



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## The Close Binary Fraction as a Function of Stellar Parameters in APOGEE

**Citation for published version:**

Mazzola, CN, Badenes, C, Moe, M, Koposov, SE, Kounkel, M, Kratter, K, Covey, K, Walker, MG, Thompson, TA, Andrews, B, Freeman, PE, Anguiano, B, Carlberg, JK, Lee, NMD, Frinchaboy, PM, Lewis, HM, Majewski, S, Nidever, D, Nitschelm, C, Price-Whelan, AM, Roman-Lopes, A, Stassun, KG & Troup, NW 2020, 'The Close Binary Fraction as a Function of Stellar Parameters in APOGEE: A Strong Anti-Correlation With Abundances', *Monthly Notices of the Royal Astronomical Society*.  
<https://doi.org/10.1093/mnras/staa2859>

**Digital Object Identifier (DOI):**

[10.1093/mnras/staa2859](https://doi.org/10.1093/mnras/staa2859)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Monthly Notices of the Royal Astronomical Society

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# The Close Binary Fraction as a Function of Stellar Parameters in APOGEE: A Strong Anti-Correlation With $\alpha$ Abundances

Christine N. Mazzola,<sup>1\*</sup> Carles Badenes,<sup>1</sup> Maxwell Moe,<sup>2</sup> Sergey E. Koposov,<sup>3,4</sup> Marina Kounkel,<sup>5</sup> Kaitlin Kratter,<sup>2</sup> Kevin Covey,<sup>5</sup> Matthew G. Walker,<sup>6</sup> Todd A. Thompson,<sup>7</sup> Brett Andrews,<sup>1</sup> Peter E. Freeman,<sup>8</sup> Borja Anguiano,<sup>9</sup> Joleen K. Carlberg,<sup>10</sup> Nathan M. De Lee,<sup>11,12</sup> Peter M. Frinchaboy,<sup>13</sup> Hannah M. Lewis,<sup>9</sup> Steven Majewski,<sup>9</sup> David Nidever,<sup>14,15</sup> Christian Nitschelm,<sup>16</sup> Adrian M. Price-Whelan,<sup>17</sup> Alexandre Roman-Lopes,<sup>18</sup> Keivan G. Stassun,<sup>12</sup> Nicholas W. Troup<sup>19</sup>

<sup>1</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>2</sup>Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA

<sup>3</sup>Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

<sup>4</sup>Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

<sup>5</sup>Department of Physics and Astronomy, Western Washington University, 516 High St., Bellingham, WA 98225, USA

<sup>6</sup>McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

<sup>7</sup>Department of Astronomy, The Ohio State University, 140 W. 18th Ave., Columbus, OH 43210, USA., and Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 W. Woodruff Ave., Columbus, OH 43210, USA.

<sup>8</sup>Department of Statistics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

<sup>9</sup>Department of Astronomy, University of Virginia, 530 McCormick Road, Charlottesville, VA 22904-4325, USA

<sup>10</sup>Space Telescope Science Institute, 3700 San Martin Dr. Baltimore MD 21218

<sup>11</sup>Department of Physics, Geology, and Engineering Technology, Northern Kentucky University, Highland Heights, KY 41099

<sup>12</sup>Department of Physics & Astronomy, Vanderbilt University, Nashville, TN 37235

<sup>13</sup>Department of Physics & Astronomy, Texas Christian University, Fort Worth, TX 76129, USA

<sup>14</sup>Department of Physics, Montana State University, P.O. Box 173840, Bozeman, MT 59717-3840

<sup>15</sup>NSF's National Optical-Infrared Astronomy Research Laboratory, 950 North Cherry Ave, Tucson, AZ 85719

<sup>16</sup>Centro de Astronomía (CITEVA), Universidad de Antofagasta, Avenida Angamos 601, Antofagasta 1270300, Chile

<sup>17</sup>Center for Computational Astrophysics, Flatiron Institute, Simons Foundation, 162 Fifth Avenue, New York, NY 10010, USA

<sup>18</sup>Departamento de Física, Facultad de Ciencias, Universidad de La Serena, Cisternas 1200, La Serena, Chile

<sup>19</sup>Department of Physics, Salisbury University, Salisbury, MD 21801

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

We use observations from the APOGEE survey to explore the relationship between stellar parameters and multiplicity. We combine high-resolution repeat spectroscopy for 41,363 dwarf and subgiant stars with abundance measurements from the APOGEE pipeline and distances and stellar parameters derived using *Gaia* DR2 parallaxes from Sanders & Das (2018) to identify and characterise stellar multiples with periods below 30 years, corresponding to  $\Delta RV_{\max} \gtrsim 3 \text{ km s}^{-1}$ , where  $\Delta RV_{\max}$  is the maximum APOGEE-detected shift in the radial velocities. Chemical composition is responsible for most of the variation in the close binary fraction in our sample, with stellar parameters like mass and age playing a secondary role. In addition to the previously identified strong anti-correlation between the close binary fraction and  $[\text{Fe}/\text{H}]$  we find that high abundances of  $\alpha$  elements also suppress multiplicity at most values of  $[\text{Fe}/\text{H}]$  sampled by APOGEE. The anti-correlation between  $\alpha$  abundances and multiplicity is substantially steeper than that observed for Fe, suggesting C, O, and Si in the form of dust and ices dominate the opacity of primordial protostellar disks and their propensity for fragmentation via gravitational stability. Near  $[\text{Fe}/\text{H}] = 0$  dex, the bias-corrected close binary fraction ( $a < 10$  au) decreases from  $\approx 100$  per cent at  $[\alpha/\text{H}] = -0.2$  dex to  $\approx 15$  per cent near  $[\alpha/\text{H}] = 0.08$  dex, with a suggestive turn-up to  $\approx 20$  per cent near  $[\alpha/\text{H}] = 0.2$ . We conclude that the relationship between stellar multiplicity and chemical composition for sun-like dwarf stars in the field of the Milky Way is complex, and that this complexity should be accounted for in future studies of interacting binaries.

**Key words:** binaries: close – binaries: spectroscopic – stars: abundances

## 1 INTRODUCTION

The accurate characterisation of stellar multiplicity remains a key priority in stellar astrophysics. Interacting binaries, defined as those that are close enough to transfer mass and experience significant deviations from single stellar evolution, are responsible for a wide array of phenomena in time-domain astronomy. These include, but are not limited to, cataclysmic variables, novae, all Type Ia and many core-collapse supernovae, high- and low-mass X-ray binaries, and the majority of gravitational wave sources in the LIGO and LISA passbands (for a review, see [De Marco & Izzard 2017](#)). The formation rates of these sources in a variety of stellar populations are determined by the initial conditions for stellar multiplicity: the multiplicity fraction, and the distribution of periods, mass ratios, and eccentricities. It is now clear that these fundamental statistics of stellar multiplicity are strong functions of stellar properties like mass and composition, and that they are not independent of each other (see [Duchêne & Kraus 2013](#); [Moe & Di Stefano 2017](#), for reviews). This realisation sets the stage for the challenging observational problem of identifying and characterising all the relevant correlations between stellar properties and multiplicity statistics in the field.

Fortunately, modern spectroscopic surveys are well suited to this task. The Apache Point Observatory Galactic Evolution Experiment 2 (APOGEE-2, [Majewski et al. 2017](#)), one of the constituent surveys in the fourth instalment of the Sloan Digital Sky Survey (SDSS-IV, [Gunn et al. 2006](#); [Blanton et al. 2017](#)), collected multi-epoch data for 437,485 stars with its high-resolution ( $R \sim 22,500$ ) multiplexed infrared spectrograph as part of Data Release 16 (DR16) ([Wilson et al. 2019](#)). This constitutes the most comprehensive sample of the detailed compositions of Milky Way stars to date. The APOGEE Stellar Parameter and Chemical Abundances Pipeline (ASPCAP, [García Pérez et al. 2016](#); [Jönsson et al. 2020](#)) has measured reliable stellar parameters for each of these stars, including calibrated abundances of as many as 20 or more elements, and precise radial velocities (RVs, [Nidever et al. 2015](#)) for each individual visit spectra. Leveraging the time-domain component of the survey, [Badenes et al. \(2018\)](#) identified a strong anti-correlation between the multiplicity fraction at short periods and stellar metallicity in DR13 of APOGEE ([Albaret et al. 2017](#)) - see also [Grether & Lineweaver \(2007\)](#); [Raghavan et al. \(2010\)](#); [Gao et al. \(2014, 2017\)](#); [Yuan et al. \(2015\)](#); [El-Badry & Rix \(2018\)](#); [El-Badry et al. \(2018b\)](#); [Pawlak et al. \(2019\)](#); [Liu \(2019\)](#); [Price-Whelan et al. \(2020\)](#); [Miglio et al. \(2020\)](#). Further analysis by [Moe et al. \(2019\)](#) established that the metal-poor ( $[\text{Fe}/\text{H}] \sim -1$  dex) dwarfs observed by APOGEE are  $\sim 4$  times more likely to have short-period ( $P \lesssim 30$  yr, or  $a \lesssim 10$  AU) binary companions than the metal-rich ( $[\text{Fe}/\text{H}] \sim 0.5$  dex) dwarfs, and that this trend likely extends to the lower metallicities characteristic of halo stars. This anti-correlation has now been firmly established using large numbers of sparsely sampled RV curves ([Gao et al. 2014, 2017](#); [Badenes et al. 2018](#); [Price-Whelan et al. 2020](#)), smaller numbers of systems with known orbital periods (from both complete orbital solutions and eclipses, [Moe et al. 2019](#)), and common-proper-motion binaries with projected separations measured by Gaia ([El-Badry & Rix 2018](#)). This has profound implications for the rates of interacting binaries in

the Universe (e.g. [Paczynski 1971](#); [Iben Jr. & Tutukov 1984](#); [Suda et al. 2013](#); [de Mink & Belczynski 2015](#); [De Marco & Izzard 2017](#); [Breivik et al. 2019](#); [Stanway et al. 2020](#)) and for the physics of star formation and disk fragmentation (e.g., [Kratler et al. 2010a](#); [Duchêne & Kraus 2013](#); [Moe & Di Stefano 2017](#); [Moe & Kratter 2018](#); [Kounkel et al. 2019](#)).

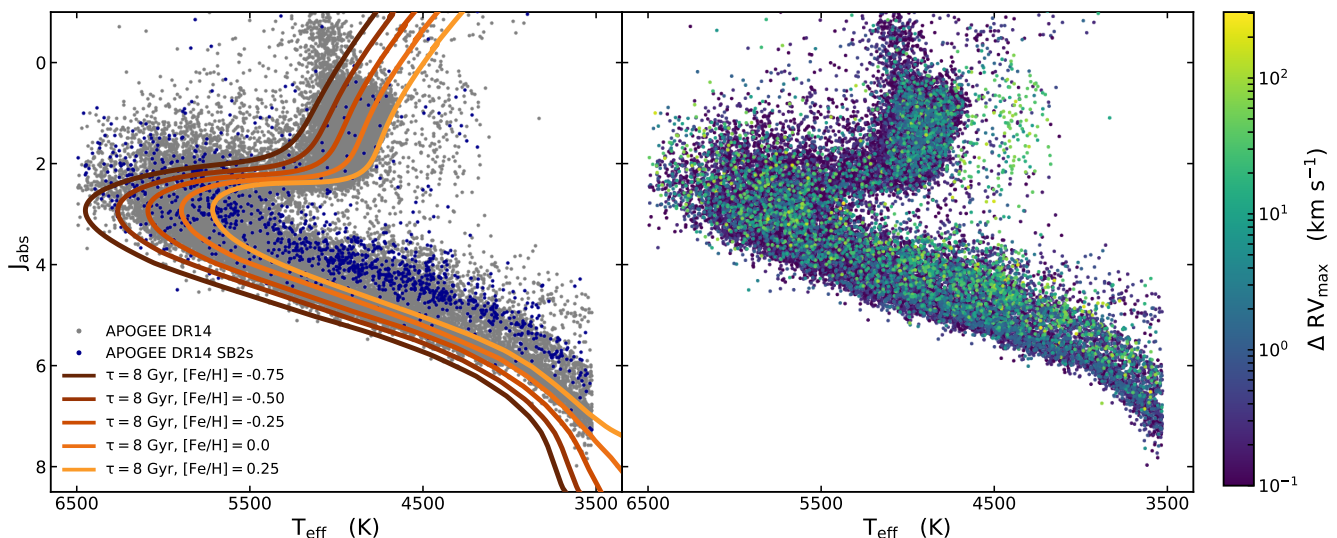
Here we continue to explore the relationship between stellar parameters and stellar multiplicity using public data from APOGEE, complemented by Gaia Data Release 2. In order to avoid the details of the interplay between stellar evolution and multiplicity described by [Badenes et al. \(2018\)](#), we restrict our analysis to dwarf and subgiant stars. We examine a wide array of stellar parameters, paying special attention to the abundances of  $\alpha$  elements. In Section 2, we detail our sample selection and method to account for double-lined spectroscopic binaries (SB2s). Section 3.1 describes our completeness corrections. In Section 3.2 we describe the broad view of the relationship between stellar multiplicity and stellar parameters in our sample. In Section 3.3, we examine in more detail the impact of chemical composition on stellar multiplicity. We discuss our results in Section 4 and summarise in Section 5.

