to constrain much more detailed models of binary-star formation and evolution (e.g., Moe & Kratter 2018; El-Badry & Rix 2019). However, the most precise measurements of the binary-star population properties will benefit from joint analysis of all stellar surveys, which cover a range of stellar types, wavelengths, and measurement techniques.

As a step toward large-scale population inference, we focus here on multi-epoch spectroscopic data from the APOGEE surveys. This survey sequence has predominantly targeted red giant stars (although with DR16 there are now many main-sequence stars, see below). Because of operational revisit decisions to reach signal-to-noise thresholds, the surveys deliver some time-domain information, although they were not designed with binary-star characterization as the highest priority. The fundamental data are  $R \sim 20,000$   $H$ -band spectroscopy taken with the primary purpose of mapping the Milky Way in elemental abundances and kinematics. This survey is not the perfect target for binary-star identification, but it arguably has one of the best combinations of spatial reach around the Milky Way, coverage of the color–magnitude diagram, and multi-epoch data.

The challenge of working with data that were *not* taken primarily for binary characterization is that the data are sparse, time baselines are variable, and most individual systems are not characterized uniquely (in orbital parameters). Indeed, period-fitting tasks generically produce multimodal likelihood functions and posterior pdfs for periods, amplitudes, and phases, and the Kepler problem is no different. These multimodal likelihood functions are a nightmare for optimization or sampling. For this reason, we created a custom sampler, *The Joker* (Price-Whelan et al. 2017), that performs brute-force rejection sampling using a large, initial prior sampling. Because this sampler does not produce Markov chains, but rather just uses dense prior samples, it does not get “stuck” in local optima. Instead, it samples the full parameter space, with zero

autocorrelation among samples. To proceed, however, the sampler makes strong assumptions—for example, that the source is a single-lined spectroscopic binary (SB1) with only one companion. The advantage is that it generates full samplings over orbital parameters for arbitrarily sampled radial-velocity data.

It is important to clarify that we use *The Joker* to sample *every* APOGEE target as if it were an SB1 (see Section 7 for some discussion of the implications of this assumption). Discovery of which stars really do have companions then becomes a post-processing step on the confidence of the characterization. This project deliberately conflates discovery with characterization. Doing so has the added advantage that we can deliver full posterior samplings even for stars that are not obviously in binary systems; these can then be used for statistical studies and subthreshold searches when new data arrive (e.g., from future surveys or follow-up). In what follows, we run *The Joker* on all of APOGEE DR16 (subject to some quality cuts), and release the resulting catalog of posterior samples, along with some summary metadata for the subsample with good, uniquely determined companions and orbits. We demonstrate the use of the samplings with a few examples of interesting objects and a simple hierarchical probabilistic populations inference.

This article is similar to Price-Whelan et al. (2018), but should be viewed as a strict replacement (rather than an improvement) of the catalogs and results of that work. We have made substantial improvements to the methodology (improvements to *The Joker*; see Appendix A), and the APOGEE data in DR16 are of higher quality, substantially larger in volume, and have more observation epochs.

## 2. Data

We use spectroscopic data from data release 16 (DR16) of the APOGEE surveys (Majewski et al. 2017; Ahumada et al. 2019; H. Jönsson et al. 2020, in preparation). APOGEE is a component of the Sloan Digital Sky Survey IV (SDSS-IV; Gunn et al. 2006; Blanton et al. 2017); its main goal is to survey the chemical and dynamical properties of stars across much of the Milky Way disk by obtaining high-resolution ( $R \sim 22,500$ ; Wilson et al. 2019), infrared ( $H$ -band) spectroscopy of hundreds of thousands of stars. The primary survey targets are selected with simple color and magnitude cuts (Zasowski et al. 2013, 2017), but the survey uses fiber-plugged plates that cover only a small fraction of the available area, which leads to extremely nonuniform coverage of the Galactic stellar distribution (see, e.g., Figure 1 in Ahumada et al. 2019).

DR16 is the first SDSS data release to contain APOGEE data observed with a duplicate of the APOGEE spectrograph on the 2.5 m Irénée du Pont telescope (Bowen & Vaughan 1973) at Las Campanas Observatory, providing access to targets in the Southern Hemisphere. For the first time, this data release also contains calibrated stellar parameters for dwarf stars (H. Jönsson et al. 2020, in preparation). These two facts mean that DR16 contains nearly three times more sources with calibrated stellar parameters than the previous public data release, DR14 (Abolfathi et al. 2018; Holtzman et al. 2018). The interested reader may refer to Section 4 of Ahumada et al. (2019) for many more details about APOGEE DR16.

Most APOGEE stars are observed multiple times in separate “visits” that are combined before the APOGEE data reduction pipeline (Nidever et al. 2015; Zamora et al. 2015;

García Pérez et al. 2016) determines stellar parameters and chemical abundances for each source. While the visit spectra naturally provide time-domain velocity information about sources (thus enabling searches for massive companions), studying stellar multiplicity is not the primary goal of the survey: the cadence and time baseline for a typical APOGEE source are primarily governed by trying to schedule a set number of visits determined by signal-to-noise thresholds for the faintest targets in a given field. A small number of fields (five) were designed specifically for companion studies and have  $>10$  visits spaced to enable binary-system characterization.

While some past studies have made use of other fields with large numbers of visits to study binary-star systems (Troup et al. 2016; Fernández-Trincado et al. 2019), a consequence of this strategy is that the time resolution and number of visits for the vast majority of APOGEE sources in DR16 is not sufficient for fully determining companion orbital properties, as illustrated below. Still, the large number of targets in APOGEE and the dynamic range in stellar and chemical properties offers an exciting opportunity to study the *population* of binary-star systems as a function of these intrinsic properties, even if most individual systems are poorly constrained. We have previously developed tools to enable such studies (Price-Whelan et al. 2017), as summarized in Section 3 below. Here, we describe quality cuts we apply to the APOGEE DR16 catalogs before proceeding, and modifications to the visit-level velocity uncertainties to account for the fact that they are generally underestimated by the APOGEE data reduction pipeline.

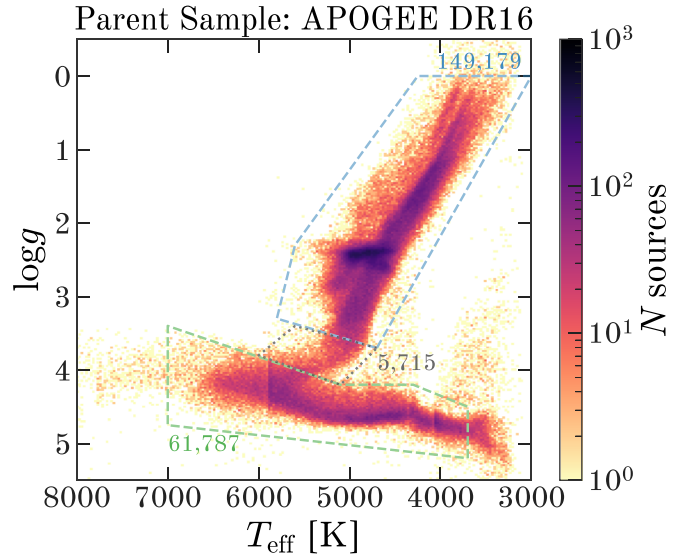
### 2.1. Quality Cuts and Defining a Parent Sample

The primary goal of this *Article* is to produce a catalog of posterior samplings in Keplerian orbital parameters for *all* high-quality APOGEE sources in DR16 with multiple, well-measured radial velocities. We therefore impose a set of quality cuts to subselect APOGEE DR16 sources by rejecting sources or visits using the following APOGEE bitmasks (Holtzman et al. 2018; H. Jönsson et al. 2020, in preparation):

1. Source-level (allStar) STARFLAG must not contain VERY\_BRIGHT\_NEIGHBOR, SUSPECT\_RV\_COMBINATION (bitmask values: 3, 16).
2. Source-level (allStar) ASPCAPFLAG must not contain TEF\_BAD, LOGG\_BAD, VMICRO\_BAD, ROTATION\_BAD, VSINI\_BAD (bitmask value: 16, 17, 18, 26, 30).
3. Visit-level (allVisit) STARFLAG must not contain VERY\_BRIGHT\_NEIGHBOR, SUSPECT\_RV\_COMBINATION, LOW\_SNR, PERSIST\_HIGH, PERSIST\_JUMP\_POS, PERSIST\_JUMP\_NEG (bitmask value: 3, 9, 12, 13, 16).

These bitmasks are designed to remove the most obvious data reduction or calibration failures that would directly impact the visit-level radial-velocity determinations. However, we later impose a stricter set of quality masks when showing results in Section 6.2. After applying the above masks, we additionally reject any source with  $<3$  visits. Our final parent sample contains 232,495 unique sources, selected from the 437,485 unique sources in all of APOGEE DR16. Of the  $\approx 200,000$  sources removed, the vast majority were dropped because they had  $<3$  visits ( $\approx 17,000$  were removed by the quality cuts).

Figure 1 shows the sources in our parent sample—i.e., APOGEE sources with three or more visits that pass the quality



**Figure 1.** Two spectroscopic (ASPCAP) stellar parameters—effective temperature,  $T_{\text{eff}}$ , and log-surface gravity,  $\log g$ —of the APOGEE DR16 sources that pass our quality cuts. These sources represent our “parent sample.” Pixel coloration indicates the number of sources in each bin of stellar parameters. Outlined regions roughly identify the red giant branch (upper polygon, blue), subgiant branch (middle polygon, black), and (FGK-type) main sequence (lower polygon, green). Numbers next to each selection polygon indicate the number of sources in each.

cuts described above—as a function of spectroscopic stellar parameters  $T_{\text{eff}}$ , effective temperature, and  $\log g$ , the log-surface gravity. While the majority of sources are giant-branch stars ( $>150,000$ ), a substantial number of main-sequence stars are present ( $>60,000$ ), thanks to the APOGEE data reduction pipeline improvements for DR16 (H. Jönsson et al. 2020, in preparation). Figure 2 shows some statistics about the time coverage of the visits for sources in our parent sample. About half of the sources have a small number of visits spread over a small time baseline (the time spanned from the first to last visit for each source): 50% of sources have  $<5$  visits over  $<100$  days. About 7% of sources (15,366) have  $\geq 10$  visits over  $\geq 100$  days.

### 2.2. Visit Velocity Uncertainty Calibration

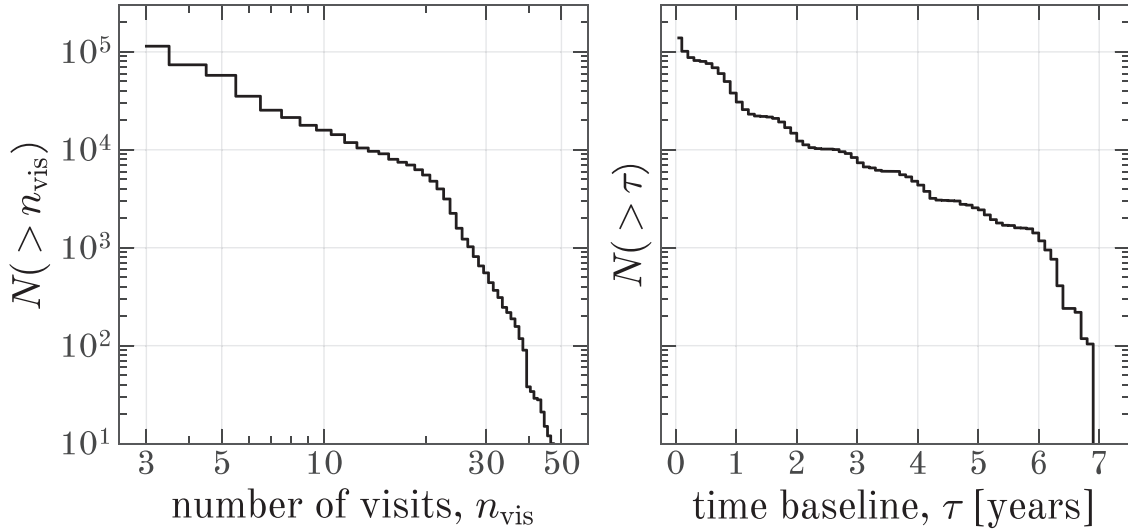
The significance of apparent radial-velocity variations, especially when considering low-mass or long-period companions, will depend strongly on the accuracy of the visit-velocity measurement errors. However, the catalog-level APOGEE visit-velocity errors (VRELERR in the allVisit file) are known to be underestimated (e.g., Cottaar et al. 2014; Badenes et al. 2018). Here, we adopt the relation defined in A. G. A. Brown et al. (2020, in preparation) to scale up the visit-velocity errors:

$$\sigma_v^2 = (3.5(\text{VRELERR})^{1.2})^2 + (0.072 \text{ km s}^{-1})^2, \quad (1)$$

where  $\sigma_v$  is the adopted visit-velocity error for a given visit, and VRELERR is the uncertainty reported in the APOGEE DR16 catalog. The form of this expression,  $(a \sigma_v^b)^2 + c^2$ , comes from the assumption that the visit-velocity noise variance values have a finite minimum value (i.e., a noise “floor”), and should either be increased or decreased by a multiplicative factor that scales with the pipeline-reported value of the uncertainty. The values of the constants in this function ( $a$ ,  $b$ ,  $c$ ) come from robustly fitting this relation to the observed visit velocity



## Complementary Cumulative Number of Sources



**Figure 2.** Some statistics of APOGEE DR16 visits. Left: number of sources with more than a given number of visits,  $n_{\text{vis}}$ . While  $\approx 50\%$  of sources have three visits, (114,263, 57,593, 15,862) sources have  $>(3, 5, 10)$  visits, respectively. A very small number of sources have  $>50$  visits. Right: number of sources with a time baseline,  $\tau$ , longer than given (on the horizontal axis). While  $\approx 50\%$  of sources have a time baseline  $\tau \lesssim 56$  days, (88,737, 9743) sources have  $\tau > (100, 1000)$  days.

scatters for each source as a function of their median visit velocity error values. This effectively applies a floor to the visit-velocity errors of  $72 \text{ m s}^{-1}$  and globally scales up the error values.