## 2 SAMPLE SELECTION

The DR14 version of the APOGEE `allStar` file contains spectral parameters for 277,371 entries ([Abolfathi et al. 2018](#); [Holtzman et al. 2018](#); [Jönsson et al. 2018](#)). We first note that there are only 258,475 unique APOGEE IDs amongst these 277,371 entries. The duplicate entries are a result of a star being observed in different fibre plugplates with different field centres, which are not automatically combined by the pipeline. Each `allStar` entry corresponds to a combined spectrum and its measured stellar parameters, and is uniquely described by an APOGEE ID and a field location ID.

From APOGEE DR14, we removed stars with the `STAR_BAD` flag set in the ASPCAP bitmask ([Holtzman et al. 2015](#)) and those targeted as telluric calibrators (bit 9 in both the `apogee_target2` and `apogee2_target2` masks; [Zasowski et al. 2013, 2017](#)). Star cluster members (bit 9 in `apogee_target1` and `apogee2_target1` and bit 10 in `apogee_target2` and `apogee2_target2`) and commissioning stars (bit 1 in `STARFLAG`, [Holtzman et al. 2015](#)) were removed as well. Finally, we required acceptable ( $\neq -9999$ , APOGEE's default for a bad value) uncalibrated effective temperatures ( $T_{\text{eff}}$ ) and surface gravities ( $\log(g)$ ) to maximise our ability to distinguish dwarfs from giants in DR14. As noted in [Holtzman et al. \(2018\)](#), dwarfs in APOGEE DR14 do not have calibrated  $\log(g)$  values, so we do not make cuts on the calibrated parameters. In order to estimate the dereddened  $JHK_s$  magnitudes, we used the value of  $A_K$  adopted for targeting purposes (`AK_TARG`, [Zasowski et al. 2013, 2017](#)).

For each APOGEE ID/location ID combination, we identified the individual visits from the `allVisit` file that were included in its combined APOGEE spectrum (the `VISITS_PK` indices, [Holtzman et al. 2015](#); [Nidever et al. 2015](#)). We imposed an additional quality cut, requiring two or more of these visits to have a  $S/N \geq 40$ . If a star had duplicate APOGEE IDs, all of the acceptable visit RVs from its vari-



**Figure 1.** Left panel: An HR diagram for our sample with APOGEE DR14 uncalibrated  $T_{\text{eff}}$  and the absolute 2MASS  $J$  magnitude, calculated using the Sanders & Das (2018) distance estimates. The grey points are for our main sample, and the dark blue are for objects identified as likely SB2s. The coloured lines are MIST isochrone tracks for  $\tau = 8$  Gyr and various metallicities. Right panel: The same HR diagram but with a colourbar on  $\Delta RV_{\text{max}}$ . Points with  $\Delta RV_{\text{max}} \geq 1 \text{ km s}^{-1}$  are plotted on top for clarity.

ous plugplate fields were concatenated. This meant that objects with at least one acceptable visit in two or more fields could be included. For these stars, we averaged any duplicate stellar parameters with valid values from the pipeline.

Both the APOGEE data reduction pipeline (Nidever et al. 2015) and ASPCAP (García Pérez et al. 2016) assume that each source can be modelled by a single stellar spectrum. Stellar companions within the range of mass ratios  $q = M_2/M_1$  that can make a significant contribution to the observed flux (double-lined spectroscopic binaries or SB2s) can therefore introduce biases in the spectral fits; see El-Badry et al. (2018a) for a discussion. To identify these stars, we examined the APOGEE cross-correlation functions (CCFs), following the procedure described in Kounkel et al. (2019). Two approaches were considered: using CCFs that APOGEE provides natively in its data releases, and recalculating the CCFs by cross-matching the spectra with the best-fit PHOENIX synthetic spectrum (Husser et al. 2013), using the reported  $RV_{\text{TEFF}}$  and  $RV_{\text{LOGG}}$  parameters. In most cases, the deconvolution of multiple components from the CCFs occurred in the same sources, with comparable RVs. In this way, we identified 3656 likely SB2s within APOGEE DR14, of which 1512 were in our quality-cut sample. From the CCFs for these stars, we determined the RV of the highest peak at each epoch and used this as a more reliable estimate for the RV of the photometric primary. After applying our quality cuts, we were left with 1495 likely SB2s, which we kept in our sample with spectral parameters from APOGEE/ASPCAP and RVs from our CCF analysis. Details about the downloadable tables of these SB2s are available in Appendix C.

In a final step, we restricted our sample to  $\log(g/\text{cm s}^{-2}) \geq 3.25$ ,  $[\text{Fe}/\text{H}] \geq -1.0$  dex. This simple cut in  $\log(g)$  will not purely select dwarfs, but it is sufficient for our purposes in eliminating most stars on the red giant branch. We also imposed an additional requirement of acceptable values ( $\neq$

$-9999$ ) for  $[\alpha/\text{Fe}]$ ,  $[\alpha/\text{H}]$ ,  $[\text{O}/\text{H}]$ ,  $[\text{Mg}/\text{H}]$ , and  $[\text{Si}/\text{H}]$ . This left us with 41,363 unique APOGEE targets, 1278 of which were identified as SB2s, and 3896 (131 SB2s, 3765 non-SB2s) had duplicate entries and so their stellar parameters were averaged. The fraction of SB2s in this sample is  $1278/41,363 = 3.1 \pm 0.1$  per cent, consistent with the  $2.8 \pm 0.2$  per cent value measured in young stellar objects by Kounkel et al. (2019).

Unlike Kounkel et al. (2019), which focused primarily on the young stellar objects, most of the sources deconvolved as SB2s in this work are main sequence stars, and their CCFs are not affected by variability due to star spots. Therefore, it is possible to reliably include sources with quality flag 3 in addition to 4 in the list of likely SB2s (see Table 5 and Section 4.1 in Kounkel et al. (2019) for an explanation of these flags). Thus, we caution against blindly comparing these fractions. El-Badry et al. (2018b) used a more sophisticated method based on *The Payne* (Ting et al. 2019), to identify SB2s from RV shifts among dwarf stars in APOGEE DR12. Their measured SB2 fraction from this method is  $663/20,142 = 3.3 \pm 0.1$  per cent, which is consistent with our results. These authors also found SB2s by making multi-component spectral fits, and found a higher SB2 fraction of  $2645/20,142 = 13.1 \pm 0.2$  per cent. However, many of the systems identified by this method had small or negligible RV shifts and therefore this higher SB2 fraction is hard to compare with what we measure in our RV-selected sample.

We cross-matched our final sample of APOGEE targets with the catalogue from Sanders & Das (2018), who calculated Bayesian posteriors on distance  $d$ , mass  $M$ , and age  $\tau$ , by fitting PARSEC isochrones to a combination of Gaia DR2 parallaxes, broadband photometry, and the spectral parameters derived by ASPCAP. Sanders & Das (2018) give non-NAN values of  $d$ ,  $M$  and  $\tau$  for the vast majority (41,014, or 99 per cent) of the stars in our sample. We use these distance estimates to plot absolute 2MASS magnitudes  $J_{\text{abs}}$  vs.

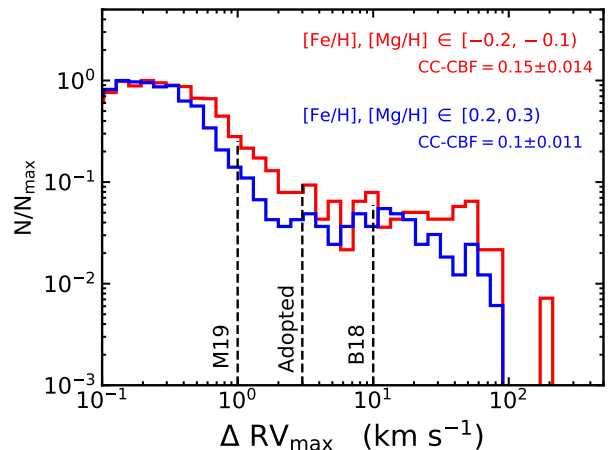
uncalibrated APOGEE  $T_{\text{eff}}$  in Fig. 1. The left panel shows the bulk of our sample in grey with the SB2s over-plotted in dark blue. Isochrones from the MESA Isochrone and Stellar Tracks Collaboration (MIST; Paxton et al. 2011, 2013, 2015; Dotter 2016; Choi et al. 2016) are shown for  $\tau = 8$  Gyr and a range of representative metallicities. According to Sanders & Das (2018), the age distribution in our sample peaks around 8 Gyr, which is in good agreement with the main sequence turn-off point shown in Fig. 1. The majority of the SB2s lie above the single star isochrone tracks, as expected for systems with a measurable flux contribution from both components. The right panel shows the same HR diagram coloured by the maximum shift in the RVs,  $\Delta RV_{\text{max}}$  (see Badenes & Maoz 2012; Maoz et al. 2012; Badenes et al. 2018; Moe et al. 2019), with stars that have  $\Delta RV_{\text{max}} \geq 1$  km s $^{-1}$  plotted on top for clarity. Here too we find a significant excess of objects with large RV variability to have locations above the single-star isochrones.

### 3 RESULTS

#### 3.1 Stellar multiplicity, $\Delta RV_{\text{max}}$ distributions, and completeness corrections

Following Badenes & Maoz (2012), Maoz et al. (2012), and Badenes et al. (2018), we use  $\Delta RV_{\text{max}}$  as a figure of merit to evaluate the sparsely sampled RV curves from APOGEE. Most (42.9 per cent) of the stars in our sample have 3 visits, with 36.4 per cent having 2 and the rest having 4 or more. While this is not enough to define a full orbital solution for most stars (see Price-Whelan et al. 2018, 2020, for discussions), values of  $\Delta RV_{\text{max}}$  above a certain threshold can securely identify large numbers of short-period binaries. In Fig. 2, we show the distribution of  $\Delta RV_{\text{max}}$  in two groups of  $N \sim 2000$  stars with constant  $[\text{Fe}/\text{H}]$  and  $[\text{Mg}/\text{H}]$ . This example illustrates the two main features of  $\Delta RV_{\text{max}}$  distributions derived from high quality data: a core of low  $\Delta RV_{\text{max}}$  values dominated by measurement errors and an extended tail of high  $\Delta RV_{\text{max}}$  values dominated by stars with companions in short-period orbits, clearly defined and cleanly separated from the core. We refer the reader to the discussions in Badenes et al. (2018) for the role of measurement errors, metallicity, and RV jitter in the APOGEE  $\Delta RV_{\text{max}}$  distributions. Here we focus on two closely related issues: the completeness corrections and the threshold value of  $\Delta RV_{\text{max}}$  to single out multiple systems.