### 3. Methods

As illustrated above, a large number of sources in APOGEE DR16 have few visits that span a short time baseline. For most sources, we therefore expect that even if visit-level radial-velocity variations are detected with high significance, the companion orbital parameters will be very uncertain—i.e., the posterior probability distribution function (pdf) over orbital parameters will generally be multimodal with many modes of comparable integrated probability (e.g., Price-Whelan et al. 2017). Still, in unison, or within the context of a hierarchical model that utilizes the individual posterior samplings, the combination of all of these individually weakly constrained binary-star orbits provides information about the population of binary stars. We have previously defined and implemented a custom Monte Carlo sampler for precisely this problem: *The Joker* (Price-Whelan et al. 2017). *The Joker* is designed to deliver converged posterior samplings over Keplerian orbital parameters given radial-velocity observations, even when the observations are sparse or very noisy. Its prior application to APOGEE DR14 (Price-Whelan et al. 2018) resulted in a released catalog of over 5000 binary-star systems; this article and companion catalog supersede the previous work and its catalog.

As before, we use a parameterization of the two-body problem in which the radial velocity,  $v$ , of an observed star in a binary system (referred to as the *primary*, even if it is less massive than its companion) can be expressed as:

$$v(t) = K \zeta(t; P, e, \omega, M_0, t_0) + v_0, \quad (2)$$

where  $K$  and  $v_0$  are linear parameters (in time  $t$ ). The other parameters—period  $P$ , eccentricity  $e$ , argument of periastron  $\omega$ ,

reference time  $t_0$ , and mean anomaly at reference time  $M_0$ —enter through the nonlinear function

$$\zeta(t; P, e, \omega, M_0, t_0) = \cos(f + \omega) + e \cos \omega, \quad (3)$$

where  $f$  is the true anomaly, and we always set the reference time,  $t_0$ , to the minimum observation time for a given set of radial-velocity observations.

#### 3.1. Updates to The Joker

Since our initial paper defining *The Joker*, we identified a conceptual error in the assumptions made about the prior over the linear parameters ( $K, v_0$ ) in Price-Whelan et al. (2017). We previously assumed that adopting a sufficiently broad, Gaussian prior over the linear parameters meant that we could ignore an explicit definition of this prior. In particular, this allowed us to drop any terms related to the prior over these parameters in the marginal likelihood expression (Equation (11) in Price-Whelan et al. 2017). This assumption is not correct, and can lead to unexpected behavior when applied to data that are very noisy or have a small time baseline compared to the samples of interest. We have rewritten the expression for the marginal likelihood that underlies *The Joker* in Appendix A, based on the notation in D. W. Hogg et al. (2020, in preparation).

Another issue with the assumption in the original implementation of *The Joker* is that the prior over the velocity semi-amplitude,  $K$ , was identical at all period and eccentricity values. This implies vastly different prior beliefs about the companion mass as a function of orbital period. For example, a zero-mean Gaussian prior on  $K$  with a standard deviation of  $30 \text{ km s}^{-1}$  transforms to reasonable prior beliefs about companion mass at periods around 1 yr, but gives substantial prior probability to companion masses  $>100 M_\odot$  at periods  $>10^4$  days. We therefore, by default, adopt a new (also Gaussian) prior on the semi-amplitude with a variance,  $\sigma_K^2$ , that scales with the period and

**Table 1**  
Summary and Description of Parameters and Prior pdfs

Parameter	Prior	Description
<i>Nonlinear parameters</i>		
$P$	$\ln P \sim \mathcal{U}(2, 16384)\text{day}$	Period
$e$	$e \sim \text{Beta}(0.867, 3.03)$	Eccentricity
$t_0$	Fixed (minimum time)	Reference time
$M_0$	$M_0 \sim \mathcal{U}(0, 2\pi)\text{rad}$	Mean anomaly at reference time
$\omega$	$\omega \sim \mathcal{U}(0, 2\pi)\text{rad}$	Mrgument of pericenter
$s$	$\ln s \sim \mathcal{N}(\mu_s, \sigma_s^2)$	Extra “jitter” added in quadrature to each visit-velocity error
<i>Linear parameters</i>		
$K$	$K \sim \mathcal{N}(0, \sigma_K^2)\text{ km s}^{-1}$ $\sigma_{K,0} = 30\text{ km s}^{-1}$ $P_0 = 365\text{ days}$	Velocity semi-amplitude, where $\sigma_K$ is given by Equation (4) (see Equation (4)) (see Equation (4))
$v_0$	$v_0 \sim \mathcal{N}(0, \sigma_{v_0}^2)\text{ km s}^{-1}$ $\sigma_{v_0} = 100\text{ km s}^{-1}$	System barycentric velocity

**Note.**  $\text{Beta}(a, b)$  is the beta distribution with shape parameters  $(a, b)$ ,  $\mathcal{U}(a, b)$  the uniform distribution over the domain  $(a, b)$ , and  $\mathcal{N}(\mu, \sigma^2)$  is the normal distribution with mean  $\mu$  and variance  $\sigma^2$ .

eccentricity such that

$$\sigma_K^2 = \sigma_{K,0}^2 \left( \frac{P}{P_0} \right)^{-2/3} (1 - e^2)^{-1}, \quad (4)$$

where  $\sigma_{K,0}$  and  $P_0$  are additional hyperparameters that must be specified. This new prior on  $K$  has the advantage that, at fixed primary mass, it has a fixed form in companion mass that does not depend on period or eccentricity.

We have also made a number of improvements to the Python implementation of the sampler.<sup>29</sup> For example, the prior distributions over all parameters are now specified as `pymc3` (Salvatier et al. 2016) distribution objects, meaning that the priors over the nonlinear Keplerian parameters ( $P$ ,  $e$ ,  $\omega$ , and  $M_0$ ; see Price-Whelan et al. 2017) are now fully customizable. Using `pymc3` also enables compatibility with more efficient Markov chain Monte Carlo methods, such as Hamiltonian Monte Carlo (HMC), which is useful for seamlessly transitioning from generating posterior orbit samples with *The Joker* to HMC when the system parameters are highly constrained (as discussed below).

### 3.2. Assumptions, Caveats, and Known Failures

The assumptions that underlie the sampling procedure described above are enumerated in Section 3.1 in Price-Whelan et al. (2018). We therefore briefly rephrase the most important implications of these assumptions as a set of caveats and points of caution for any users of the posterior samplings released with this article.

First, despite allowing for a per-source, additive (in variance), extra uncertainty parameter ( $s$  in Table 1), we are very sensitive to outliers—or more generally, any visit velocities with strongly non-Gaussian noise properties. In APOGEE, this might occur for blended sources, multiple-star systems with more than one star that contribute significantly to

infrared flux, or sources with stellar parameters that are beyond the grid of template spectra (Nidever et al. 2015).

Second, related to the first point, we make the strong assumption that all sources are single-lined, i.e., that any binary companions do not contribute significantly to the observed spectra. This is wrong in general: some systems will be double-lined (sometimes only subtly, but detectable, e.g., El-Badry et al. 2018), and we will be biased in cases where these sources are not properly filtered by the quality cuts done above. This also means that we will definitely miss systems with similar masses (and especially “twin” binaries; El-Badry et al. 2019). We expect this assumption to be worse on the main sequence than on the giant branch, because any lower-mass companion to a giant-branch star will have a luminosity hundreds or thousands of times fainter.

Third, the adopted prior pdf over systemic velocities,  $v_0$ , is reasonable for considering all APOGEE sources (i.e., for a mixture of kinematically disk-like and halo-like sources), but may lead to biased posterior samplings for sources in globular clusters or dwarf galaxies. For such systems, it would be safer to rerun *The Joker* with specialized priors over systemic velocity for each individual host system.

Fourth, in this work, we assume that all systems are binary systems, so triples or higher-order stellar multiplets will generally have incorrect samplings. We will generate samplings for systems that are consistent with being hierarchical triple-star systems, but this is beyond the scope of the general-use catalog considered here.

Finally, related to the fourth point, we assume that all velocity variations represent orbital motion (i.e., not pulsation or similar intrinsic stellar phenomena). As shown later, this assumption is violated when  $\log g \lesssim 1$ , where stellar surface jitter masquerades as orbital motion and leads to (a small amount of) contamination in our sample of binary-star systems.

## 4. Running *The Joker* on APOGEE DR16

For each of the 232,495 sources in the parent catalog of sources selected from APOGEE DR16 (Section 2), we run *The Joker* (Price-Whelan et al. 2017) in order to generate posterior samplings over the Keplerian orbital parameters, including an additional per-source uncertainty or “jitter” parameter that is added in quadrature to the (adjusted) APOGEE visit-velocity uncertainties (see above). We start by generating a cache of 100,000,000 prior samples for the nonlinear parameters generated from the prior pdf summarized in Table 1 (top rows). For each source, we iteratively read random blocks of samples out of this cache, evaluate the marginal likelihood, and rejection sample to produce posterior samplings in the nonlinear parameters. We repeat this iterative process until we reach the total number of samples in the cache, or until we obtain a requested number of prior samples,  $M_{\min}$ ; this is an arbitrary parameter that we set to  $M_{\min} = 512$  for this work. In practice, this is done by parallelizing the sampling (over sources) and takes about 8 hr to run on 720 cores on our local compute cluster (at the Flatiron Institute) for the entire sample.

After this procedure, the sources are in one of two stages of completion: Sources either have  $M_{\min} = 512$  posterior samples (227,999 sources) and are complete, or  $M_{\min} < 512$  posterior samples (4496 sources) and more samples are needed. For sources that require more samples, these can be split again into two classes—sources with unimodal samplings in period, and

<sup>29</sup> <https://github.com/adrn/thejoker>

sources with multimodal samplings—following the criteria described in Price-Whelan et al. (2017). For incomplete sources with multimodal samplings (2787 sources), these would need to be rerun with a much larger number of prior samples to reach 512 posterior samples, but here we mark these sources as incomplete. For incomplete sources with unimodal samplings (1709 sources), we use the samples returned from *The Joker* to initialize a Hamiltonian Monte Carlo (HMC) to continue generating posterior samples. We use `pymc3` with the *No-U-Turn Sampler* (NUTS; Homan & Gelman 2014), using the dense mass-matrix tuning prescription implemented in `exoplanet` (Foreman-Mackey et al. 2019), and run four chains in parallel, each for 1000 tuning steps and 4000 steps subsequently. We use these four chains and `pymc3` to compute the Gelman-Rubin convergence statistic,  $\hat{R}$  (Gelman & Rubin 1992), for each parameter for each source; If  $\hat{R} < 1.1$  for all parameters, we downsample to 512 samples and mark that source as MCMC-completed. If the HMC sampling fails (e.g., chains diverge substantially), we fall back to the sampling from *The Joker*—this generally only happens for data with serious systematic issues (e.g., one outlier visit velocity with an unreasonable velocity measurement).

At this point, we now have up to 512 posterior samples of the parameters listed in Table 1 for most of the 232,495 APOGEE sources in the parent sample. These full samplings will be released as a Value-Added Catalog (VAC) with the SDSS DR16+ “mini” data release planned for 2020 July.

## 5. A Catalog of Binary Stars

For some science cases and exploration, it is useful to define catalogs of systems with likely companions from the posterior samplings generated here, as we illustrate below. This requires making decisions and imposing hard cuts on the samplings returned by the above procedure. However, most of the orbital parameter samplings for most sources are highly uncertain and multimodal, and it is therefore not possible to define simple cuts on physical parameters (e.g., companion mass) to produce a simple catalog. In the previous iteration of this work (Price-Whelan et al. 2018), we defined cuts based on percentiles computed from the distribution of  $\ln K$  values after comparing to running the same pipeline on a control sample of data generated with purely Gaussian noise properties and time information taken from the APOGEE DR14 visits. Here, we use both a selection based on the posterior samples in  $K$ , as well as a likelihood ratio comparing the Keplerian orbit model assumed by *The Joker* with a (robust) constant-velocity model for each source.

Following Price-Whelan et al. (2018), we again compute the first-percentile values of the posterior samples in  $\ln K$  for each source and refer to these values as  $P_{1\%}(\ln K)$ ; this amounts to an estimate of the 99% confidence lower limit on the (log) velocity semi-amplitude. We again also generate a simulated control sample of data to assess contamination when making selections using this quantity. For each APOGEE source in our parent sample, we take the maximum *a posteriori* sample returned from the procedure defined above and subtract the orbit computed from this sample from the visit velocity data. We then rerun *The Joker* on all of the residual data, and compute  $P_{1\%}(\ln K)$  for each star in the control sample. We find that  $<5\%$  of the control sample passes a cut of  $P_{1\%}(\ln K) > 0$  (i.e., a conservative cut to require that sources have a velocity semi-amplitude  $>1 \text{ km s}^{-1}$ ).