We estimate completeness corrections on these  $\Delta RV_{\text{max}}$  distributions with a Monte Carlo sampler similar to that used by Moe et al. (2019). Our sampler simulates a population of  $N$  systems, with the fraction of systems in binaries determined by a free parameter called the multiplicity fraction  $f_m$ . Each system is assigned a visit history (number of visits and time lags between visits) from a random star in our APOGEE DR14 dwarf/subgiant sample. For each simulated binary, we draw the main orbital parameters (period and eccentricity) from the observational distributions measured for field solar-type binaries (period, Raghavan et al. 2010; eccentricity, Moe & Di Stefano 2017), select a random orbital inclination and initial phase, and generate RVs by sampling the projected orbit with the visit history, adding RV errors from a user-specified distribution. For each simulated single star, we set all RVs to zero and add errors from

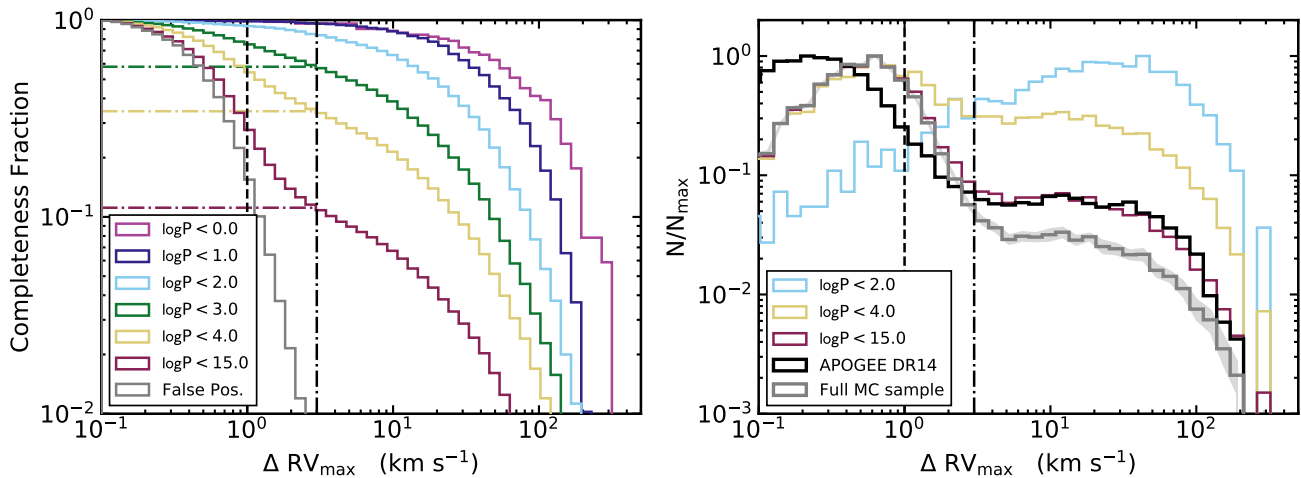


**Figure 2.** Normalised distributions of  $\Delta RV_{\text{max}}$  for two boxes with  $N \sim 2000$  in  $[\text{Fe}/\text{H}]$ - $[\text{Mg}/\text{H}]$  space from Fig. 4. The  $\Delta RV_{\text{max}}$  thresholds from Moe et al. (2019), Badenes et al. (2018), and the present work are shown as dashed lines.

the same distribution. The code is described in more detail in Badenes et al. (2018) – here we list the specific choices made for the present work. We simulate  $N = 50,000$  stars with  $f_m = 0.5$ . Each star is assigned  $\log(g/\text{cm s}^{-2}) = 4.25$ , the median value for our sample, which corresponds to a critical Roche Lobe Overflow period of  $P_{\text{crit}} = 0.49$  days in a  $1 M_{\odot}$  binary with  $q = 1$ . The primary mass  $M$  is randomly drawn from the distribution of Sanders & Das (2018) mass estimates for our sample (shown in the second diagonal panel of Fig. 4). For the mass ratio  $q$ , we assume a flat distribution with a twin excess fraction of 25 per cent for systems with  $0.95 \leq q \leq 1.0$  (Moe & Di Stefano 2017). The RV errors are drawn from a Student’s  $t$  distribution (scipy.stats.t) with degrees of freedom 3.5, location 0 and scale 0.25. Appendix A discusses these choices and their effects on the completeness corrections in more detail.

In Table 1 and the left panel of Fig. 3, we show the cumulative fraction of systems with  $\Delta RV_{\text{max}}$  above a given value in several period ranges in our Monte Carlo simulation. Assuming the underlying period and eccentricity distributions are not too different from the assumed ones, the completeness correction that needs to be applied to recover the total number of binaries with periods below a certain value is the inverse of these cumulative fractions. Our results are consistent with those of Moe et al. (2019) (included in Table 1), who applied a similar approach to a sample of APOGEE DR13 dwarfs. The grey and red histograms in Fig. 3 show the cumulative fractions for all non-binary and binary systems, respectively. The false positive rate for binaries in a given period range at a given value of  $\Delta RV_{\text{max}}$  is the ratio between the relevant cumulative fraction and the grey histogram at that value of  $\Delta RV_{\text{max}}$ .

These curves inform our choice of  $\Delta RV_{\text{max}}$  threshold value. A conservative value like the  $10$  km s $^{-1}$  chosen by Badenes et al. (2018) is virtually free of false positives, but results in low detection efficiencies and correspondingly large completeness corrections, which can lead to issues when dealing with small samples of systems with a specific set of



**Figure 3.** Left panel: Completeness fraction versus  $\Delta RV_{\max}$  for Monte Carlo generated samples at various period limits, using the APOGEE DR14 time lags from our sample. The horizontal dot-dashed lines are the completeness fractions for the relevant  $\log(P/\text{days})$  samples given a threshold  $\Delta RV_{\max} \geq 3 \text{ km s}^{-1}$ . The vertical dashed line is at  $\Delta RV_{\max} = 1 \text{ km s}^{-1}$  for comparison. The grey histogram shows the completeness fraction for false positives—systems that are not binaries but show some RV variation due to simulated RV error. Right panel: normalised  $\Delta RV_{\max}$  distributions for MC data. The fainter lines are for several  $\log(P/\text{days})$  samples, colour-coded as in the left panel. The grey histogram is for the full MC sample (binaries and non-binaries included), with shading indicating  $1\sigma$  intervals from bootstrapping the sample ( $N_{\text{boots}} = 25$ ). The black histogram is for our APOGEE DR14 sample.

**Table 1.** Completeness fractions for selected  $\log(P/\text{days})$  and  $\Delta RV_{\max}$  thresholds.

$\log(P/\text{days})$ threshold	$\Delta RV_{\max} \geq 1 \text{ km s}^{-1}$	$\Delta RV_{\max} \geq 3 \text{ km s}^{-1}$	$\Delta RV_{\max} \geq 10 \text{ km s}^{-1}$
$\log P \leq 0.0$	1.00	0.96	0.88
$\log P \leq 2.0$	0.93	0.84	0.66
$\log P \leq 4.0$ , this work	0.55	0.34	0.21
$\log P \leq 4.0$ , from Moe et al. (2019)	0.57	0.40	0.24
$\log P \leq 15.0$	0.29	0.11	0.07
False Positives	0.17	0.005	0.00

stellar parameters. For the dwarf and subgiant stars that we examine here, which have low RV jitter (Hekker et al. 2008) and relatively narrow  $\Delta RV_{\max}$  distribution cores (Badenes et al. 2018), we propose a more reasonable value of  $3 \text{ km s}^{-1}$ . Using the uncertainties reported by the APOGEE data reduction pipeline, the median RV uncertainty for our sample is  $\sigma_{RV} \sim 0.04 \text{ km s}^{-1}$ , though these uncertainties are almost certainly underestimated (see discussions in Holtzman et al. 2018; Badenes et al. 2018, and sources within). We can instead consider a more reasonable value of  $\sigma_{RV} \sim 0.2 \text{ km s}^{-1}$ , obtained from roughly fitting the observed  $\Delta RV_{\max}$  distribution core to those simulated by our MC with Gaussian error distributions with mean of 0 and varying spreads (similar to APOGEE DR13, see Fig. 9 of Badenes et al. 2018). Regardless, our threshold remains far larger than what can be explained with typical RV uncertainties alone. This threshold yields a detection efficiency of  $\approx 34$  per cent for systems with  $\log(P/\text{days}) \leq 4.0$  and  $\approx 84$  per cent for  $\log(P/\text{days}) \leq 2.0$ , with an overall false positive rate of  $\approx 0.1$  per cent. Compared to Moe et al. (2019), who chose a threshold value of  $\Delta RV_{\max} \geq 1 \text{ km s}^{-1}$ , we expect a false positive rate about 30x

lower, with only a modest loss of  $\approx 20$  per cent in detection efficiency.

In the context of our APOGEE sample, completeness corrections for systems with  $\log(P/\text{days}) > 4$  ( $a > 10 \text{ AU}$ ) are unwarranted for several reasons. These long-period binaries will rarely produce detectable RV variability in APOGEE, and are often difficult to characterise using sparsely sampled RV curves. Moreover, the anti-correlation between stellar multiplicity and  $[\text{Fe}/\text{H}]$  weakens beyond  $a > 50 \text{ AU}$  and disappears beyond  $a > 200 \text{ AU}$  (Moe et al. 2019; El-Badry & Rix 2018), and this might apply to other stellar parameters. Therefore, in the remainder of this work we will quote completeness-corrected binary fractions for systems with  $\log(P/\text{days}) \leq 4$ , which we identify as ‘close binaries’. For reference, a  $1 M_{\odot}$  star of solar composition at the tip of the red giant branch has a critical period for Roche Lobe overflow of  $\log(P/\text{days}) \sim 2.8$ .

In the right panel of Fig. 3 we compare the simulated  $\Delta RV_{\max}$  distribution from our MC run to the observed distribution in the APOGEE sample. We estimate  $1\sigma$  intervals on the simulated distribution (shown as the grey shading) by bootstrapping the sample with  $N_{\text{boot}} = 25$ ,

$N_{\text{sys}} = 40,000$ . We also show the  $\Delta RV_{\text{max}}$  distributions in three different subsets of simulated systems: those with  $\log(P/\text{days}) < 15$  (all binaries),  $\log(P/\text{days}) < 4$  (all close binaries), and  $\log(P/\text{days}) < 2$ . We do not attempt to provide an accurate match to the observed  $\Delta RV_{\text{max}}$  distribution, as this would require a complete characterisation of the correlations between stellar properties, multiplicity, and RV errors, but we note that the shape and extent of the tail in our simulation is very similar to what we see in the APOGEE sample. We also note that our choice of RV error distribution is conservative, as shown by the comparison between the simulated and observed core shapes.

Binaries in general, and twins in particular, can be detected further away than single stars in magnitude-limited samples due to Malmquist bias (see Fig. 1). Conversely, it is more difficult to detect RV variability of twin SB2s if their absorption features are significantly blended (but see El-Badry et al. 2018b, for an alternative approach). In their analysis, Moe et al. (2019) estimated that these two effects bias the close binary fraction measured by APOGEE by  $\approx 30$  per cent in opposite directions and therefore approximately cancel each other. However, they relied solely on the APOGEE pipeline RV measurements, while we applied a CCF method to identify SB2s and more accurately measure their RVs. Our Malmquist bias in favour of detecting twin binaries should therefore be slightly greater than our inefficiency in the detection of RV variability in SB2s. We compensate for this by reducing our completeness-corrected close binary fractions by 10 per cent to make the reported values more representative of volume-limited samples. This results in an estimated detection efficiency of 0.38 for  $\Delta RV_{\text{max}} \geq 3 \text{ km s}^{-1}$  and  $\log(P/\text{days}) \leq 4$ , which we adopt for the remainder of this work. Using this completeness correction, the close binary fractions we recover from the  $\Delta RV_{\text{max}}$  distributions shown in Fig. 2 are  $0.15 \pm 0.014$  and  $0.1 \pm 0.011$ .

### 3.2 The Impact of Stellar Parameters on the Close Binary Fraction

We are now in a position to examine the impact of stellar parameters on the completeness-corrected close binary fractions measured in our sample of APOGEE dwarfs and subgiants. To do this, we choose a few representative parameters among those measured by APOGEE and Sanders & Das (2018):  $T_{\text{eff}}$ ,  $M$ ,  $[\text{Fe}/\text{H}]$ ,  $[\text{Mg}/\text{H}]$ ,  $[\text{Si}/\text{H}]$ ,  $\tau$ , the vertical action  $J_z$ , and the galactocentric radial velocity  $v_R$ . The vertical action is defined as

$$J_z = \frac{1}{2\pi} \oint dz v_z \quad (1)$$

where  $z$  and  $v_z$  are the position and galactocentric vertical velocity for the star along its orbit. As an indicator of a star's vertical displacement,  $J_z$  is unaffected by orbital phase as compared to  $z$  or  $v_z$ , and it is a tracer of the birth location of stars in the Milky Way disk that is more robust to radial migration than galactocentric radius (Vera-Ciro et al. 2014).