For each APOGEE source, after generating the posterior samplings, we compute and store the maximum (over posterior samples) *unmarginalized* log-likelihood value for the Keplerian orbit model, i.e., for one source with  $N$  visit-velocity measurements  $v_n$  at times  $t_n$ ,

$$\ln \hat{L}_1 = \sum_n^N \ln \mathcal{N}(v_n | v(t_n; \theta_{\max}), \sigma_n^2), \quad (5)$$

where  $\sigma_n$  are the adjusted visit-velocity errors (Equation (1)), and  $v(\cdot)$  is given by Equation (2) and is evaluated using the parameters for the maximum likelihood posterior sample,  $\theta_{\max}$ .

For each source, we then also compute the maximum log-likelihood value for the visit data under a model that assumes that the visit velocities are drawn from a constant velocity with Gaussian uncertainties but allowing for  $<20\%$  outliers; we refer to this model as a *robust constant-velocity* model for the visit velocities. Using the same notation for the visit-velocity data as above, the likelihood under this model is given by

$$\begin{aligned} \hat{L}_2 = & \prod_n^N [(1 - f) \mathcal{N}(v_n | \mu_v, \sigma_n^2) \\ & + f \mathcal{N}(v_n | \mu_v, \sigma_n^2 + \Sigma)], \end{aligned} \quad (6)$$

where  $f$  is the outlier fraction,  $\mu_v$  is the constant-velocity value, and  $\Sigma$  is the variance of the outlier model component (assumed to be large). We optimize this likelihood using the BFGS algorithm (Nocedal & Wright 2006) with bounds on the parameters such that  $f \in (0, 0.2)$ ,  $\mu_v \in (-500, 500) \text{ km s}^{-1}$ , and  $\Sigma \in (0, 3000) (\text{km s}^{-1})^2$ . We store the optimized log-likelihood value of the robust constant-velocity model and refer to this as  $\ln \hat{L}_2$ .

We define a catalog of binary-star systems based on the posterior samplings generated from *The Joker* by selecting sources for which

$$P_{1\%}(\ln K) > 0 \quad (7)$$

$$(\ln \hat{L}_1 - \ln \hat{L}_2) > 4.6. \quad (8)$$

The cut on the log-likelihood ratio comes from the (adopted) condition that the maximum Kepler model likelihood should be  $>100$  times the maximum robust constant-velocity likelihood value. Of the 232,495 total sources in our parent sample, 19,635 sources pass the selection above; those are binary systems where we can provide meaningful orbit parameter samplings. Summary information for the samplings generated from all sources in the parent sample is included in Table B1; in this table, the Boolean mask `binary_catalog` can be used to select the 19,635 sources that pass the binary-star selections defined in Equation (8).

Figures 3 and 4 show some examples of binary systems that passed the selection above. In each figure, each row is a different APOGEE source, the two figures show some example short baseline (Figure 3) and long baseline (Figure 4) cases. The left column of panels in each figure show the radial velocity data (black markers) for randomly chosen sources with (3, 5, 7, 9) visits (from top to bottom), and the blue lines show radial-velocity orbits computed from the posterior samplings generated by *The Joker*. The right column of panels show the same posterior samples (blue markers), but in a projection of the parameter space (period  $P$  and minimum companion mass  $M_{2,\min}$ ). To compute  $M_{2,\min}$ , we use the posterior samplings from *The Joker* along with primary stellar masses computed by

**Table B1**  
Description of the Metadata Table Containing Summary Information for All APOGEE Sources in the Parent Sample

Metadata for All Sources in the Parent Sample		
Column Name	Unit/Format	Description
APOGEE_ID		APOGEE source identifier
n_visits		Number of visits that pass our quality cuts
MAP_P	days	$P$ , orbital period
MAP_P_err	days	
MAP_e		$e$ , orbital eccentricity
MAP_e_err		
MAP_omega	rad	$\omega$ , argument of pericenter
MAP_omega_err	rad	
MAP_M0	rad	$M_0$ , phase at reference epoch
MAP_M0_err	rad	
MAP_K	km s <sup>-1</sup>	$K$ , velocity semi-amplitude
MAP_K_err	km s <sup>-1</sup>	
MAP_v0	km s <sup>-1</sup>	$v_0$ , systemic velocity
MAP_v0_err	km s <sup>-1</sup>	
MAP_s	km s <sup>-1</sup>	$s$ , excess variance parameter
MAP_s_err	km s <sup>-1</sup>	
t0_bmjd		Reference epoch (Barycentric MJD)
baseline	days	Time baseline of the visits
MAP_ln_likelihood		The log-marginal-likelihood value of the MAP sample
MAP_ln_prior		The log-prior value of the MAP sample
max_unmarginalized_ln_likelihood		Maximum value of the unmarginalized likelihood
max_phase_gap		Maximum gap in phase of the data for the MAP sample
periods_spanned		Number of (MAP) periods spanned by the data
phase_coverage		Phase coverage of the data folded on the MAP period
phase_coverage_per_period		Maximum number of data points within a single MAP period
unimodal	boolean	True if the sampling is unimodal in period
joker_completed	boolean	True if the source has 512 samples from <i>The Joker</i>
mcmc_completed	boolean	True if the source has 512 samples from MCMC
mcmc_success	boolean	True if the MCMC sampling converged
gelman_rubin_max		Maximum (over parameters) value of $\hat{R}$
robust_constant_ln_likelihood		Maximum log-likelihood value of a robust constant-velocity fit to the visit data
binary_catalog	boolean	True if the source is in the binary catalog (Section 5)
(232,495 rows)		

**Note.** All columns ending in `_err` are estimates of the standard deviation of the posterior samples around the MAP sample, but are only computed for unimodal samplings.

(This table is available in its entirety in FITS format.)

Queiroz et al. (2019) by sampling over the reported uncertainties on prior mass (assuming a Gaussian noise distribution on primary mass).

### 5.1. The Gold Sample

The majority of binary-star systems that comprise the catalog defined above have strongly multimodal samplings in orbital properties. While this is useful for binary *population* studies, it is more difficult to summarize the system orbital properties and their trends with stellar properties, as it is not possible to simply compress the samplings. We therefore construct an additional catalog for the subset of sources that pass a more stringent set of quality cuts and have converged, unimodal, or bimodal posterior samplings. We define this *Gold Sample* as sources that pass the following cuts:

1. Matches to a Gaia source within 2'' of the reported 2MASS sky position.
2. Has a stellar mass measurement in the STARHORSE catalog (Queiroz et al. 2019).
3. No additional Gaia sources within 2'' with a  $G$ -band magnitude difference  $\Delta G > -5$  (to remove sources that

would lie within the APOGEE fiber and appreciably contaminate the spectrum).

4. No additional Gaia sources within 10'' with a  $G$ -band magnitude difference  $\Delta G > 2.5$  (to remove bright neighbor stars).
5.  $-0.5 < \log g < 5.5$  (reliable stellar parameters).
6.  $3500 \text{ K} < T_{\text{eff}} < 10,000 \text{ K}$  (reliable stellar parameters).<sup>30</sup>
7.  $-2.5 < [\text{M}/\text{H}] < 0.5$  (reliable stellar parameters).
8.  $s_{\text{MAP}} < 0.5 \text{ km s}^{-1}$  (a small inferred excess variance).
9.  $n_{\text{vis}} > 5$  (more than five visit spectra).

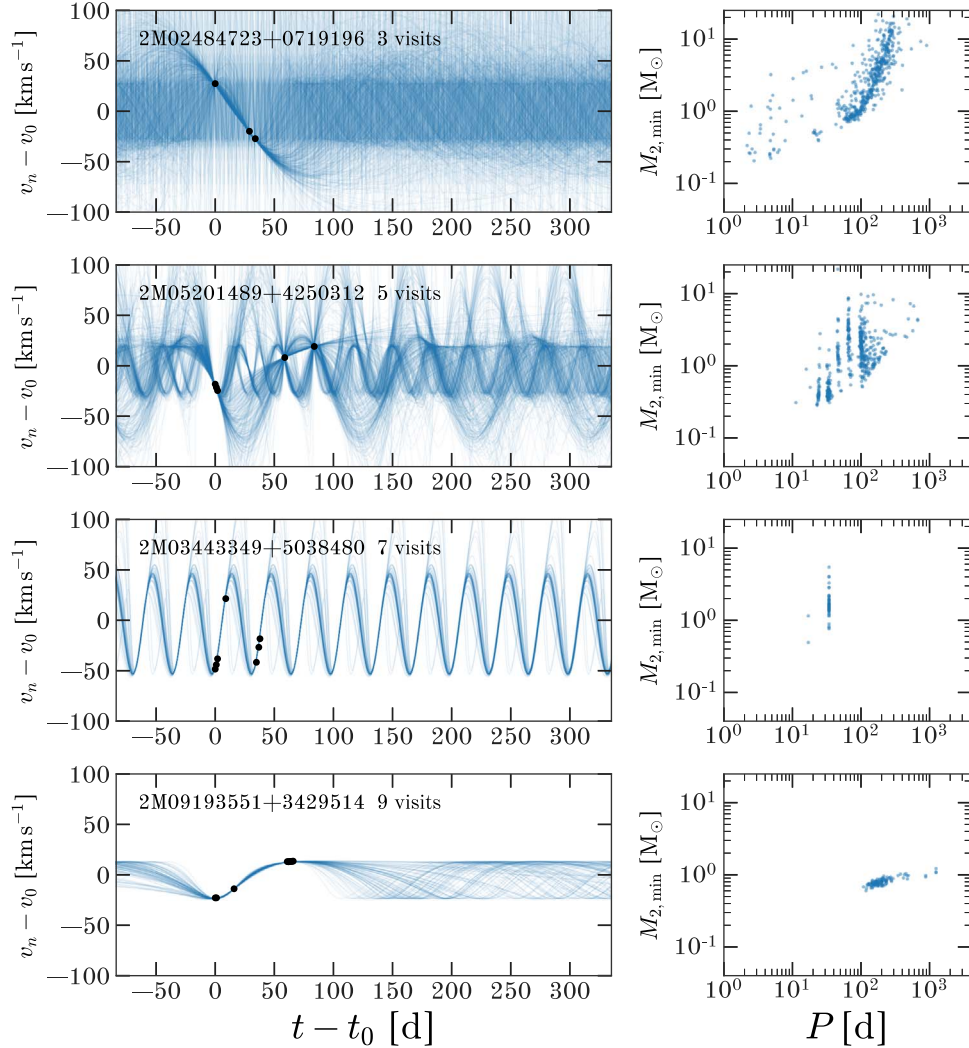
Here,  $s_{\text{MAP}}$  is the maximum *a posteriori* (MAP) sample in the excess variance parameter.

We identify sources with unimodal posterior samplings as sources for which the MCMC procedure succeeded (see Section 3). We identify sources with bimodal samplings using the same procedure as Price-Whelan et al. (2018). Briefly, we use a  $k$ -means clustering algorithm with  $k = 2$  to identify

<sup>30</sup> While the new APOGEE MARCS grids extend below 3500 K (Ahumada et al. 2019; H. Jönsson et al. 2020, in preparation), there are few calibrator stars with such low temperatures, so the stellar parameters are less reliable for stars with  $T_{\text{eff}} < 3500 \text{ K}$ .



## Example Binary Systems: Short Baseline



**Figure 3.** Example binary-star systems that pass the selection and are included in the catalog released here for APOGEE sources with short visit baselines ( $\tau < 100$  days). Each row is an APOGEE source (indicated on the left panel). Left panels show the visit velocity data (black markers; error bars are typically smaller than the marker) and radial-velocity orbits computed from the posterior samples (blue lines). Right panels show the same samples in period  $P$  and minimum companion mass  $M_{2,min}$ .

clusters of samples in orbital period, and assess whether the samplings in each cluster are unimodal by checking whether all samples in each cluster lie within the mode size defined in Price-Whelan et al. (2017). In total, the *Gold Sample* contains 1032 systems with unimodal samplings, and 127 systems with bimodal samplings. Summary information and a list of sources in the *Gold Sample* are included in Table B2; this sample is also available for investigation through a Filtergraph portal.<sup>31</sup>

## 6. Results

The epoch baselines for most APOGEE sources,  $\tau \lesssim 1$  yr, imply that most of the  $\sim 20,000$  binary systems (and certainly the  $\sim 1000$  gold sample systems) will have  $P \lesssim$  years and  $a \lesssim$  few au. However, the overall binary-star population extends from close binaries to systems with  $a \gtrsim 20,000$  au, with a broad, approximately log-normal period distribution

centered at  $\sim \log(250 \text{ yr})$  (Raghavan et al. 2010). The binary systems we identified and study here are commonly referred to as “close binaries” (Badenes et al. 2018; Moe & Kratter 2018), representing the closest  $\sim 20\%$ – $40\%$  of all bound binary-star systems. With this in mind, we use our sample to study some simple population properties of these binary systems, as well as to highlight some interesting systems in our sample.

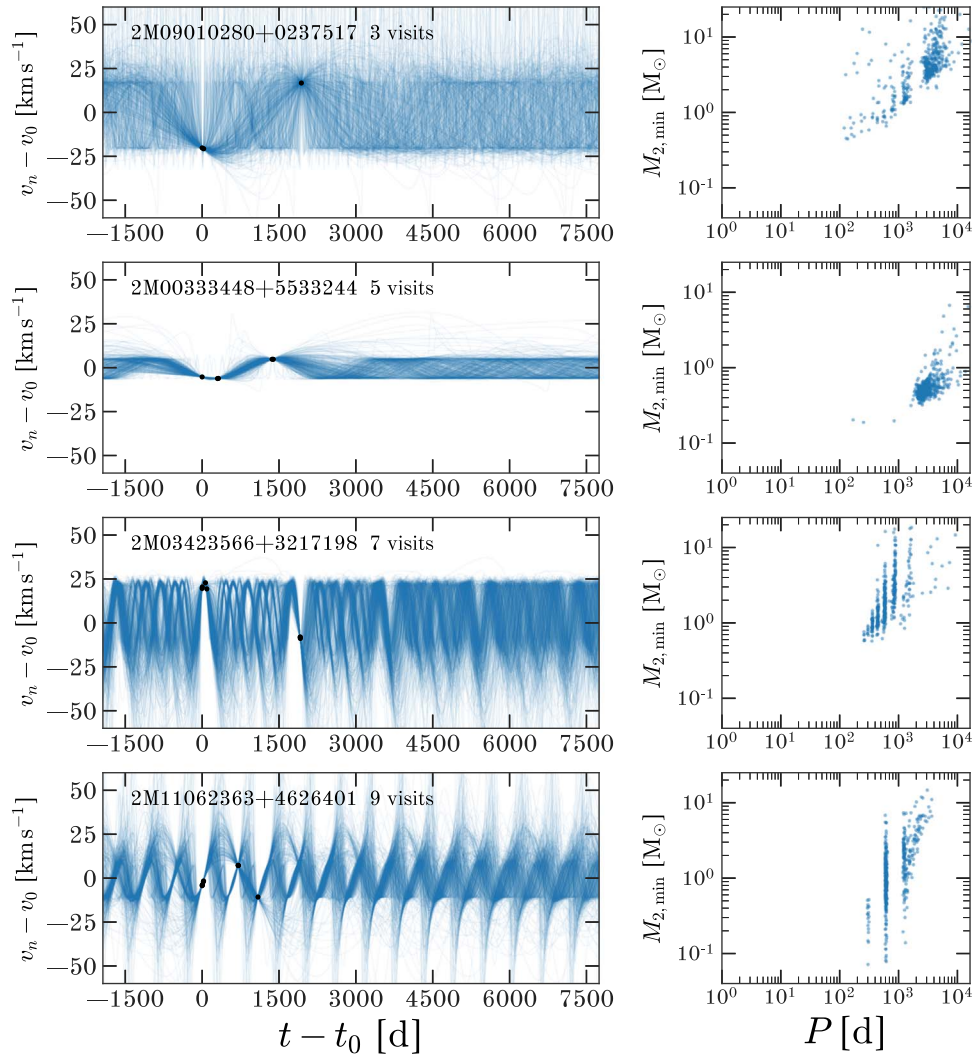
### 6.1. Close Binary Fraction Trends with Stellar Properties

Our catalog of binary stars is not complete, in the sense that the cuts we have made on the orbital-parameter samplings will impart nonuniform selection biases that depend on the true orbital properties of binaries and on the cadence of observations (visits) for each source. However, because of the simple target selection and observation strategy used by APOGEE (Zasowski et al. 2013, 2017) and APOGEE-2S (south; R. Beaton et al. 2020, in preparation; F. Santana et al. 2020, in preparation), we do not expect these binary-star selection

<sup>31</sup> [https://filtergraph.com/apogee\\_dr16\\_binaries/](https://filtergraph.com/apogee_dr16_binaries/)



## Example Binary Systems: Long Baseline



**Figure 4.** Same as Figure 3, but for APOGEE sources with long visit baselines ( $\tau > 1000$  days).

biases to depend strongly on stellar parameters (e.g., metallicity, surface gravity). We can therefore still use this catalog to study the *relative* close binary fraction within our sample, but caution against interpreting the binary fractions discussed below in an absolute sense. Note that here we use “binary fraction” to mean the observed fraction of *detected* binary systems, not the intrinsic or birth fraction of binary systems.

Figure 5 (right) shows a near-infrared, binned color-magnitude diagram for all APOGEE sources in the subset of our sample that cross-match to the Gaia DR2 astrometric catalog (Gaia Collaboration et al. 2016, 2018; Lindegren et al. 2018) and have a parallax signal-to-noise  $\varpi/\sigma_{\varpi} > 8$ ; we compute the absolute  $H$ -band magnitudes by converting parallax into distance as  $d = 1/\varpi$ . Each pixel is colored by the ratio of the number of binary-star systems identified by the selections defined above (Equation (8)) over the total number of sources in the parent sample; pixels with fewer than four stars in the parent sample are white. The solid (orange) line shows a 5 Gyr MIST isochrone (Dotter 2016; Choi et al. 2016; Paxton et al. 2011, 2013, 2015), with metallicity indicated in the figure legend. The dashed line shows the equal-mass binary

sequence for this isochrone, 0.75 mag above the main sequence. Note that, as expected, the observed binary fraction is higher in this region of the CMD, and is also higher toward younger and more massive main-sequence stars. Observe also that the red clump (around  $(J - K) \approx 0.6$ ,  $H \approx -1.8$ ) has a *low* binary fraction, as noted in Badenes et al. (2018). The width of the binary sequence (in  $H$  magnitude) can be explained by the spread in metallicities in the APOGEE sample.

The top left panel of Figure 5 shows another view of the observed binary fraction, here shown as a 1D function of surface gravity for stars in the giant branch selection region shown in Figure 1. There is a clear and sharp dip in the occurrence of binary systems near the red clump, and the close-binary fraction decreases appreciably with decreasing surface gravity (i.e., increased size). Both of these features are likely signatures of companion engulfment: as red giant stars evolve up the giant branch, any companions with orbital semimajor axes smaller than a few times the surface size of the primary star could be consumed (Ivanova et al. 2013), thus leading to an overall decrease in the binary fraction for larger stars. Note that the apparent upturn of the observed binary fraction at low

**Table B2**  
Description of the Data Table Containing Summary Information for All Sources in the *Gold Sample*

Metadata for the <i>Gold Sample</i>		
Column Name	Unit/ Format	Description
APOGEE_ID		APOGEE source identifier
mass	$M_{\odot}$	Mass of the primary, from the STARHORSE catalog
mass_err	$M_{\odot}$	Uncertainty on the primary mass
m2_min_1	$M_{\odot}$	1st percentile value of $M_{2,\min}$
m2_min_5	$M_{\odot}$	5th percentile value of $M_{2,\min}$
m2_min_16	$M_{\odot}$	16th percentile value of $M_{2,\min}$
m2_min_50	$M_{\odot}$	Median value of $M_{2,\min}$
m2_min_84	$M_{\odot}$	84th percentile value of $M_{2,\min}$
m2_min_95	$M_{\odot}$	95th percentile value of $M_{2,\min}$
m2_min_99	$M_{\odot}$	99th percentile value of $M_{2,\min}$
:		All columns from Table B1
:		All columns from APOGEE <i>allStar</i> file
:		All columns from Gaia DR2 <i>gaia_source</i> table
(1032 rows)		

(This table is available in its entirety in FITS format.)

$\log g$  is most likely dominated by contamination: intrinsic radial-velocity noise is interpreted as a binary system detection by the simplistic definition of binary fraction defined above.

The bottom-left panel shows the same, but as a function of effective temperature (as a proxy for the mass of the primary) for stars in the main-sequence selection region shown in Figure 1. As has been noted in many past studies, we find that the observed binary fraction (detected companion frequency) increases with stellar mass (effective temperature) along the main sequence (e.g., Duchêne & Kraus 2013). Like the compilation of companion frequencies shown in Duchêne & Kraus (2013) and a number of studies of local samples of binary systems (e.g., Eggleton & Tokovinin 2008; Raghavan et al. 2010; Gao et al. 2014), we find that the binary fraction increases steeply for primary stellar masses  $M_1 \gtrsim 1.1 M_{\odot}$  ( $T_{\text{eff}} \gtrsim 6000$  K).

Figure 6 shows the total fraction of detected close binaries for main-sequence stars as a function of bulk metallicity,  $[M/H]$ . As has been recently emphasized, we find that the observed fraction of close binaries is significantly anticorrelated with metallicity (El-Badry & Rix 2019; Moe et al. 2019), in disagreement with previous work (with a much smaller sample) that had found no dependency with metallicity (Jenkins et al. 2015). We find a shallower dependence of these properties, with a slope of  $\approx -0.1$  as compared to the previously determined slope of  $-0.2$ , which was also for close binaries (Moe et al. 2019). While we do not expect there to be strong selection biases that imprint on these quantities, we emphasize that we have not corrected for detection efficiency or completeness with our sample.

## 6.2. Orbital Parameter Trends with Stellar Properties

Figure 7 shows the inferred orbital periods for all stars in the *Gold Sample* as a function of orbital eccentricity (left) and

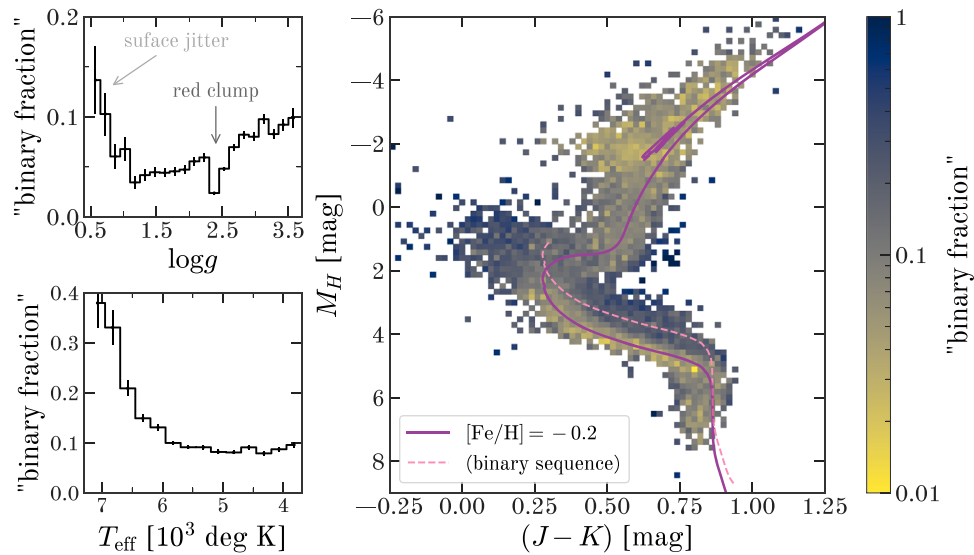
surface gravity (right), where the range of periods is limited (at large  $P$ ) by the length of the APOGEE surveys. The left panel clearly shows the impact of tidal circularization (e.g., Zahn 1977; Meibom & Mathieu 2005): the eccentricities of systems with  $P \lesssim 10$  days are much more peaked near  $e = 0$  as compared to systems with larger orbital periods (e.g.,  $P \gtrsim 100$  days). However, for systems with giant-star members, circularization occurs at longer periods (e.g., Price-Whelan & Goodman 2018): systems with low eccentricities and  $P > 30$  days tend to have lower  $\log g$  values (darker points) as compared to systems with  $P < 10$  days. In this panel, we visually inspected the small number of systems with  $P \lesssim 5$  days and  $e \gtrsim 0.4$  and found that most of these cases appear to result from (unexplained) systematic errors with the visit velocity data that were not removed by the quality cuts. Some of these cases also arise from failures of our assumption of binarity: when the data shows signs of more complex velocity variations, such as from a higher-order multiple-star system, the samplings returned by *The Joker* will be erroneous.

The right panel of Figure 7 shows orbital periods of systems in the *Gold Sample* (with unimodal or bimodal samplings) as a function of surface gravity. The diagonal lines in this panel show the orbital periods at which the orbital pericenter of a  $0.3 M_{\odot}$  companion is equal to the stellar surface radius for a  $1.1 M_{\odot}$  primary star (the median red giant branch mass in our sample) at the given surface gravity and eccentricity indicated. The apparent lack of systems with orbital periods shorter than or within  $\sim 1$  dex of these critical lines (with  $\log g \gtrsim 1$ ) suggests that a substantial fraction of binaries must merge or disrupt during the evolution of the primary star. The paradoxical points with seemingly small orbital periods at small  $\log g$  (i.e., suggesting they orbit within the surface of the primary star) are likely due to contamination from asteroseismic modes that manifest as velocity “jitter” in low-surface-gravity giant stars. This jitter can reach amplitudes of  $> 1 \text{ km s}^{-1}$  already by  $\log g \approx 1$  and likely increases toward even lower values of  $\log g$  (e.g., Hekker et al. 2008).