Several of these parameters are precisely determined by APOGEE ( $T_{\text{eff}}$ , chemistry), whereas others represent fundamental stellar properties ( $M$ ,  $\tau$ ) or are related to galactic dynamics that may prove interesting ( $J_z$ ,  $v_R$ ). Of course many of these parameters, like  $\tau$  and  $[\text{Fe}/\text{H}]$ , have substan-

tial internal correlations that cannot be properly examined without a multivariate analysis. Moreover, we are restricted to the parameter ranges covered by APOGEE, which are very broad for some parameters like  $[\text{Fe}/\text{H}]$ , but quite narrow for others that are of high interest for stellar multiplicity, like  $M$ . Finally, not all these parameters are equally well constrained by the observations. Stellar ages, for example, are notoriously hard to estimate without asteroseismic data (e.g., see Ness et al. 2016; Pinsonneault et al. 2018). We also note that both ASPCAP and Sanders & Das (2018) assume single star models, which can introduce biases in some parameters (see El-Badry et al. 2018a, for a discussion).

With all these caveats in mind, we present our view of the impact of stellar parameters on the close binary fraction in Fig. 4. This triangle plot shows the completeness-corrected close binary fraction as a two-dimensional histogram mapped on each pairwise combination of parameters. The one-dimensional terminal plots show the full distribution of each parameter in the APOGEE sample (black histograms) and the completeness-corrected close binary fraction as a function of that parameter alone (blue histograms with shaded error bars). We required a minimum of ten objects per bin in order to extend our measurements through the sample's full range of parameter space. Uncertainties are not shown in the 2D histograms, but they scale as  $\sigma_f/c$ , where  $c = 0.38$  is the completeness-correction discussed in Section 3.1, and  $\sigma_f$  is the uncertainty from the binomial process on each measurement,

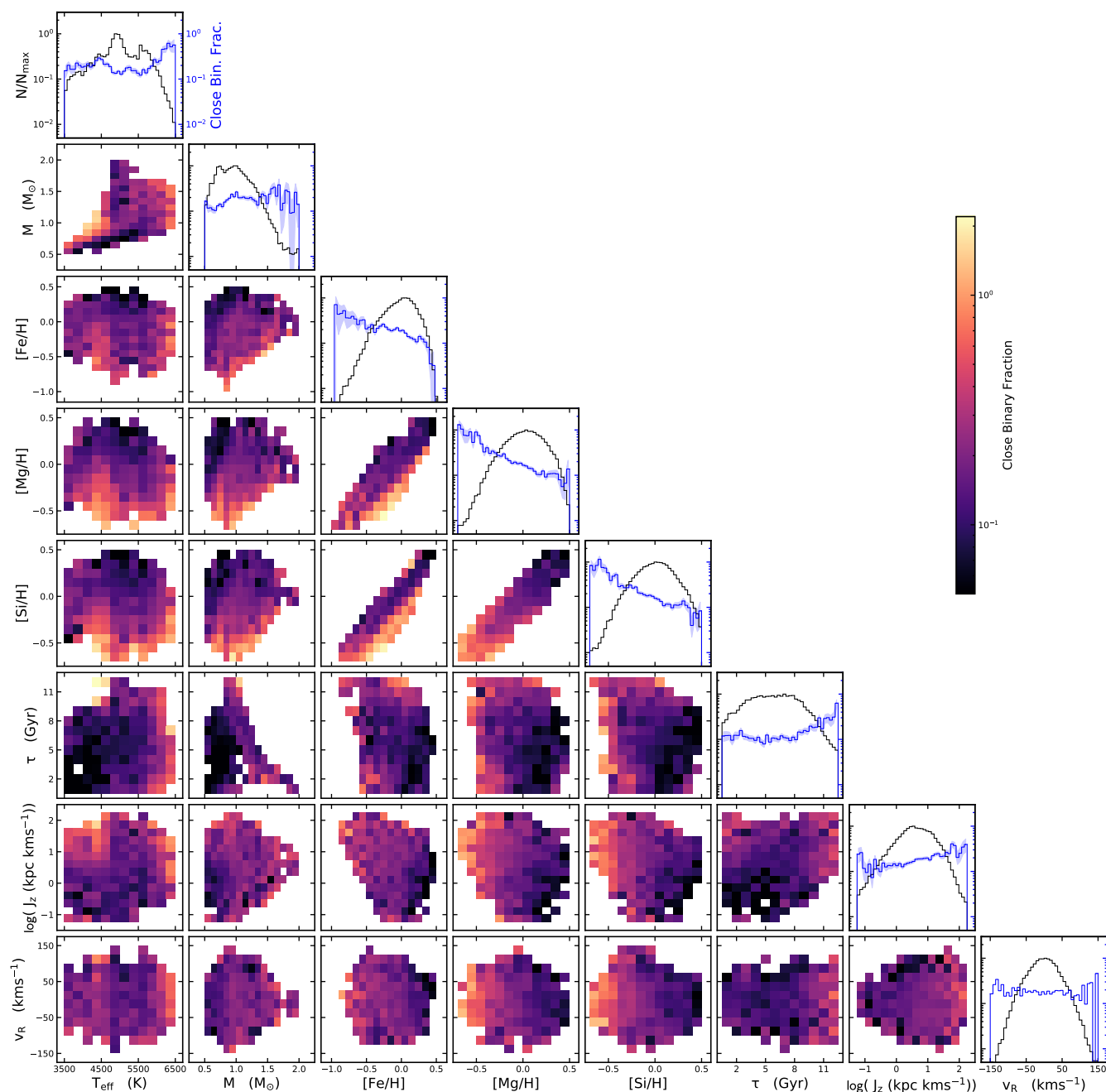
$$\sigma_f = \sqrt{\frac{f(1-f)}{N}} \quad (2)$$

where  $f$  is the fraction of systems with  $\Delta RV_{\text{max}} \geq 3 \text{ km s}^{-1}$ , and  $N$  is the total number of systems in that bin. Measurements made with small- $N$  samples will be noisy due to the  $\sqrt{1/N}$  factor, but the RV variable fraction  $f$  also introduces a  $\sqrt{f(1-f)}$  factor. For a bin with  $N = 10$ , we can consider two cases: (1)  $f = 0.2$  and (2)  $f = 0.8$ . In both instances, the binomial process uncertainty is  $\sigma_f/c = 0.33$ . The completeness-corrected close binary fractions are (1)  $f_m = 0.52 \pm 0.33$  and (2)  $f_m = 2.09 \pm 0.33$ , showing that it is possible to measure variations in the close binary fraction even in bins with  $N$  as small as 10.

Note also that our completeness-correction can result in close binary fractions that are in excess of 100 per cent, and we indeed see bins with values of  $f_m \sim 2.0$  in Fig. 4. We assumed the same period distribution for the entire simulated sample, and this assumption is most likely not valid across our diverse APOGEE sample. From the ASAS-SN Catalogue of Variable Stars, Jayasinghe et al. (2020) found that metal-poor eclipsing binaries were skewed towards shorter periods than metal-rich systems at fixed temperature. A shift towards shorter periods for metal-poor stars results in an over-correction from the completeness estimate, leading to our excessively large close binary fractions. Future studies of the period distribution as a function of chemistry and metallicity will be useful for addressing this issue.

The salient features of Fig. 4 can be summarised as follows:

(i) The parameters related to chemical composition ( $[\text{Fe}/\text{H}]$ ,  $[\text{Mg}/\text{H}]$ , and  $[\text{Si}/\text{H}]$ ) emerge as the dominant drivers of stellar multiplicity in our sample. The completeness-



**Figure 4.** Two-dimensional histograms showing the completeness-corrected close binary fraction as a function of many of the parameters considered in this work. Along the diagonal, the black lines are the normalised histograms of each parameter, and shown in blue is the completeness-corrected close binary fraction as a function of that parameter alone. The blue shaded region shows the uncertainties on the completeness-corrected close binary fraction (equation 2).

corrected close binary fractions as a function of these parameters (blue 1D histograms in the diagonal panels) show clear monotonic downward trends, with dynamic ranges in excess of an order of magnitude, that are distinctly larger than for any other parameters. The gradients due to this downward trend are the most striking feature in all the 2D histograms that include chemical composition parameters. While the trends are uniform and monotonic in the 1D histograms, the 2D histograms reveal a great deal of complexity in the

relationship between stellar multiplicity and chemical composition, which we examine in further detail in Section 3.3.

(ii) Even though stellar mass (and by proxy,  $T_{\text{eff}}$ ) is known to have a strong effect on the close binary fraction of field dwarfs (Lada 2006; Duchêne & Kraus 2013; Moe & Di Stefano 2017), this is not clear in the APOGEE sample. The close binary fraction ( $\log(P/\text{days}) \lesssim 4$ ) for Solar-mass stars scales as  $M^{0.5}$  (Moe & Di Stefano 2017), so we expect the close binary fraction to increase by a factor of 2 across the



**Table 2.** Fit parameters and the number of  $\sigma$  (estimated via bootstrapping) for the parameter to be consistent with 0.

	$\log f_m = b + aX$		$\log f_m = c + bX + aX^2$			$\chi_{\text{lin}}^2/\chi_{\text{quad}}^2$
	$b$	$a$	$c$	$b$	$a$	
$T_{\text{eff}}$ (K)	-1.595	1.9e-7 2.7 $\sigma$	3.318	-0.002 5.4 $\sigma$	1.9e-7 5.8 $\sigma$	2.7
$M$ ( $M_{\odot}$ )	-0.815	0.11 2.1 $\sigma$	-1.329	0.993 3.1 $\sigma$	-0.34 2.5 $\sigma$	1.5
[Fe/H] (dex)	-0.787	-0.595 8.6 $\sigma$	-0.782	-0.436 4.7 $\sigma$	0.196 1.2 $\sigma$	0.98
[Mg/H] (dex)	-0.806	-1.32 13.0 $\sigma$	-0.752	-0.851 6.6 $\sigma$	0.627 2.7 $\sigma$	3.5
[Si/H] (dex)	-0.77	-1.15 9.1 $\sigma$	-0.755	-1.011 3.6 $\sigma$	0.187 0.5 $\sigma$	1.3
$\tau$ (Gyr)	-1.355	0.077 2.8 $\sigma$	-0.696	-0.127 3.5 $\sigma$	0.013 4.4 $\sigma$	4.7
$\log(J_z/\text{kpc km s}^{-1})$	-0.76	0.142 5.0 $\sigma$	-0.783	0.045 1.1 $\sigma$	0.062 2.7 $\sigma$	1.4
$v_R$ ( $\text{km s}^{-1}$ )	-0.662	1.0e-5 0.1 $\sigma$	-0.757	1.0e-4 0.4 $\sigma$	1.0e-5 2.4 $\sigma$	1.1

sample’s mass range. From the 1D histogram, we observe the close binary fraction increasing by a factor of  $\sim 1.5$ . However, the close binary fraction measurements in the high- $M$  bins are noisy, and the mass estimates themselves are poorly constrained compared to APOGEE  $T_{\text{eff}}$ , so our measurement alone cannot be considered to be at odds with previous work. We do detect a noticeably higher close binary fraction for the hottest ( $T_{\text{eff}} \gtrsim 6000$  K) stars, which Price-Whelan et al. (2020) also found in a sample of binaries in APOGEE DR16. While this spike may be due to larger primary masses, the correlation at lower temperatures seem weaker. This might be due to the overlap between dwarfs and subgiants below 6000 K (apparent in Fig.1). In the  $T_{\text{eff}}-M$  2D histogram, there appears to be a region of increased binaries around 4000 K and 1  $M_{\odot}$ , though this is more likely to be a result of erroneous mass estimates, given the temperature and mass values.