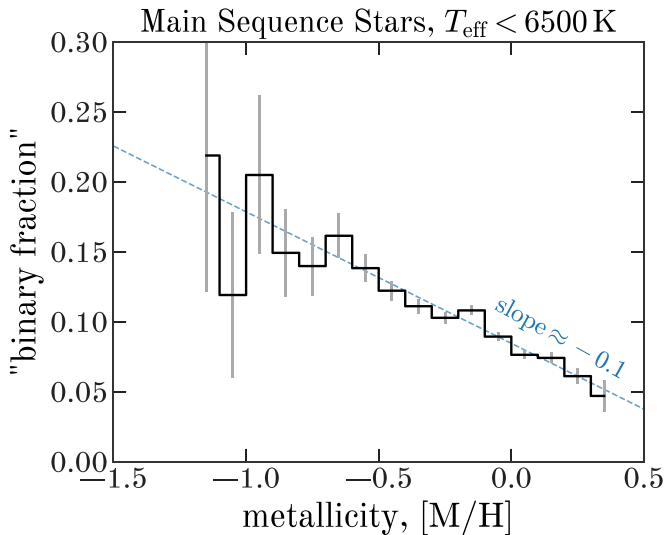
Also note the gradients in eccentricity (i.e., marker color) with respect to orbital period in the right panel of Figure 7. At a given surface gravity, there tend to be more black points (low eccentricity) closer to the stellar surface (closer to the diagonal lines). However, near the red clump ( $2 \lesssim \log g \lesssim 3$ ), there appears to be an overabundance of low-eccentricity points at orbital periods  $P \gtrsim 100$  days. Figure 8 shows histograms of maximum *a posteriori* MAP eccentricity values for sources in the *Gold Sample* with MAP period values  $P > 50$  days in four bins of surface gravity, from the main sequence (farthest left) to the upper giant branch (farthest right). Note that, around the red clump (third panel from the left), there are significantly more  $e < 0.1$  systems than expected (i.e., as compared to the previous bin). We therefore posit that systems with a primary around the red clump that have low eccentricities (especially  $e \lesssim 0.2$ ) have likely already ascended the giant branch, but asteroseismic observations would be required to test this hypothesis.

## 6.3. Interesting Low- and High-mass Companions

By construction, all sources in the *Gold Sample* have primary stellar-mass estimates,  $M_1$ , from the STARHORSE catalog (Queiroz et al. 2019). For these systems, we can then also convert the inferred orbital parameters into measurements of the minimum companion mass,  $M_{2,\min}$  (i.e.,  $M_2 \sin i$  where



**Figure 5.** Close binarity across the color-magnitude diagram. Upper left: observed close-binary fraction as a function of spectroscopic surface gravity,  $\log g$ , for all stars that pass the giant-branch selection indicated in Figure 1. Lower left: observed binary fraction as a function of spectroscopic effective temperature,  $T_{\text{eff}}$ , for all stars that pass the main-sequence selection indicated in Figure 1. Right: extinction-corrected 2MASS color-magnitude diagram (CMD) for all APOGEE sources, colored by the fraction of sources identified as binary-star systems (Section 5). Solid (dark purple) line shows a MIST isochrone for a 5 Gyr stellar population with  $[\text{Fe}/\text{H}] = -0.2$ , and dashed line indicates the corresponding equal-mass binary sequence for main-sequence stars. Panels in this figure are meant to be illustrative only, since the observed binary fraction is not the same as the true, complete binary fraction (see Section 5 for selection criteria).



**Figure 6.** Observed close-binary fraction as a function of bulk metallicity,  $[\text{M}/\text{H}]$ . Binary fraction is anticorrelated with metallicity, here measured with a slope of  $-0.1$ . Normalization of our observed binary fraction is set by the selection function of the APOGEE surveys and detection efficiency of our selection criteria (see Section 6.1). As the binary fraction is uncorrected for completeness, we tend to measure smaller binary fractions at, e.g., solar metallicity, as compared to other work (e.g., Raghavan et al. 2010).

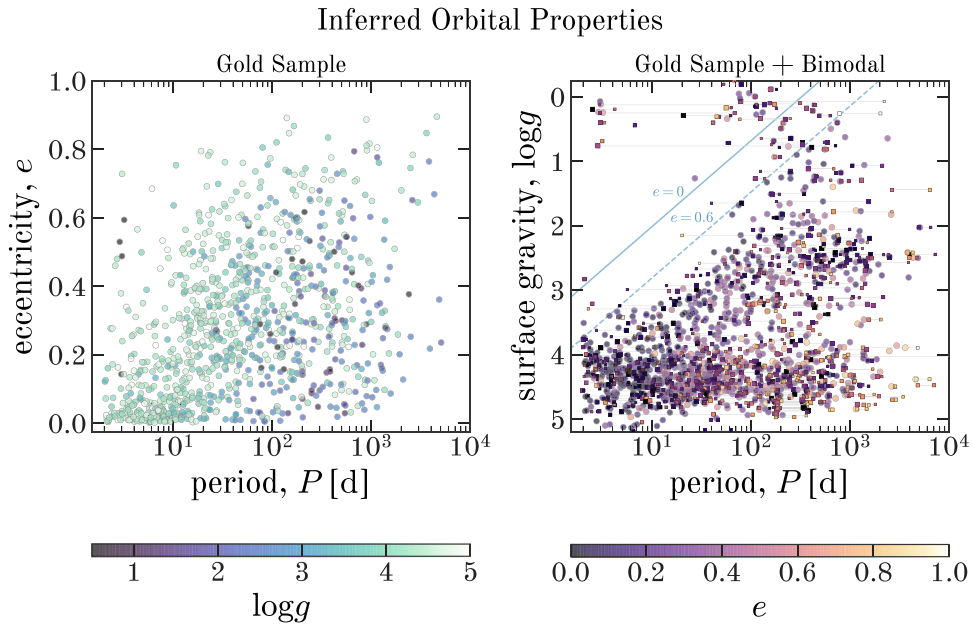
the unknown inclination  $i$  is set to  $90^\circ$ ). Figure 9 (left panel) shows these minimum companion-mass estimates as a function of the STARHORSE primary masses for all sources in the *Gold Sample*. While the uncertainties in these quantities are not shown for most sources, the eight highlighted systems (red markers with error bars) show typical values of the errors on the masses (but note that the errors will be strongly correlated). The two dashed (blue) lines show the approximate hydrogen-burning limit (lower horizontal line), and the upper curve shows the line of equality where the minimum companion mass

is equal to the primary mass. Of these, 95 systems have  $M_{2,\text{min}} < 80 M_J$ . Some of these may be high-inclination stellar-mass systems, but all should be considered brown dwarf candidates. Based on the quality cuts applied to define the parent sample (which should remove sources with blended spectral lines), systems with  $M_{2,\text{min}} > M_1$  should not exist in the sample if the companion is luminous. The 40 systems with  $M_{2,\text{min}} > M_1$  are therefore excellent candidate compact object companions and will be discussed in a separate paper (A. M. Price-Whelan et al. 2020, in preparation).

The right panel of Figure 9 shows the ratio of the primary stellar radius over the (projected) system semimajor axis as a function of (minimum) mass ratio. Here, the curved, dashed line shows an estimate of the Roche radius (Eggleton 1983). Systems above this line are likely interacting. One such system (2M08160493+2858542), indicated by the square (orange) marker in this panel, appears to be strongly photometrically variable in data from the ASAS-SN survey (Shappee et al. 2014; Jayasinghe et al. 2019). However, most other candidate interacting systems do not have ASAS-SN light curves and could instead be followed up with TESS (Ricker et al. 2014).

Figure 10 shows the radial velocity data (black markers)—underplotted (blue lines) with orbits computed from posterior samples—for the four highlighted systems below the  $80 M_J$  line in Figure 9 (left). The left panels show the time series. The right panels show the same data and orbits, but now phase-folded using the MAP period value. The inferred minimum companion masses are indicated in each right panel. These systems were chosen from a vetted subsample of all substellar companion candidates in order to highlight systems with a range of companion masses, eccentricities, and numbers of observations.

Figure 11 shows the same, but for the four highlighted systems above the  $M_{2,\text{min}} = M_1$  line in Figure 9. The companions in the systems shown in the top two rows are just barely consistent with being high-mass neutron stars (e.g., Cromartie et al. 2020), but the systems shown in the bottom



**Figure 7.** Left: inferred orbital periods,  $P$ , and eccentricities,  $e$ , for *Gold Sample* systems with unimodal samplings. Markers are colored by surface gravity,  $\log g$ . Right: similar to left panel, but showing orbital periods and APOGEE  $\log g$  measurements. Circular markers again indicate the 1032 sources with unimodal samplings, and square markers indicate the 127 sources with bimodal samplings, where the mean of each mode is plotted and connected by a horizontal gray line and the relative sizes of the square markers indicate the fraction of the samples that lie in each mode. Diagonal (blue) lines show the orbital period of a  $0.3 M_{\odot}$  companion with an orbital pericentric radius equal to the surface size of a  $1.1 M_{\odot}$  star with the given surface gravity and labeled eccentricity,  $e$ .

two rows contain candidate noninteracting black hole companions (these are discussed in more detail in the companion paper; A. M. Price-Whelan et al. 2020, in preparation). The recently discovered noninteracting black hole–giant star system (that made use of APOGEE data; Thompson et al. 2019) does not appear in our candidates, because it only has three visit spectra with APOGEE data alone—and thus does not pass the strict cuts we used to construct the *Gold Sample*.

#### 6.4. Hierarchical Inference of the Eccentricity Distribution

Most of the results highlighted above make use of the catalogs created from defining selections on the posterior samplings generated with *The Joker*. However, the real power in the individual system posterior samplings is that they enable further *hierarchical* modeling of binary-star population properties without having to make hard cuts on the samples.

In performing a hierarchical inference, the idea is to replace the rigid priors over per-source parameters (i.e., the priors we used to do the samplings described in Section 3) with a parameterized prior that represents the population of binaries. After defining hyperpriors over the hyperparameters of the population model, inference of the population properties is equivalent to, e.g., generating posterior samplings over the hyperparameters. In future work, we will use the per-source samplings generated in this work to construct a joint model for the period, eccentricity, and mass ratio distributions of binary stars as a function of stellar parameters, but this will also require modeling the sample selection function and our detection efficiency. While the full hierarchical inference is out of scope for this article, here we demonstrate how this could be done using a simpler, toy problem: namely, inferring the eccentricity distribution of long-period binary stars using a parametric model.

We first select all 53,790 sources with samplings from *The Joker* (or subsequent MCMC) with  $4000 \text{ K} < T_{\text{eff}} < 7000 \text{ K}$

and  $-2 < [\text{M}/\text{H}] < 0.5$  in order to select main-sequence, FGK-type stars. We then only keep 8599 sources that have  $>128$  samples with  $P > 100$  days and  $K > 1 \text{ km s}^{-1}$ , of which 158 sources have unimodal samplings (i.e., generated with MCMC). The second criterion (on velocity semi-amplitude,  $K$ ) would not be necessary if we instead constructed a mixture model with components to represent stars with companions and the background population, but here, for simplicity, we instead just make a cut on the posterior samplings. For each  $j$  source in this sample of binaries with FGK-type primary stars, we then have  $M_j$  (up to 512) samples in orbital eccentricity,  $e_{jm}$ . We parameterize the eccentricity distribution using a beta distribution,  $B(a, b)$ , and use the importance-sampling trick to reweight the  $e_{jm}$  samples to infer the hyperparameters  $(a, b)$  (Hogg et al. 2010).

In detail, we would like to evaluate or generate samples from the posterior probability distribution over the hyperparameters of the eccentricity distribution

$$p(a, b|D) \propto p(D|a, b) p(a, b), \quad (9)$$

where  $D$  represents all visit data for all sources, i.e.,

$$p(D|a, b) = \prod_j p(D_j|a, b), \quad (10)$$

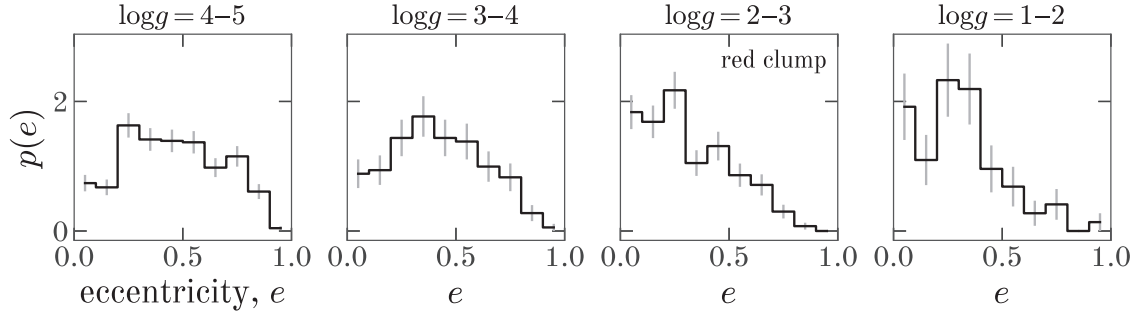
and  $(a, b)$  are the parameters of our assumed beta distribution model. There is no obvious way to evaluate this posterior probability—especially the likelihood term—given the hyperparameters. However, recall that we have samplings from the per-source posterior distributions over eccentricity,

$$p(e_j|D_j, \alpha_0) \propto p(D_j|e_j) p(e_j|\alpha_0) \quad (11)$$

$$e_{jm} \sim p(e|D_j, \alpha_0), \quad (12)$$

where  $\alpha_0$  is meant to represent the parameters of our assumed “interim” prior over eccentricity that we used to generate the per-source samplings (see Table 1), and here “ $\sim$ ” means “is



Gold Sample Systems with  $P > 50$  d

**Figure 8.** Histograms of MAP eccentricity values for systems with MAP  $P > 50$  days. Each panel contains systems with primary-star surface gravities indicated in the title.

sampled from.” Through math that is explained in more detail in other work (e.g., Hogg et al. 2010; Price-Whelan et al. 2018), it is determined that the hierarchical likelihood can be written as a sum over the ratio of probabilities

$$p(D_j|a, b) \approx \frac{\mathcal{Z}}{M_j} \sum_k \frac{p(e_{jm}|a, b)}{p(e_{jm}|\alpha_0)}, \quad (13)$$

where  $\mathcal{Z}$  is a normalization constant. In practice, we evaluate the hierarchical log-likelihood,  $\ln p(D_j|a, b)$ , as

$$\ln p(D|a, b) = \sum_j [\log \text{sumexp}(\ln p(e_{jm}|a, b) - \ln p(e_{jm}|\alpha_0) - \ln M_j)]. \quad (14)$$

With a method for evaluating the hierarchical likelihood (Equation (9)), we now just need to specify prior probability distributions over the hyperparameters of the beta distribution ( $a, b$ ). For each parameter, we use a uniform distribution over the domain (0.1, 10). We implement this model, including the sums over the eccentricity samples, within the context of a `pymc3` model and use the built-in NUTS sampler to generate posterior samples in the parameters of the eccentricity distribution. We run the sampler for 1000 steps to tune, then run four chains in parallel, each for an additional 2000 steps. We assess convergence again by computing the Gelman-Rubin statistic,  $\hat{R}$ , and find that all parameters have converged samplings at the end of our run.