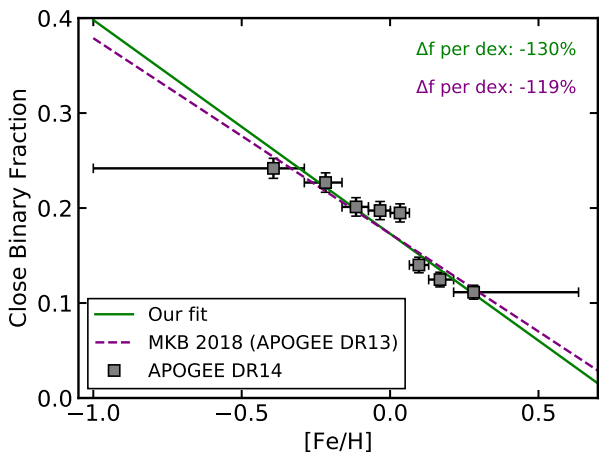
(iii) Stellar age shows a modest upward trend, though this is hard to interpret. Stellar ages are poorly constrained in general, and age estimates for SB2s are particularly prone to errors: SB2 systems may be mis-classified as overly young (100s Myr) or overly old ( $> 10$  Gyr), because stars that are offset from the MS, like the high- $\Delta RV_{\text{max}}$  objects in Fig. 1, may be classified along stellar pre-MS or post-MS tracks. This could account for the apparent increase in multiplicity fraction for  $\tau > 8$  Gyr. There is also a well established (though complex) correlation between  $[\alpha/\text{Fe}]$ ,  $[\text{Fe}/\text{H}]$ , and age (Mackereth et al. 2017, 2019, and sources within), which is often used in studies of galactic dynamical evolution. A more complete treatment of these correlations is required before we can comment on any trends between age and the close binary fraction.

(iv) The 1D histogram for  $\log J_z$  shows a significant correlation with the completeness-corrected close binary fraction, but this could simply be due to the fact that the outer disk

is more metal-poor (e.g., in APOGEE Hayden et al. 2015; Weinberg et al. 2019, and sources within).

(v) The galactocentric radial velocity shows the flattest distribution of the parameters studied here. In the 1D histogram, the bins at either edge in parameter space appear to have an increased binary fraction, but they are consistent with a flat distribution given their large uncertainties.

To quantify the impact of each parameter on the multiplicity fraction, we fit linear and quadratic functions to each of the blue histograms along the diagonals of Fig. 4. The best fit parameters and the ratio between the  $\chi^2$  are given in Table 2. None of the distributions are necessarily expected to follow a linear or quadratic function, but these are simple, easily-fit functions that provide an estimate of the slopes of the distributions. We then bootstrapped ( $N_{\text{boot}} = 500$ ) the fits to estimate uncertainties on the fit parameters. We can then calculate the number of  $\sigma$  required for the first and second derivatives to be consistent with zero. These values are listed in the second row for each parameter in Table 2, with significant values ( $n > 2$ ) in purple and highly significant values ( $n > 5$ ) in blue. From these values, we conclude that the chemical composition parameters show the most significant correlations with close binary fraction in our sample, though there are also clear trends with stellar age, mass,  $T_{\text{eff}}$ , and vertical action. We recover the strong anti-correlation between  $[\text{Fe}/\text{H}]$  and the completeness-corrected close binary fraction previously reported by various authors, and identify for the first time a similar effect in both sign and strength for  $\alpha$ -process elements Mg and Si. Characterising these correlations is the subject of the remainder of this paper.



**Figure 5.** Completeness-corrected close binary fraction for each  $[\text{Fe}/\text{H}]$  bin. The horizontal error bars show the  $[\text{Fe}/\text{H}]$  range of each bin, and the vertical error bars show the completeness-adjusted uncertainty,  $\sigma_f/c$ . The results of Moe et al. (2019) are over-plotted alongside a linear fit to our data in order to find the difference in the close binary fraction per dex of  $[\text{Fe}/\text{H}]$ .

### 3.3 Chemical Composition and the Close Binary Fraction

The completeness-corrected close ( $\log(P/\text{days}) \leq 4.0$ ) binary fraction as a function of  $[\text{Fe}/\text{H}]$  alone is shown in Fig. 5. We divided our sample into eight bins in  $[\text{Fe}/\text{H}]$ , chosen to contain approximately 5200 stars each. A linear fit to these data shows that the close binary fraction decreases by a factor of  $\sim 2.4$  from  $[\text{Fe}/\text{H}] = -0.5$  dex to  $[\text{Fe}/\text{H}] = 0.25$  dex. As we have seen, however, the relationship between chemical composition and stellar multiplicity is complex, and it cannot be characterised by metallicity alone. Here we consider in detail four parameters related to the abundance of  $\alpha$ -process elements:  $[\text{Mg}/\text{H}]$  and  $[\text{Si}/\text{H}]$  (already discussed in Section 3.2), plus  $[\alpha/\text{H}]$  and  $[\text{O}/\text{H}]$ . The measurements of  $[\text{C}/\text{H}]$  and  $[\text{N}/\text{H}]$  for APOGEE DR14 dwarfs are not reliable (Holtzman et al. 2018), so we did not include them in our analysis.

We begin by revisiting the two-dimensional histograms of completeness-corrected close binary fraction. Each panel in Fig. 6 shows an  $\alpha$ -process abundance measurement as a function of  $[\text{Fe}/\text{H}]$ , similar to the 2D histograms of Fig. 4, but with a lower minimum count of five stars per bin to maximise parameter space coverage. The close binary fraction again exceeds 100 per cent in multiple bins, though this still may be due to the degeneracies present in our RV variability fraction method discussed in Sec. 3.2. The anti-correlation between close binary fraction and  $[\text{Fe}/\text{H}]$  is apparent as the trend along the diagonal, and it is present for all six  $\alpha$  abundances. The additional anti-correlation with  $\alpha$  abundance is clear when looking along lines of constant  $[\text{Fe}/\text{H}]$ , manifesting as two distinct sequences:  $\alpha$ -poor with large close binary fractions, and  $\alpha$ -rich with smaller close binary fractions. The weakest effect is seen in  $[\text{O}/\text{H}]$ , but the anti-correlation is obvious for  $[\alpha/\text{H}]$ ,  $[\text{Mg}/\text{H}]$ ,  $[\text{Mg}/\text{Fe}]$ , and Si. However, especially around solar metallicity,  $[\alpha/\text{H}]$ ,  $[\text{O}/\text{H}]$ , and  $[\text{Si}/\text{H}]$

show increased close binary fractions at low and high values when looking along lines of constant  $[\text{Fe}/\text{H}]$ .

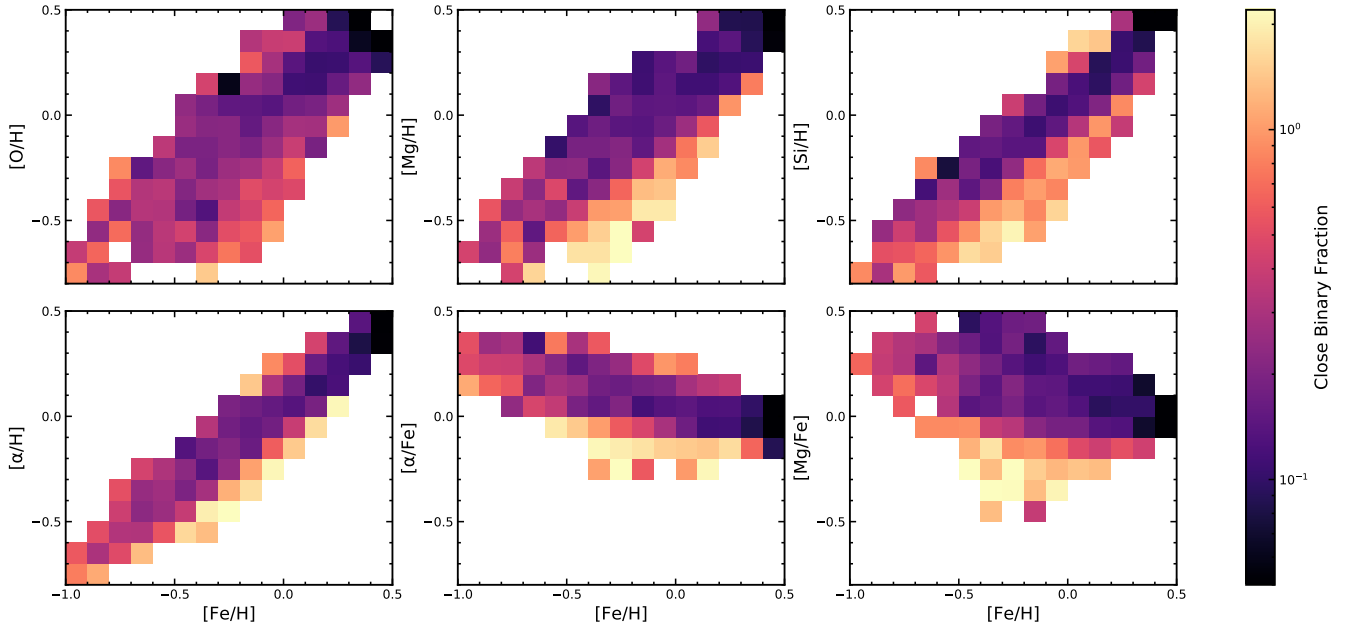
To study these effects in a regime that is not prone to numerical noise due to small numbers of stars, we use the same bins as those shown in Fig. 5 ( $N \sim 5200$  each). The grey squares in the first column of Figs. 7-8 are plotted at the median  $[\text{Fe}/\text{H}]$  and  $\alpha$  abundance for each bin, with each row showing one of the four  $\alpha$  abundance measurements from earlier, plus  $[\alpha/\text{Fe}]$ . Within each  $[\text{Fe}/\text{H}]$  bin we define “low-X” and “high-X” subsamples, with X standing for each of the five parameters we study, shown in red and blue. The dividing line between low- and high-X samples is drawn using a quadratic fit to the median with a finer grid of thirty-five bins in  $[\text{Fe}/\text{H}]$ . The second column of Figs. 7-8 shows the fraction of systems with  $\Delta\text{RV}_{\text{max}} \geq 3 \text{ km s}^{-1}$  for each of the low- and high-X subsamples as a function of  $[\text{Fe}/\text{H}]$ . Horizontal error bars indicate the extent of the  $[\text{Fe}/\text{H}]$  bins, and vertical error bars represent the binomial process uncertainty,  $\sigma_f$  (equation 2). The anti-correlation between  $[\text{Fe}/\text{H}]$  and close binary fraction is again present, but there is a significant gap between the low- and high-X subsamples in all five parameters we study. The green points in the third column of Figs. 7-8 display the ratio of RV variability fractions between the low- and high-X abundance subsamples, with horizontal error bars again indicating the extent of the  $[\text{Fe}/\text{H}]$  bins and vertical error bars denoting the uncertainty obtained via error propagation. This ratio is greater than one across every  $[\text{Fe}/\text{H}]$  bin and for every abundance considered here. The ratios generally increase with  $[\text{Fe}/\text{H}]$  for O, but the opposite appears to be true for Mg and Si.