From these posterior samples, we find  $a = 1.749 \pm 0.001$  and  $b = 2.008 \pm 0.001$ . Figure 12 shows the inferred eccentricity distribution (black curve, left panel) and the corresponding cumulative distribution function (right panel), along with some other eccentricity distributions from the literature: for example, the (global) exoplanet eccentricity distribution from Kipping (2013), the (theoretical) eccentricity distribution for a thermalized population of binaries (Jeans 1919), and a uniform distribution. At long periods, binary-star systems seem to have moderate eccentricities that disfavor circular or very eccentric values. This is in agreement with past studies that have focused on nearby samples of solar-type stars with smaller sample sizes (e.g., Duquennoy & Mayor 1991; Raghavan et al. 2010).

Future analyses should assess the impact of selection effects and detection efficiency on the inferred eccentricity distribution using these data. We also emphasize that the extremely precise constraints we obtain on the parameters of the beta distribution imply that we have sufficient data to complexify the model,

either by using nonparametric forms for the eccentricity distribution to move away from rigid models or by parameterizing variations in the distribution with stellar parameters. What we have shown here is meant to be a demonstration of hierarchical inference that utilizes the per-source posterior samplings released with this paper. We do not consider the results of this inference to be one of the key astrophysical results of this paper.

## 7. Discussion, Caveats, and Limitations

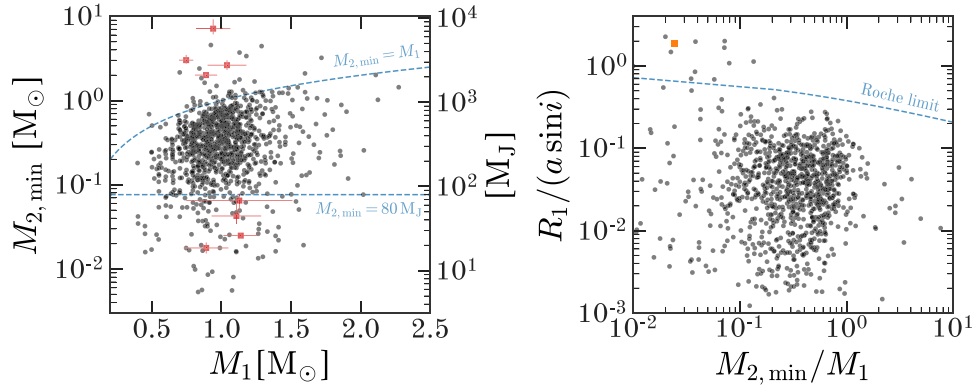
In what follows, we discuss some important caveats and considerations for interpreting the results and using the catalogs described in this article.

### 7.1. Binary Fraction Trends with Stellar Properties

We find a number of interesting trends in the binary fraction (or actually, the fraction of stars with binaries detected in this sample) as a function of stellar parameters and chemical composition that have far-reaching implications. For one, the observed binary fraction increases rapidly with decreasing metallicity, as also noted recently by El-Badry & Rix (2019) and Moe et al. (2019). Beyond the implications discussed in Moe et al. (2019), this also motivates appropriate investigation into how a large binary fraction could influence inferred properties of dwarf galaxies and globular clusters. For example, stellar binarity impacts velocity dispersion measurements, which then contaminate estimates of dark matter masses in these systems (e.g., Aaronson & Olszewski 1987; Kouwenhoven & de Grijs 2008; Martinez et al. 2011; Spencer et al. 2017, 2018; Minor et al. 2019), but also impact the long-term stability and dynamical evolution of compact stellar systems (e.g., Hut et al. 1992; Sigurdsson & Phinney 1993). A large binary fraction also impacts the chemical evolution of these systems (e.g., Eldridge et al. 2008; Eldridge & Stanway 2009).

We caution, however, that any real understanding of the binary fraction (in an absolute sense) from the *catalogs* produced here will require models for both the APOGEE selection function and for our detection efficiency. Still, we clearly now have samples of binary stars that are sufficiently large to begin placing strong constraints on theories of binary-star formation and evolution, as well as their impact on Galactic stellar populations.

## Gold Sample: Primary and Companion Masses



**Figure 9.** Left: minimum companion masses,  $M_{2,\min}$ , as a function of primary mass,  $M_1$ , for all stars in the *Gold Sample*. Upper dashed (blue) line shows the line of equality where  $M_{2,\min} = M_2$ . Systems near or above this threshold have candidate compact-object companions, as the more massive secondary is fainter than the primary. Lower dashed (blue) line shows the approximate hydrogen-burning limit. Sources below this threshold are candidate substellar objects. Eight highlighted points (red) are shown below in Figures 10 and 11. Right: ratio of the primary stellar radius to the (projected) system orbital semimajor axis as a function of minimum mass ratio. Sources above the dashed (blue) line are likely interacting and may be photometrically variable.

### 7.2. Selection, Completeness, False Positives, and Contamination

As cautioned above, to properly use our catalog of systems for performing statistical tests (for example, to interpret the binary-fraction values in Figure 5 in an absolute sense), it is critical to have estimates of our detection efficiency or completeness as a function of binary orbital parameters and stellar parameters. While we have not provided estimates of the completeness, we do release the full posterior samplings in orbital parameters for *all* sources, along with open source software that could be used to construct these estimates. In general, our completeness will likely be a strong function of velocity semi-amplitude, period, and eccentricity (which imply functions of companion mass, inclination, and separation). It will also depend on mass ratio, as many sources with bright companions will be dropped by the quality cuts imposed on the APOGEE data. As a consequence of this, the systems considered in this work are primarily close-binary systems with intermediate mass ratios.

One way to construct completeness estimates would be to repeat the analysis done in this *Article* using simulated systems with known parameters. The sampler used to generate the posterior orbit samplings is open source and released as a Python package (Price-Whelan et al. 2017; Price-Whelan & Hogg 2019). The pipeline software used to define and analyze the APOGEE parent sample is likewise open source and is also available as a Python package, *hq* (Price-Whelan 2019).

### 7.3. Below-threshold Searches

Another problem with a traditional catalog of detected binary companions (like ours) is that there are many sources for which the hypothesis of a single star is clearly rejected, but not sufficiently strongly rejected for the source to qualify as a reliable binary “detection.” A user who has informative data for a source that would take a subthreshold source above threshold (in terms of our criteria for including a source in the catalog) has no recourse if only given the rigidly thresholded catalog entries. Even for well-detected binary systems, in a traditional catalog of orbital parameters, there is no simple way to combine the results with new data to improve or adjust parameter estimates. The samplings provided here can be used

to solve these problems. They can be used to combine the APOGEE information with new data, without requiring an *ab initio* reanalysis of the original APOGEE data. An example of this kind of use is the following: the posterior samples delivered here from the APOGEE data are declared to be the *prior samples* for a new rejection sampling. In this new rejection sampling, the likelihood we use here is replaced with a likelihood from the new data (say, for example, astrometric data), computed now at each of these new prior samples. The output of this rejection sampling will be a sampling of a posterior pdf that accounts correctly for both the APOGEE data and the new data.

In detail, following the notation in Appendix A, converting the posterior samples released here into posterior samples over the APOGEE data,  $D_1$ , *plus* some new velocity data,  $D_2$ , requires generating samples from the posterior pdf

$$p(\theta|D_1, D_2) \propto p(D_1|\theta) p(D_2|\theta) p(\theta) \quad (15)$$

$$\propto p(D_2|\theta) [p(D_1|\theta) p(\theta)], \quad (16)$$

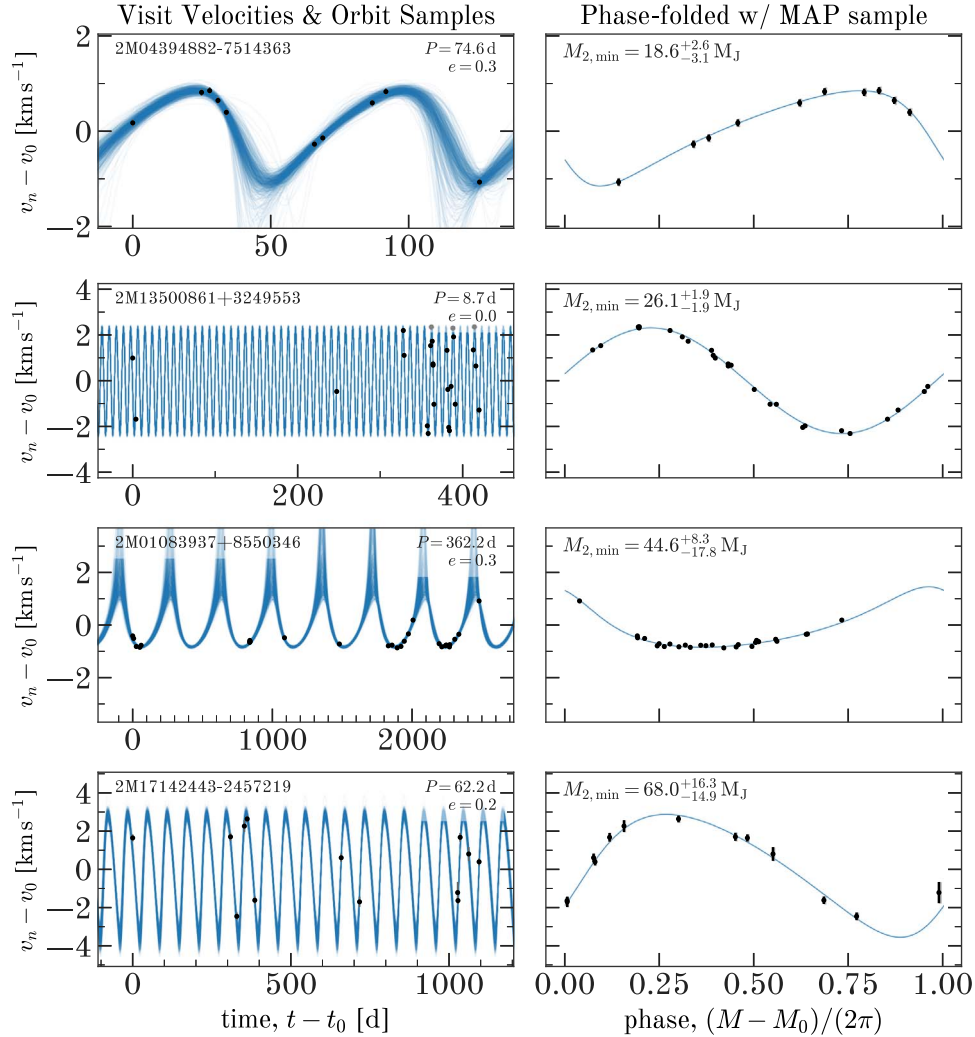
where we have assumed that the new data are independent of the APOGEE data. Note that the terms in brackets in Equation (16) are proportional to the posterior probability of the parameters given only the APOGEE data, i.e., the distribution we have generated samples from using *The Joker*. We can therefore use the posterior samples generated from the *The Joker* with the APOGEE data to rejection-sample using the new marginal likelihood,  $p(D_2|\theta)$ , to generate samples from  $p(\theta|D_1, D_2)$ .

The ability to perform subthreshold searches or add external information is limited by detailed shape of the posterior pdf and the number of samples we deliver (here, 512). If the new data are highly informative, or favor a mode in the posterior pdf that is not well-sampled, the posterior samplings we deliver will not be sufficiently dense to provide support for the updated posterior pdf. This puts limitations on the scope of subthreshold searches, but at least they are possible with these outputs, in some regime of applicability.

### 7.4. Stability of the Posterior Samplings

For any finite-sized radial-velocity data set for a binary-star system, the posterior pdf over Keplerian orbital parameters will

## Example Candidate Brown Dwarf Companions



**Figure 10.** Example binary-star systems from the *Gold Sample* with low-mass companions that are candidate substellar objects. Each row is a different source (indicated in the left panels). Left panels show the raw visit velocity data (black markers) underplotted with orbits computed from the posterior samples for each source (blue lines). Right panels show the same data phase-folded with the MAP orbit sample period and underplotted with an orbit computed from the MAP sample. Minimum companion mass for each system is indicated in the right panels.