These results reinforce our finding that  $\alpha$  element abundances have a strong impact on the close binary fraction. To further investigate to what extent this effect is separate from the  $[\text{Fe}/\text{H}]$  effect, we calculated the difference in subsample means for four of the abundance measurements, defined as

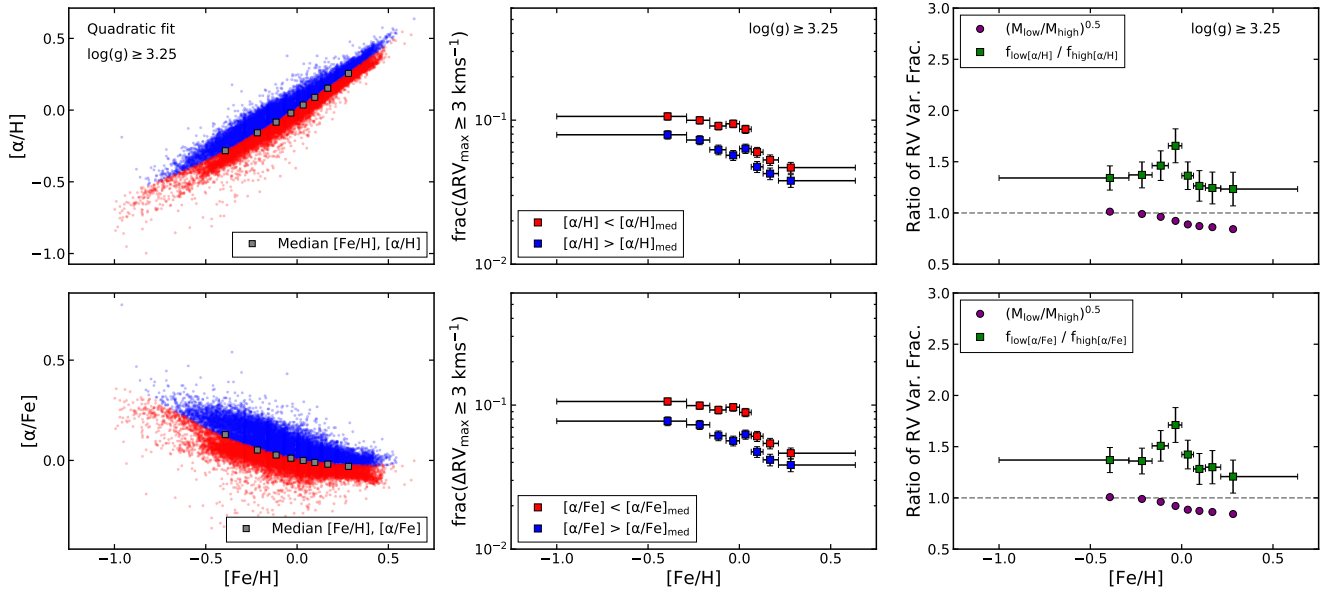
$$\Delta[\text{Y}/\text{H}]_{\text{mean}} = \text{mean}([\text{Y}/\text{H}]_{\text{high}}) - \text{mean}([\text{Y}/\text{H}]_{\text{low}}) \quad (3)$$

where Y can be Fe or one of the  $\alpha$ -process abundances. We plot these differences in Fig. 9, which shows that the differences in mean  $[\text{Fe}/\text{H}]$  are essentially 0 for all of the bins except the first and eighth, while the differences in mean  $[\text{X}/\text{H}]$  for  $\alpha$ , O, Mg, and Si remain substantial, although they do decrease as  $[\text{Fe}/\text{H}]$  increases. In other words, while some systematic differences in  $[\text{Fe}/\text{H}]$  exist between the high-X and low-X samples that we have defined for the  $\alpha$  element abundances, they are too small to account for the effect that we see. Of course, our high-X and low-X samples are not exactly comparable in every aspect, but the effect we observe is too large to be due to other (i.e., non-chemistry related) factors. To illustrate this, we also show in the third column of Figs. 7-8 the magnitude of the effect due to systematic differences in the stellar mass measured by Sanders & Das (2018) between the high-X and low-X samples as a function of  $[\text{Fe}/\text{H}]$ . These systematic differences, while present, are again too small to explain the disparity in RV variability fraction between the low- and high- $\alpha$  subsamples.

Another way to disentangle the  $[\text{Fe}/\text{H}]$  and  $\alpha$  effects is to examine trends with  $\alpha$  abundances in a narrow range of  $[\text{Fe}/\text{H}]$  (Fig. 10). For each  $\alpha$  abundance, we select a subset of the full sample that spans  $-0.075 \leq [\text{Fe}/\text{H}] \leq 0.075$  and  $-0.2 \leq [\text{X}/\text{H}] \leq 0.2$ , shown in dark red in the first column of Fig. 10. We then divide this subsample into eight bins



**Figure 6.** Two-dimensional histogram showing the completeness-corrected close binary fraction as a function of  $[\text{Fe}/\text{H}]$  and various  $\alpha$  abundances.



**Figure 7.** First row: [left] distribution of  $[\alpha/\text{H}]$  and  $[\text{Fe}/\text{H}]$ , with the grey points placed at the median  $[\text{Fe}/\text{H}]/[\alpha/\text{H}]$  of each bin and the red points showing the “low- $[\alpha/\text{H}]$ ” subsample and blue the “high- $[\alpha/\text{H}]$ ” subsample; [center] the fraction of systems with  $\Delta\text{RV}_{\text{max}} \geq 3$  km s $^{-1}$  for the low- and high- $[\alpha/\text{H}]$  subsamples, with the horizontal error bars showing the range of  $[\text{Fe}/\text{H}]$  in each bin and the vertical error bars showing the uncertainty, equation (2); and [right] the ratio of the low- $\alpha$  to high- $\alpha$  bins’ RV variability fraction alongside the ratio of median masses between the low- $\alpha$  and high- $\alpha$  bins. The second row is the same but for  $[\alpha/\text{Fe}]$ .

across the relevant  $\alpha$  abundance, with histograms for each bin shown in the second column and the number of objects in each bin listed in the coloured text. Cumulative  $\Delta\text{RV}_{\text{max}}$  histograms for four of these bins are shown in the third column. The fourth column displays the completeness-corrected close binary fraction as a function of each  $\alpha$ -process abundance

measurement, with the horizontal error bars indicating the edges of the bins and the vertical error bars representing the completeness-adjusted uncertainty,  $\sigma_f/c$ . This analysis reveals that the close binary fraction in this narrow  $[\text{Fe}/\text{H}]$  range is clearly anti-correlated with Mg. For the other three abundances, there is a general downward trend, but the de-

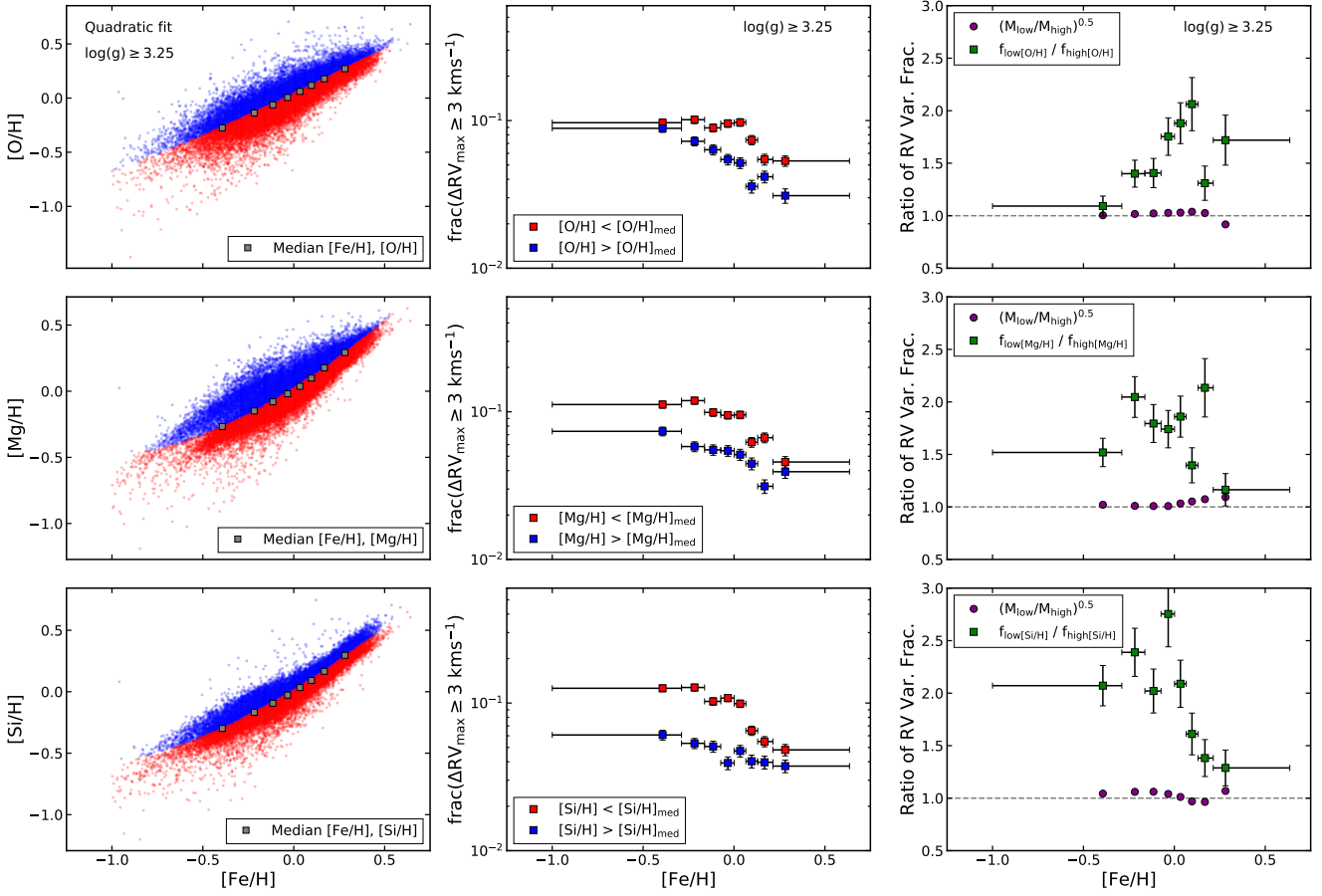


Figure 8. Same as Fig. 7, but for [O/H], [Mg/H], and [Si/H].

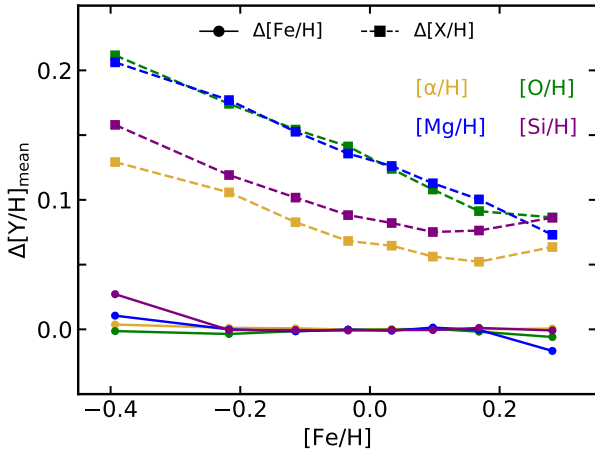


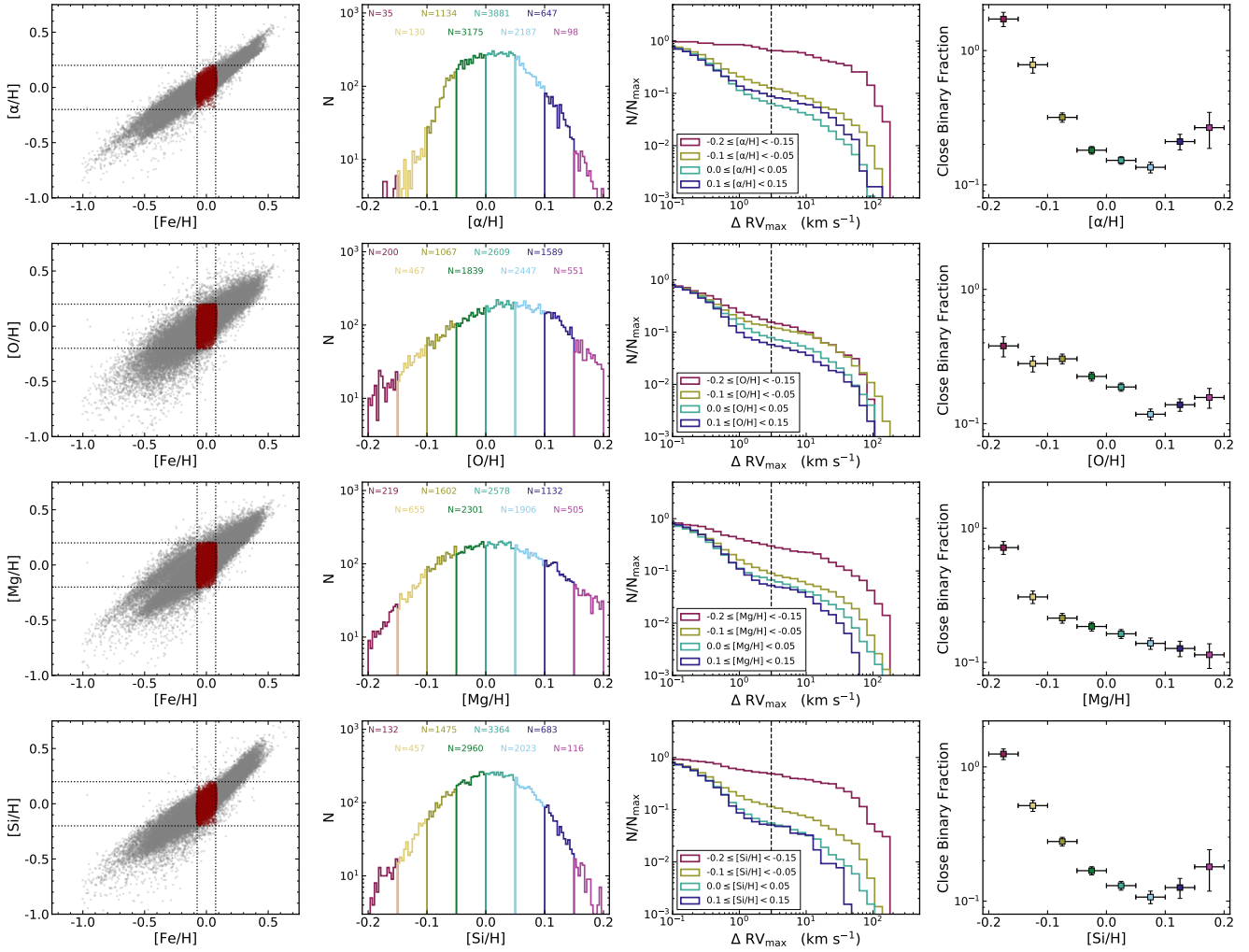
Figure 9. Differences in the mean [Y/H] for each sample, with  $\Delta[Y/H]_{\text{mean}} = \text{mean}([Y/H]_{\text{high}}) - \text{mean}([Y/H]_{\text{low}})$ .

tailed behaviour is more complex. For  $[\alpha/H]$ , O, and Si, it appears to reach a minimum around 0.075 dex, and then steadily increases once again. This turnaround is weakly present in O, but it is clear in  $[\alpha/H]$  and Si. Similar to the

third panel of Figs. 7-8, we compared the ratio of the median mass for each bin against the bin with the minimum close binary fraction. Across all abundances, the difference in median mass between bins is insignificant compared to the difference in observed RV variability.