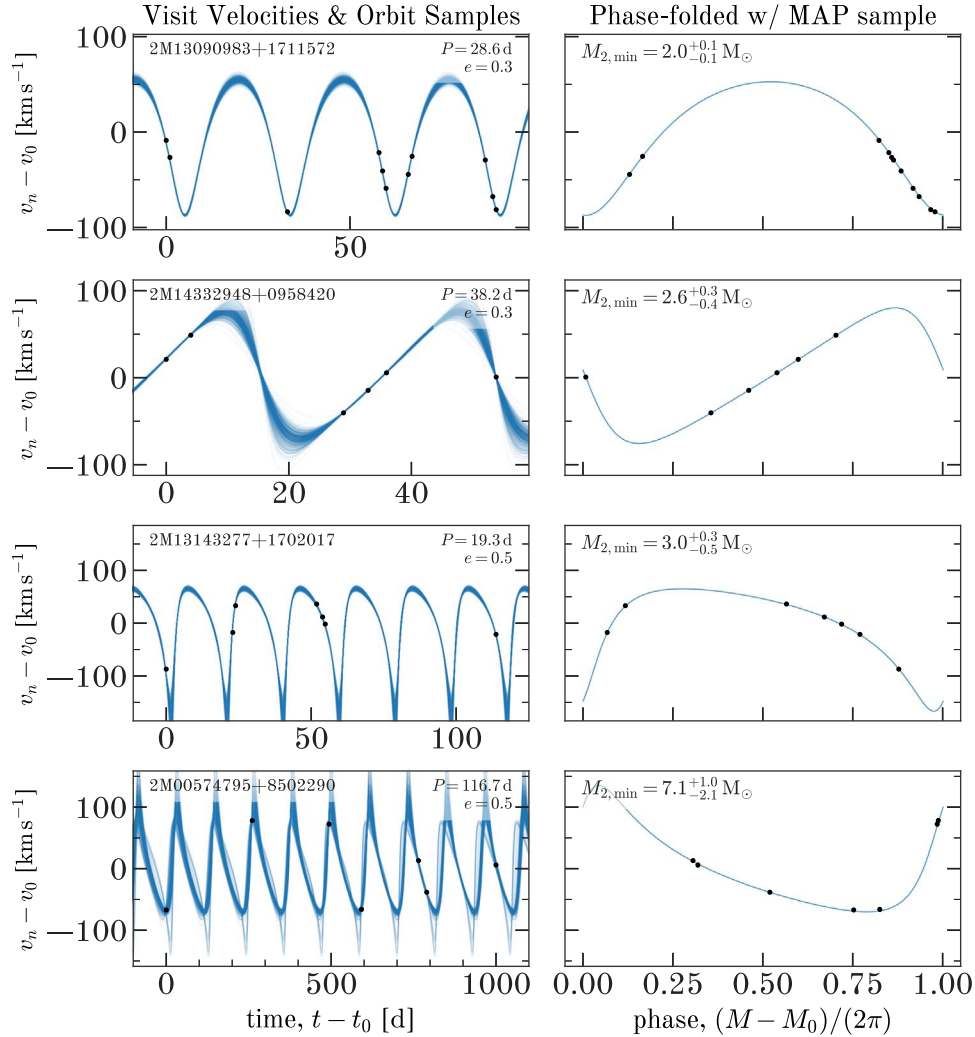
be multimodal. Why, then, do we discuss systems as having “unimodal” or “bimodal” samplings? For a given data set, the relative amplitudes or integrated probabilities in the modes of the posterior pdf can be vastly different, depending on the precision of the data, number of data points, and phase coverage. When the data are very informative, typically one or a few modes dominate, meaning that for a finite-resolution sampling (like those delivered here), the posterior pdf can appear to be unimodal or only slightly multimodal. However, even in these cases, there are many other less significant modes “hidden” by the finite sizes of the samplings. This fact can lead to paradoxical or surprising effects for a given source when a sampling is recomputed after including new data, or when (especially precise) data change slightly, such as after updates to data reduction procedures.

One implication of this is that adding a new velocity measurement for a source, subtly changing all of the velocity values, or even just changing the error bars on the data, can cause the strength of a posterior pdf mode to change dramatically. An example of this behavior can be seen with

the source 2M13090983 + 1711572 (top row of Figure 11), which appears in our *Gold Sample* with an orbital period  $P \approx 28.6$  days. This source also appeared in our previous catalog of systems with unimodal samplings from APOGEE DR14 (Price-Whelan et al. 2018), but with an orbital period  $P \approx 14.3$  days. Though the same 16 visits appear in the APOGEE DR14 and DR16 catalogs for this source, our new, stricter quality cuts for this work (Section 2) filtered out five suspect velocity measurements, and the remaining 11 visit velocities used to generate the samplings used here led to a posterior pdf with different mode amplitudes.

The stability of the posterior samplings described above has some consequences for users of the catalogs and samplings released here. For one, the procedure described above for performing subthreshold searches (Section 7.3) is risky, because for some sources our samplings are not sufficiently dense to fully capture the support of the posterior pdf and its detailed multimodal structure. For these sources, it may be necessary to rerun the sampler on the updated data instead of post-processing our samples. Another consequence is that, as

## Example Candidate Compact Object Companions



**Figure 11.** Same as Figure 10, but for example binary-star systems from the *Gold Sample* with faint, high-mass companions that are candidate compact objects.

the velocity data for a given source evolve in terms of the measured values, error properties, or total number of visits, some results presented here will change in detail. These caveats are important to keep in mind for any users of the samplings and catalogs released with this article.

## 8. Conclusions

Our key results and conclusions are summarized below:

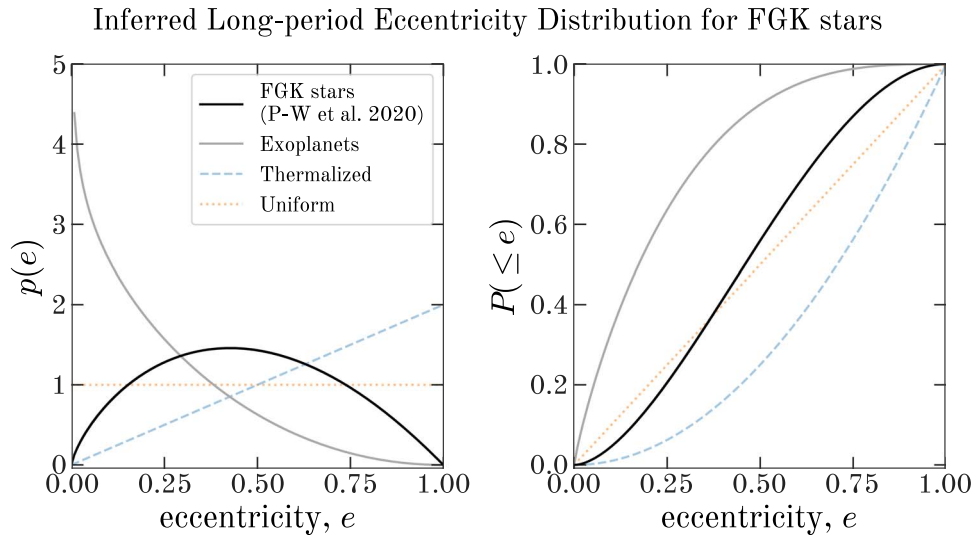
*The close-binary fraction depends on the stellar parameters of the primary star.* Figure 5 shows the observed (uncorrected for completeness) estimate of the close-binary fraction over the color–magnitude diagram, utilizing all 232,495 APOGEE sources used in this work. There is a notable dearth of close companions around the red clump, where companions may have been engulfed when the primary star ascended the upper giant branch. We also rediscover trends in the binary fraction with effective temperature along the main sequence (i.e., stellar mass), and show that the binary fraction on the giant branch depends on the surface gravity (i.e., surface size) of the primary, also indicating that companion engulfment is an important outcome of close-binary stellar evolution.

*The binary fraction is anticorrelated with metallicity.* We find that the binary fraction decreases linearly with bulk metallicity, with a slope of  $-0.1 \text{ dex}^{-1}$ , over the domain  $-1 \lesssim [\text{M}/\text{H}] \lesssim +0.4$ , as shown in Figure 6. This result and the previous result are technically about the observed or detected binary fraction within this sample, not the absolute binary fraction, integrating over all possible amplitudes and mass ratios.

*We detect the clear signature of tidal circularization in field main-sequence and red giant branch stars.* Figure 7 (left) shows inferred periods and eccentricities for systems with uniquely determined orbits in the *Gold Sample*. The abundances of low-eccentricity sources at short periods is a manifestation of tidal circularization, which should operate at longer periods for larger, more convective (i.e., giant-branch) stars.

*We identify 95 candidate brown dwarf companions.* Using simplistic cuts on the *Gold Sample* of high-quality sources with unimodal posterior samplings, we identify candidate brown dwarf companions by selecting sources with median minimum companion mass ( $M_{2,\text{min}}$ ) values below the hydrogen-burning limit,  $M_{2,\text{min}} < 80 M_J$ . Figure 9 (left) shows these sources as points below the horizontal





**Figure 12.** Left: inferred eccentricity distribution for binary-star systems with FGK-type primary stars (black line). We use the posterior samplings for 8599 sources within a toy hierarchical model of the eccentricity distribution, represented as a beta distribution. We also plot common eccentricity distributions from the literature, such as the global exoplanet model from Kipping (2013), shown as a solid, gray line, the (theoretical) thermalized population (Jeans 1919; dashed, blue line), and a uniform distribution (dotted, orange line). Right: same as left panel, but showing cumulative distribution functions instead.

dashed line. We highlight a few of these systems in Figure 10.

We identify 40 candidate noninteracting compact-object companions. Again using simplistic cuts on the *Gold Sample* of high-quality sources with unimodal posterior samplings, we identify candidate compact-object companions by selecting sources with  $M_{2,\min} > M_1$ . Figure 9 (left) shows these sources as points above the upper, curved dashed line. We highlight a few of these systems in Figure 11. As we emphasize above, these systems are just candidates, but are sufficiently interesting that we believe they are worthy of follow-up. We encourage a community effort to continue monitoring these systems to determine their nature.

The binary-star eccentricity distribution is peaked at moderate eccentricities. We execute a toy hierarchical inference using the posterior samplings for  $\sim 8600$  FGK-type main-sequence stars to infer the eccentricity distribution of long-period ( $P > 100$  days), intermediate-mass main-sequence star binary systems. By representing the distribution using a beta distribution, we derive precise posterior constraints on the parameters using this hierarchical model, and find  $a = 1.749 \pm 0.001$  and  $b = 2.008 \pm 0.001$ . Figure 12 shows our inferred eccentricity distribution, indicating that long-period binary-star systems prefer moderate eccentricities.

We release a sample of 20,000 binary-star systems and posterior samplings over orbital parameters for 232,495 APOGEE sources. Finally, we release a catalog of 19,635 high-confidence binary-star systems (Table B1). The majority of these systems have poorly constrained orbital parameters, but we release posterior samplings over these parameters for all 232,495 APOGEE sources, in order to enable other probabilistic inferences with these data. We also define and release a *Gold Sample* containing 1032 systems with high-quality, unimodal posterior samplings that can be used and summarized more simply.

It is a pleasure to thank Borja Anguiano (Virginia), Evan Bauer (KITP), Lars Bildsten (KITP), Dan Foreman-Mackey (Flatiron), Will Farr (Flatiron), Marla Geha (Yale), Adam Jermyn (Flatiron), and Josh Winn (Princeton) for useful comments and discussions. A.P.W. acknowledges support and space from the Max-Planck-Institut für Astronomie during initial work on this project. H.M.L. acknowledges support from NSF grant AST 1616636. P.M.F. acknowledges support for this research from the National Science Foundation (AST-1311835 & AST-1715662). J.G.F.-T. is supported by FONDECYT No. 3180210 and Becas Iberoamérica Investigador 2019, Banco Santander Chile. D.A.G.H. acknowledges support from the State Research Agency (AEI) of the Spanish Ministry of Science, Innovation and Universities (MCIU) and the European Regional Development Fund (FEDER) under grant AYA2017-88254-P. S.H. is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award AST-1801940. T.C.B. acknowledges partial support from grant PHY 14-30152 (Physics Frontier Center/JINA-CEE), awarded by the U.S. National Science Foundation. We thank the anonymous referee for constructive comments that improved this manuscript. Support for this work was provided by NASA through Hubble Fellowship grant #51386.01 awarded to R.L. B. 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.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is [www.sdss.org](http://www.sdss.org).

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto

de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

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

*Software:* Astropy (Astropy Collaboration et al. 2013, 2018), apred (Nidever et al. 2015), ASPCAP (García Pérez et al. 2016), exoplanet (Foreman-Mackey et al. 2019), gala (Price-Whelan 2017), IPython (Pérez & Granger 2007), numpy (Walt et al. 2011), pymc3 (Salvatier et al. 2016), schwimmbad (Price-Whelan & Foreman-Mackey 2017), scipy (Jones et al. 2001), theano (Al-Rfou et al. 2016), thejoker (Price-Whelan et al. 2017; Price-Whelan & Hogg 2019).

## Appendix A

### Update to the Marginal Likelihood Expression for *The Joker*

As noted above (see Section 3.1), the assumptions in Price-Whelan et al. (2017) that led to the simplified form of the marginal likelihood (Equation (11) in Price-Whelan et al. 2017) are not valid. We therefore here rederive the marginal likelihood that forms the basis of the implementation of *The Joker* used in this work.

For each APOGEE source, we have a set of  $N$  radial-velocity measurements (visits)  $v_n$  at times  $t_n$  with uncertainties  $\sigma_n$ . Under the assumption that the source is in a binary-star system, our model for the true radial velocity of the source at any time is given by Equation (2) above. Our goal (as in Price-Whelan et al. 2017) is to analytically marginalize over the linear parameters in Equation (2)—( $K, v_0$ )—under the assumption that the uncertainties on each radial-velocity measurement are Gaussian and independent. To do this, we must write down expressions for the likelihood, and for the prior probability distribution over the parameters that we will be marginalizing over (i.e., the linear parameters). We have already made the assumption that our likelihood has a Gaussian form, so to do this marginalization conveniently and analytically, we additionally assume that the prior pdf also has a Gaussian form. Under these assumptions, the solution to this marginalization is given in D. W. Hogg et al. (2020, in preparation).

To see the relation between the specific problem solved by *The Joker* and the derivation in D. W. Hogg et al. (2020, in preparation), it is convenient to repackage our data and linear

parameters as

$$\mathbf{y} = (v_1, v_2, \dots, v_N)^T \quad (\text{A1})$$

$$\mathbf{x} = (K, v_0)^T \quad (\text{A2})$$

$$\mathbf{C} = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & \sigma_N^2 \end{pmatrix}, \quad (\text{A3})$$

and to assume that the mean and covariance matrix of the prior over the linear parameters are given by  $\boldsymbol{\mu}, \boldsymbol{\Lambda}$ , respectively. Given a set of nonlinear parameters  $\boldsymbol{\theta} = (P, e, \omega, M_0)$ , we will also need to define a design matrix,  $\mathbf{M}$ ,

$$\mathbf{M} = \begin{pmatrix} \zeta(t_1; \boldsymbol{\theta}) & 1 \\ \zeta(t_2; \boldsymbol{\theta}) & 1 \\ \vdots & \vdots \\ \zeta(t_N; \boldsymbol{\theta}) & 1 \end{pmatrix}, \quad (\text{A4})$$

where  $\zeta(\cdot)$  is given by Equation (3). With these assumptions, the marginalized likelihood,  $Q$ , for a source, given nonlinear parameters  $\boldsymbol{\theta}$ , can be written

$$Q = \int d\mathbf{x} \mathcal{N}(\mathbf{y}|\mathbf{M} \cdot \mathbf{x}, \mathbf{C}) \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Lambda}) \quad (\text{A5})$$

$$= \int d\mathbf{x} \mathcal{N}(\mathbf{x}|\mathbf{a}, \mathbf{A}) \mathcal{N}(\mathbf{y}|\mathbf{b}, \mathbf{B}) \quad (\text{A6})$$

$$= \mathcal{N}(\mathbf{y}|\mathbf{b}, \mathbf{B}) \quad (\text{A7})$$

$$\mathbf{b} = \mathbf{M} \cdot \boldsymbol{\mu} \quad (\text{A8})$$

$$\mathbf{B} = \mathbf{C} + \mathbf{M} \cdot \boldsymbol{\Lambda} \cdot \mathbf{M}^T, \quad (\text{A9})$$

where the integral becomes simple because the second normal distribution,  $\mathcal{N}(\mathbf{y}|\mathbf{b}, \mathbf{B})$ , does not depend on the integration variables,  $\mathbf{x}$ , and the integral over the first normal distribution is 1 (D. W. Hogg et al. 2020, in preparation). A final point that is relevant to additional enhancements discussed in Section 3.1 (and is exploited in this work) is to note that  $\boldsymbol{\mu}$  and  $\boldsymbol{\Lambda}$  can depend on the nonlinear parameters.

## Appendix B

### Data Tables

The primary data product released with this article are the posterior samplings generated for each of 232,495 sources in APOGEE DR16. These samplings will be released in the upcoming intermediate SDSS data release “DR16+” (expected in mid-2020). However, we also compute summary information and statistics about these samplings and provide these data in Table B1. We also define a *Gold Sample* of high-quality, uniquely solved binary-star systems (see Section 6.2) and release summary information along with cross-matched data from Gaia DR2 and the STARHORSE catalog of stellar parameters in Table B2.















## ORCID iDs

Adrian M. Price-Whelan  <https://orcid.org/0000-0003-0872-7098>

David W. Hogg  <https://orcid.org/0000-0003-2866-9403>

Hans-Walter Rix  <https://orcid.org/0000-0003-4996-9069>

Rachael L. Beaton  <https://orcid.org/0000-0002-1691-8217>

Hannah M. Lewis  <https://orcid.org/0000-0002-7871-085X>  
 David L. Nidever  <https://orcid.org/0000-0002-1793-3689>  
 Carles Badenes  <https://orcid.org/0000-0003-3494-343X>  
 Rodolfo Barba  <https://orcid.org/0000-0003-1086-1579>  
 Timothy C. Beers  <https://orcid.org/0000-0003-4573-6233>  
 Joleen K. Carlberg  <https://orcid.org/0000-0001-5926-4471>  
 Nathan De Lee  <https://orcid.org/0000-0002-3657-0705>  
 Peter M. Frinchaboy  <https://orcid.org/0000-0002-0740-8346>  
 D. A. García-Hernández  <https://orcid.org/0000-0002-1693-2721>  
 Paul J. Green  <https://orcid.org/0000-0002-8179-9445>  
 Steven R. Majewski  <https://orcid.org/0000-0003-2025-3147>  
 Keivan G. Stassun  <https://orcid.org/0000-0002-3481-9052>  
 Guy S. Stringfellow  <https://orcid.org/0000-0003-1479-3059>  
 Nicholas W. Troup  <https://orcid.org/0000-0003-3248-3097>

## References

- Aaronson, M., & Olszewski, E. 1987, in IAU Symp. 117, Dark Matter in the Universe, ed. J. Kormendy & G. R. Knapp (Dordrecht: D. Reidel), 153
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, *PhRvL*, **116**, 061102
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019, *PhRvX*, **9**, 031040
- Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, *ApJS*, **235**, 42
- Ahumada, R., Allende Prieto, C., Almeida, A., et al. 2019, arXiv:1912.02905
- Al-Rfou, R., Alain, G., Almahairi, A., et al. 2016, arXiv:1605.02688
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, **156**, 123
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33
- Badenes, C., Mazzola, C., Thompson, T. A., et al. 2018, *ApJ*, **854**, 147
- Blanton, M. R., Bershad, M. A., Abolfathi, B., et al. 2017, *AJ*, **154**, 28
- Bowen, I. S., & Vaughan, A. H. J. 1973, *ApOpt*, **12**, 1430
- Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, **823**, 102
- Cottaar, M., Covey, K. R., Meyer, M. R., et al. 2014, *ApJ*, **794**, 125
- Cromartie, H. T., Fonseca, E., Ransom, S. M., et al. 2020, *NatAs*, **4**, 72
- Dotter, A. 2016, *ApJS*, **222**, 8
- Duchêne, G., & Kraus, A. 2013, *ARA&A*, **51**, 269
- Duquennoy, A., & Mayor, M. 1991, *A&A*, **500**, 337
- Eggleton, P. P. 1983, *ApJ*, **268**, 368
- Eggleton, P. P., & Tokovinin, A. A. 2008, *MNRAS*, **389**, 869
- El-Badry, K., & Rix, H.-W. 2019, *MNRAS*, **482**, L139
- El-Badry, K., Rix, H.-W., Tian, H., Duchêne, G., & Moe, M. 2019, *MNRAS*, **489**, 5822
- El-Badry, K., Rix, H.-W., Ting, Y.-S., et al. 2018, *MNRAS*, **473**, 5043
- Eldridge, J. J., Izzard, R. G., & Tout, C. A. 2008, *MNRAS*, **384**, 1109
- Eldridge, J. J., & Stanway, E. R. 2009, *MNRAS*, **400**, 1019
- Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017, *PASA*, **34**, e058
- Fernández-Trincado, J. G., Mennickent, R., Cabezas, M., et al. 2019, *A&A*, **631**, A97
- Foreman-Mackey, D., Czekala, I., Luger, R., et al. 2019, dfm/exoplanet v0.2.3, Zenodo, doi:10.5281/zenodo.1998447
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, **616**, A1
- Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, *A&A*, **595**, A1
- Gao, S., Liu, C., Zhang, X., et al. 2014, *ApJL*, **788**, L37
- García Pérez, A. E., Allende Prieto, C., Holtzman, J. A., et al. 2016, *AJ*, **151**, 144
- Gelman, A., & Rubin, D. B. 1992, *StaSc*, **7**, 457
- Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, *AJ*, **131**, 2332
- Hekker, S., Snellen, I. A. G., Aerts, C., et al. 2008, *A&A*, **480**, 215
- Hogg, D. W., Myers, A. D., & Bovy, J. 2010, *ApJ*, **725**, 2166
- Holtzman, J. A., Hesselquist, S., Shetrone, M., et al. 2018, *AJ*, **156**, 125
- Homan, M. D., & Gelman, A. 2014, *J. Mach. Learn. Res.*, **15**, 1593
- Hut, P., McMillan, S., Goodman, J., et al. 1992, *PASP*, **104**, 981
- Ivanova, N., Justham, S., Chen, X., et al. 2013, *A&ARv*, **21**, 59
- Jayasinghe, T., Stanek, K. Z., Kochanek, C. S., et al. 2019, *MNRAS*, **485**, 961
- Jeans, J. H. 1919, *MNRAS*, **79**, 408
- Jenkins, J. S., Díaz, M., Jones, H. R. A., et al. 2015, *MNRAS*, **453**, 1439
- Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python v1.3.3, <http://www.scipy.org/>
- Kipping, D. M. 2013, *MNRAS*, **434**, L51
- Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv:1711.03234
- Kouwenhoven, M. B. N., & de Grijs, R. 2008, *A&A*, **480**, 103
- Lindgren, L., Hernández, J., Bombrun, A., et al. 2018, *A&A*, **616**, A2
- Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, *AJ*, **154**, 94
- Martinez, G. D., Minor, Q. E., Bullock, J., et al. 2011, *ApJ*, **738**, 55
- Meibom, S., & Mathieu, R. D. 2005, *ApJ*, **620**, 970
- Minor, Q. E., Pace, A. B., Marshall, J. L., & Strigari, L. E. 2019, *MNRAS*, **487**, 2961
- Moe, M., & Di Stefano, R. 2017, *ApJS*, **230**, 15
- Moe, M., & Kratter, K. M. 2018, *ApJ*, **854**, 44
- Moe, M., Kratter, K. M., & Badenes, C. 2019, *ApJ*, **875**, 61
- Nidever, D. L., Holtzman, J. A., Allende Prieto, C., et al. 2015, *AJ*, **150**, 173
- Nocedal, J., & Wright, S. J. 2006, Numerical Optimization (2nd ed.; New York: Springer)
- Paxton, B., Bildsten, L., Dotter, A., et al. 2011, *ApJS*, **192**, 3
- Paxton, B., Cantiello, M., Arras, P., et al. 2013, *ApJS*, **208**, 4
- Paxton, B., Marchant, P., Schwab, J., et al. 2015, *ApJS*, **220**, 15
- Pérez, F., & Granger, B. E. 2007, *CSE*, **9**, 21
- Price-Whelan, A. 2019, adrn/hq: The Version Used to Run on APOGEE DR16 v1.0, Zenodo, doi:10.5281/zenodo.3596084
- Price-Whelan, A., Breivik, K., D'Orazio, D., et al. 2019, *BAAS*, **51**, 206
- Price-Whelan, A., & Hogg, D. W. 2019, adrn/thejoker v1.0, Zenodo, doi:10.5281/zenodo.3596084
- Price-Whelan, A. M. 2017, *JOSS*, **2**, 388
- Price-Whelan, A. M., & Foreman-Mackey, D. 2017, *JOSS*, **2**, 357
- Price-Whelan, A. M., & Goodman, J. 2018, *ApJ*, **867**, 5
- Price-Whelan, A. M., Hogg, D. W., Foreman-Mackey, D., & Rix, H.-W. 2017, *ApJ*, **837**, 20
- Price-Whelan, A. M., Hogg, D. W., Rix, H.-W., et al. 2018, *AJ*, **156**, 18
- Queiroz, A. B. A., Anders, F., Chiappini, C., et al. 2019, arXiv:1912.09778
- Raghavan, D., McAlister, H. A., Henry, T. J., et al. 2010, *ApJS*, **190**, 1
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, *Proc. SPIE*, **9143**, 914320
- Rix, H.-W., Ting, Y.-S., Sippel, A., et al. 2019, *BAAS*, **51**, 104
- Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. 2016, *PeerJ Comput. Sci.*, **2**, e55
- Shappee, B. J., Prieto, J. L., Grupe, D., et al. 2014, *ApJ*, **788**, 48
- Sigurdsson, S., & Phinney, E. S. 1993, *ApJ*, **415**, 631
- Spencer, M. E., Mateo, M., Olszewski, E. W., et al. 2018, *AJ*, **156**, 257
- Spencer, M. E., Mateo, M., Walker, M. G., et al. 2017, *AJ*, **153**, 254
- Thompson, T. A., Kochanek, C. S., Stanek, K. Z., et al. 2019, *Sci*, **366**, 637
- Troup, N. W., Nidever, D. L., De Lee, N., et al. 2016, *AJ*, **151**, 85
- Walt, S. v. d., Colbert, S. C., & Varoquaux, G. 2011, *CSE*, **13**, 22
- Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, *PASP*, **131**, 055001
- Zahn, J. P. 1977, *A&A*, **500**, 121
- Zamora, O., García-Hernández, D. A., Allende Prieto, C., et al. 2015, *AJ*, **149**, 181
- Zasowski, G., Cohen, R. E., Chojnowski, S. D., et al. 2017, *AJ*, **154**, 198
- Zasowski, G., Johnson, J. A., Frinchaboy, P. M., et al. 2013, *AJ*, **146**, 81
- Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, *RAA*, **12**, 723