We repeated this analysis for other narrow ranges of [Fe/H]. Each subsample spanned a width of  $\Delta[Fe/H] = 0.15$  dex ( $\pm 0.075$  from the central value) and  $\Delta[X/H] = 0.4$  dex ( $\pm 0.2$  from the central value). The central values we compared, in pairs of ( $[Fe/H]$ ,  $[X/H]$ ), were  $(-0.4, -0.3)$ ;  $(-0.2, -0.2)$ ; and  $(0.2, 0.2)$  - the figures for each set are included in Appendix B. The turnaround that we see in the solar [Fe/H] sample was present in some, but not all the [Fe/H] bins. These results confirm the trends seen in Fig. 6 with larger sample sizes, and show that the multiplicity statistics for stars with a specific chemistry can be quite extreme - see for instance the prominent tail in the  $\Delta RV_{\max}$  distributions corresponding to the lowest  $\alpha$  abundances in Fig. 10. In these extreme cases, it is possible that our assumed underlying period distribution is incorrect, which would make our derived values of the completeness corrected close binary fractions incorrect. However, our reported high fractions of RV variability are robust, and clearly require a high frequency of close binary companions, regardless of the underlying period distribution.

Finally, we note that when looking along lines of con-



**Figure 10.** Various distributions for a selection of data in a narrow range around solar  $[\text{Fe}/\text{H}]$ . First row: [far left]  $[\alpha/\text{H}]$  versus  $[\text{Fe}/\text{H}]$ , where the entire sample is shown in grey and the chosen subsample is shown in dark red (boundaries  $-0.075 \leq [\text{Fe}/\text{H}] \leq 0.075$  and  $-0.2 \leq [\alpha/\text{H}] \leq 0.2$ ); [center left] histograms for the selected data, split into eight equally spaced bins across  $[\alpha/\text{H}]$ ; [center right] selected cumulative  $\Delta \text{RV}_{\text{max}}$  histograms; and [far right] the close binary fraction as a function of  $[\alpha/\text{H}]$  for the selected data, colour-coded by its  $[\alpha/\text{H}]$  bin. The horizontal error bars show the  $[\alpha/\text{H}]$  range of each bin, and the vertical error bars show the completeness-adjusted uncertainty,  $\sigma_f/c$ . The remaining rows are the same but for  $[\text{O}/\text{H}]$ ,  $[\text{Mg}/\text{H}]$ , and  $[\text{Si}/\text{H}]$ .

stant  $\alpha$  abundance in Fig. 6, the binary fraction is often *positively* correlated with  $[\text{Fe}/\text{H}]$ . To verify that this is a real effect and not just a result of binning and small number statistics, we examined the cumulative chemistry distributions for both our entire sample and just the objects with  $\Delta \text{RV}_{\text{max}} \geq 3 \text{ km s}^{-1}$ . For every  $\alpha$  abundance measurement, the cumulative  $[\text{X}/\text{H}]$  distributions across each of the eight  $[\text{Fe}/\text{H}]$  bins used in Figs. 7-8 reveal that the RV variables are always shifted towards lower  $\alpha$  abundances compared to the total population. Again, this confirms the general anti-correlation between close binary fraction and  $\alpha$  abundances. We then plotted the cumulative  $[\text{Fe}/\text{H}]$  distributions for six bins of equal width across  $-0.8 \leq [\text{X}/\text{H}] < 0.4$ . For the first three bins ( $-0.8 \leq [\text{X}/\text{H}] < -0.2$ ), the RV variables are shifted towards *higher*  $[\text{Fe}/\text{H}]$  abundances than the total population. This is true for all  $\alpha$  abundances considered here. However, the cumulative metallicity distributions for

the two bins between  $-0.2 \leq [\text{X}/\text{H}] < 0.2$  generally show a weakening of this trend, and the bin for  $0.2 \leq [\text{X}/\text{H}] < 0.4$  shows a reversal—i.e., the RV variables are shifted towards *lower*  $[\text{Fe}/\text{H}]$  than the greater population. These results lend support for a inflection point in the close binary fraction as a function of  $\alpha$  abundances around 0.1 dex.

To summarise our findings, we find a robust anti-correlation between  $\alpha$  abundances and close binary fraction, similar in strength but separate from the already established anti-correlation with  $[\text{Fe}/\text{H}]$ . The general trend of increasing  $\alpha$  abundances to decrease the close binary fraction is robust, but the details are complex, and it is likely that the effect is not completely independent from  $[\text{Fe}/\text{H}]$ , at least in some regimes. Because of this, it might not be possible to provide a simple quantitative description of the full relationship between stellar chemistry and close binary fraction in the APOGEE sample.

## 4 DISCUSSION

In Section 3 we have shown that the relationship between stellar multiplicity and stellar parameters is quite complex. A robust physical interpretation of the observed anti-correlation between  $\alpha$ -process abundances and close binary fraction thus requires careful consideration of potential systematics and internal correlations. In this section, we address two such effects that were not discussed in Sections 3.2 and 3.3, and we consider the implications that our results have for star formation.

### 4.1 Potential Systematics

**Visit Histories** Most of our targets have sparsely sampled RV curves. Among the non-SB2s, 36 per cent of objects have only two visits, and 43 per cent have three. For the suspected SB2s, 43 per cent, 35 per cent, and 22 per cent having 2, 3, and 4+ visits, respectively. This is expected; as discussed in Sec. 2, fitting a single stellar template to an SB2 can bias the fit parameters (El-Badry et al. 2018a) and also result in poorer fits overall, which are then flagged in the various APOGEE bitmasks. Because we make quality cuts on these bitmasks, we expect that fewer SB2 visits might pass our quality cuts than the overall sample, though we emphasize that our stringent cuts in S/N are still in place. In both cases, objects with duplicate `allStar` entries (as discussed in the first paragraph of Sec. 2) are biased towards more visits ( $\sim 65$  per cent with 4+ visits) and longer baselines than non-duplicated APOGEE IDs. For each  $[\text{Fe}/\text{H}]$  bin used in Figs. 7–8, we compared the low- and high- $\alpha$  subsamples across histograms of the baselines,  $\text{JD}_N - \text{JD}_1$ ; the median of the time lags between visits for each star,  $\text{median}(\text{JD}_{i+1} - \text{JD}_i)$ ; and the mean of the time lags between visits for each star,  $\text{mean}(\text{JD}_{i+1} - \text{JD}_i)$ , where JD is the Julian date of each observation for a star with  $N$  total visits. There does not appear to be any significant variation in these parameters with  $[\text{Fe}/\text{H}]$  or  $\alpha$  abundances. The fractions of stars with 2, 3, and 4+ visits for each  $[\text{Fe}/\text{H}]$  bin and low- and high- $\alpha$  subsample are also consistent with those for the entire sample.

**White Dwarf Pollution** Some portion of our sample may be post-common envelope systems with white dwarf companions, rather than two MS stars or a subgiant-MS pair. The fraction of these systems will vary with the age of the stellar population, but for short-period ( $\log(P/\text{days}) \leq 4$ ), it is  $\sim 15$  per cent at 1 Gyr and  $\sim 30$  per cent at 10 Gyr (Moe & Di Stefano 2017, see their Section 8.3 and Fig. 29). Considering our median sample age  $\tau \sim 8$  Gyr, we expect a fraction of roughly 25 per cent white dwarf companions in our sample. This fraction will also depend upon the metallicity of the stars, but it cannot explain the factor of 1.5–2 difference we see in the close binary fractions of high- $\alpha$  and low- $\alpha$  samples.

### 4.2 Implications for Binary Star Formation

Close binaries ( $a < 10$  au) likely formed via fragmentation, accretion, and inward migration in the disk, whereas wide binaries ( $a > 200$  au) probably formed via fragmentation of molecular cores (Fisher 2004; Kratter & Lodato 2016; Moe & Di Stefano 2017; Tokovinin 2017; Tokovinin & Moe 2020).

Fragmentation of molecular cores is relatively insensitive to opacity (Bate 2014), explaining why both the initial mass function (Kroupa et al. 2013) and wide binary fraction beyond  $a > 200$  au (Moe et al. 2019; El-Badry & Rix 2018) are metallicity invariant across  $-1.0 < [\text{Fe}/\text{H}] < 0.5$ . A natural consequence of such a model is that the close binary fraction increases with primary mass because massive protostellar disks are more prone to fragmentation (Kratter & Matzner 2006). Analytical models and hydrodynamic simulations also show that the propensity for disk fragmentation decreases with metallicity due to two compounding effects (Machida et al. 2009; Tanaka & Omukai 2014; Moe et al. 2019). First, optically thin cores on large spatial scales radiate via molecular transitions, and so metal-poor cores are systematically hotter and must achieve higher masses in order to collapse into disks. The systematically higher core masses toward lower metallicities do however lead to higher accretion rates onto the disks, promoting gravitational instability (Machida et al. 2009). Second, for solar abundances, protostellar disks massive enough to undergo gravitational instability are optically thick (Rafikov 2005; Clarke 2009; Kratter et al. 2010b). Decreasing the disk’s metallicity decreases its optical depth, allowing the mid-plane to radiate and cool more effectively, stimulating disk fragmentation (Tanaka & Omukai 2014; Moe et al. 2019). Note that Bate (2019) has posited a more complex explanation for the increased close binary fraction at low metallicity. While some increase in disk fragmentation is observed in the simulations, Bate (2019) also observes that metal poor cores fragment on very small scales, where the gas is also optically thick. Moreover, due to the very high rate of dynamical interactions observed in these simulations, far more interchanges between core fragmentation and disk fragmentation binaries are observed. The initial conditions in such simulations may not be representative of lower density star clusters in the solar neighbourhood.

We confirm that the close binary fraction decreases with  $[\text{Fe}/\text{H}]$ , consistent with previous observational surveys and theoretical models. Moreover, we demonstrate for the first time that the close binary fraction decreases more rapidly with  $\alpha$  than Fe for  $[\alpha/\text{Fe}] < 0.05$  dex, consistent with expectations from the two compounding effects described above. For example, optically thin cores radiate mainly through molecular CO transitions, and so the infall rates onto the disk are mainly set by  $\alpha$  abundances. In the cold ( $T < 150\text{K}$ ) midplane of disks prone to fragmentation, opacities are dominated by dust and in particular ice covered grains, which can comprise roughly 60 per cent of the solid particles volume; refractory organics are the second most important contributor in this regime (Semenov et al. 2003). While the optical properties of grains still depend on the distribution and topology of Fe, the changing abundances of O and Si will play a larger role in the bulk opacity. The disk’s temperature profile and probability of fragmentation is therefore more dependant on  $\alpha$  abundances, explaining why the close binary fraction is anti-correlated with O and Si to a larger degree than with Fe.

For  $[\alpha/\text{Fe}] > 0.05$  dex, a different picture emerges whereby the close binary fraction within  $a < 10$  au flattens to 10 per cent, independent of chemical abundance. This “floor” of a 10 per cent close binary fraction appears to be universal. For example, although the close binary fraction

increases from 15 per cent for K-dwarfs to 30 per cent for A-dwarfs, the close binary fraction of both M-dwarfs and brown dwarfs is 10 per cent, relatively constant across  $M_1 = 0.05 - 0.6 M_\odot$  (Joergens 2008; Moe & Di Stefano 2017; Murphy et al. 2018; Winters et al. 2019; Moe 2019). One possible explanation is that at least 10 per cent of protostellar disks become massive or cool enough to fragment early in their accretion evolution, regardless of their chemical composition or the final primary mass. Another possibility is that metal-rich and/or low-mass disks are entirely unsusceptible to fragmentation, and the floor of a 10 per cent close binary fraction is actually due to the small fraction of cores that fragment on large scales and subsequently decay to  $a < 10$  au via dynamical friction or exchange interactions (Lee et al. 2019; Bate 2019). In the future, measurements of how the close binary fraction of M-dwarfs changes with Fe and  $\alpha$  will help differentiate between these two scenarios.

Another consequence of these metallicity trends is that the overall companion distribution becomes skewed towards shorter separations with decreasing metallicity (see Fig. 19 in Moe et al. 2019). However, based on the DR13 sample of SDSS-APOGEE RV variables and *Kepler* eclipsing binaries, Moe et al. (2019) found that the separation distribution of solar-type binaries across  $a = 0.02 - 10$  au does not vary with metallicity at a statistically significant level. With our larger DR14 sample of SDSS-APOGEE RV variables, we find that  $\alpha$ -poor binaries are skewed toward larger  $\Delta RV_{\max}$  and thus shorter separations. With decreasing metallicity, models suggest that disks are not only more likely to fragment, but that disk fragmentation occurs at smaller separations (Machida et al. 2009; Moe et al. 2019). For example, Fig. 10 of Machida et al. (2009) shows that fragmentation occurs near 200 au at solar-metallicity but near 1 au for Population III stars. Similarly, according to Fig. 20 of Moe et al. (2019), disks with solar-metallicity are stable within  $a < 30$  au, but metal-poor disks with  $Z = 10^{-3} Z_\odot$  are capable of fragmentation near  $a = 8$  au. Our measurements are qualitatively consistent with these models, demonstrating that  $\alpha$ -poor stars not only have a higher close binary fraction, but that those close binaries are skewed toward shorter separations. Note that at very low metallicities, one might expect the relative importance of Fe vs  $\alpha$  elements to shift; if fragmentation is pushed to closer separations and correspondingly higher temperatures, the relative importance of ices and organics decreases, and the overall iron abundance might become more important (Semenov et al. 2003).

## 5 CONCLUSIONS

We have presented an analysis of the complex relationship between stellar multiplicity and stellar parameters, with an emphasis on the trends for various  $\alpha$ -process abundances. We defined a sample of 41,363 dwarf and subgiant stars from APOGEE DR14 with well-measured stellar parameters and at least 2 RV measurements. Because most objects in our sample have sparsely-sampled RV curves, we applied a threshold on the maximum RV shift,  $\Delta RV_{\max} \geq 3 \text{ km s}^{-1}$ , to calculate a fraction of RV variables. This fraction is a tracer for the close binary fraction, modulo a completeness correction that can be estimated for the APOGEE observing epochs using a Monte Carlo method with an assumed pe-

riod distribution. We analysed these completeness-corrected close binary fractions alongside a variety of stellar parameters:  $T_{\text{eff}}$ ,  $M$ , [Fe/H], [Mg/H], [Si/H],  $\tau$ ,  $J_z$ , and  $v_R$ . We report a strong anti-correlation between the close binary fraction and Mg and Si abundances, similar in strength but separate from the known anti-correlation with [Fe/H]. Other stellar parameters like  $T_{\text{eff}}$  and  $M$  also have an impact on the close binary fraction, but chemical composition is clearly the main driver of multiplicity trends in our APOGEE sample.

We further investigated the relationship between [Fe/H],  $\alpha$ -process abundances, and stellar multiplicity, measuring a slightly steeper anti-correlation between [Fe/H] and the close binary fraction across the narrower interval  $-0.4 \leq [\text{Fe}/\text{H}] \leq 0.3$  than the average slope across  $-1.0 \leq [\text{Fe}/\text{H}] \leq 0.5$  reported by Moe et al. (2019), similar to the trend found by El-Badry & Rix (2018). The observed anti-correlations between the close binary fraction and  $\alpha$ -process abundances ( $\alpha$ , O, Mg, Si) are [Fe/H]- and abundance-dependant in strength and consistency. Mg and Si in particular showed exceptionally large close binary fractions and remarkable  $\Delta RV_{\max}$  distributions, where the cores almost disappeared entirely. We also find evidence for a correlation and anti-correlation between the close binary fraction and  $[\alpha/\text{H}]$  and [Si/H] with a narrow range of our parameter space. The anti-correlation between stellar composition and close multiplicity fraction has a basis in stellar formation theory. However, low- $\alpha$  binaries are also expected to be skewed towards shorter separations, which would also result in an excess of RV variables independent of an increase in the close binary fraction. Future studies of the period distribution as a function of metallicity and chemistry will help clarify the magnitude of these two effects within our measurements.

## ACKNOWLEDGEMENTS

We are grateful to Jennifer van Saders, Jo Bovy, Jeffrey Newman, and Christian Hayes for discussions. This project was presented to the 2019 APOGEE Stellar Companions Sprint, hosted by the University of Virginia Department of Astronomy, and the attendees provided useful feedback and discussions.

C.N.M. acknowledges support from Scialog Scholar grant 24215 from the Research Corporation. C.N.M. and C.B. acknowledge support from the National Science Foundation grant AST-1909022. S.E.K. and M.G.W. acknowledge support from the National Science Foundation grant AST-1909584. M.M. and K.M.K. acknowledge support from NASA grant 80NSSC18K0726 and The Research Corporation for Science Advancement grant ID# 26077. B.A., H.M.L., and S.R.M. acknowledge support from National Science Foundation grant AST-1616636. N.D. would like to acknowledge that this material is based upon work supported by the National Science Foundation under Grant No. 1616684.

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 acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is [www.sdss.org](http://www.sdss.org).

SDSS 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, the Korean Participation Group, 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.

*Software acknowledgements:* Astropy (Robitaille et al. 2013), Jo Bovy’s apogee tools (Bovy 2016), Ipython (Perez & Granger 2007), Matplotlib (Hunter 2007)

## DATA AVAILABILITY

The stellar parameters, abundances and RVs from APOGEE DR14 were derived from the `allStar` file at [https://www.sdss.org/dr14/irspec/spectro\\_data/](https://www.sdss.org/dr14/irspec/spectro_data/). The Sanders & Das (2018) catalogue can be found at [https://www.ast.cam.ac.uk/jls/data/gaia\\_spectro.hdf5](https://www.ast.cam.ac.uk/jls/data/gaia_spectro.hdf5), and details for downloading our likely SB2s are provided in Appendix C. Data from the MC sampler are available from C.N.M. upon reasonable request.

## REFERENCES

- Abolfathi B., et al., 2018, *ApJS*, 235, 19  
 Albareti F. D., et al., 2017, *ApJS*, 233, 25  
 Badenes C., Maoz D., 2012, *ApJ*, 749, L11  
 Badenes C., et al., 2018, *ApJ*, 854, 147  
 Bate M. R., 2014, *MNRAS*, 442, 285  
 Bate M. R., 2019, *MNRAS*, 484, 2341  
 Blanton M. R., et al., 2017, *AJ*, 154, 28  
 Bovy J., 2016, *ApJ*, 817, 49  
 Breivik K., Price-Whelan A. M., D’Orazio D. J., Hogg D. W., Johnson L. C., Moe M., Morton T. D., Tayar J., 2019, preprint(arXiv:1903.05094)  
 Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, *ApJ*, 823, 102  
 Clarke C. J., 2009, *MNRAS*, 396, 1066  
 De Marco O., Izzard R. G., 2017, *Publ. Astron. Soc. Australia*, 34  
 Dotter A., 2016, *ApJS*, 222, 8  
 Duchêne G., Kraus A., 2013, *ARA&A*, 51, 269  
 Eggleton P. P., 1983, *ApJ*, 268, 368  
 El-Badry K., Rix H.-W., 2018, *MNRAS*, 482, L139  
 El-Badry K., Rix H.-W., Ting Y.-S., Weisz D. R., Bergemann M., Cargile P., Conroy C., Eilers A.-C., 2018a, *MNRAS*, 473, 5043  
 El-Badry K., et al., 2018b, *MNRAS*, 476, 528  
 Fisher R. T., 2004, *ApJ*, 600, 769  
 Gao S., Liu C., Zhang X., Justham S., Deng L., Yang M., 2014, *ApJ*, 788, L37  
 Gao S., Zhao H., Yang H., Gao R., 2017, *MNRAS*, 469, L68  
 García Pérez A. E., et al., 2016, *AJ*, 151, 144  
 Grether D., Lineweaver C. H., 2007, *ApJ*, 669, 1220  
 Gunn J. E., et al., 2006, *AJ*, 131, 2332  
 Hayden M. R., et al., 2015, *ApJ*, 808, 132  
 Hekker S., Snellen I. A. G., Aerts C., Quirrenbach A., Reffert S., Mitchell D. S., 2008, *A&A*, 480, 215  
 Holtzman J. A., et al., 2015, *AJ*, 150, 148  
 Holtzman J. A., et al., 2018, *AJ*, 156, 125  
 Hunter J. D., 2007, *CSE*, 9, 90  
 Husser T.-O., von Berg S. W., Dreizler S., Homeier D., Reiners A., Barman T., Hauschildt P. H., 2013, *A&A*, 553, A6  
 Iben Jr. I., Tutukov A. V., 1984, *ApJ*, 284, 719  
 Jayasinghe T., et al., 2020, preprint(arXiv:2006.10057)  
 Joergens V., 2008, *A&A*, 492, 545  
 Jönsson H., et al., 2018, *AJ*, 156, 126  
 Jönsson H., et al., 2020, preprint(arXiv:2007.05537)  
 Kounkel M., et al., 2019, *AJ*, 157, 196  
 Kratter K., Lodato G., 2016, *ARA&A*, 54, 271  
 Kratter K. M., Matzner C. D., 2006, *MNRAS*, 373, 1563  
 Kratter K. M., Matzner C. D., Krumholz M. R., Klein R. I., 2010a, *ApJ*, 708, 1585  
 Kratter K. M., Murray-Clay R. A., Youdin A. N., 2010b, *ApJ*, 710, 1375  
 Kroupa P., Weidner C., Pflamm-Altenburg J., Thies I., Dabringhausen J., Marks M., Maschberger T., 2013, in , *Planets, Stars and Stellar Systems*. Springer Netherlands, pp 115–242, doi:10.1007/978-94-007-5612-0\_4  
 Lada C. J., 2006, *ApJ*, 640, L63  
 Lee A. T., Offner S. S. R., Kratter K. M., Smullen R. A., Li P. S., 2019, *ApJ*, 887, 232  
 Liu C., 2019, *MNRAS*, 490, 550  
 Machida M. N., Omukai K., Matsumoto T., Inutsuka S.-I., 2009, *MNRAS*, 399, 1255  
 Mackereth J. T., et al., 2017, *MNRAS*, 471, 3057  
 Mackereth J. T., et al., 2019, *MNRAS*, 489, 176  
 Majewski S. R., et al., 2017, *AJ*, 154, 94  
 Maoz D., Badenes C., Bickerton S. J., 2012, *ApJ*, 751, 143  
 Miglio A., et al., 2020, arXiv e-prints  
 Moe M., 2019, *Mem. Soc. Astron. Italiana*, 90, 347  
 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  
 Murphy S. J., Moe M., Kurtz D. W., Bedding T. R., Shibahashi H., Boffin H. M. J., 2018, *MNRAS*, 474, 4322  
 Ness M., Hogg D. W., Rix H.-W., Martig M., Pinsonneault M. H., Ho A. Y. Q., 2016, *ApJ*, 823, 114  
 Nidever D. L., et al., 2015, *AJ*, 150, 173  
 Paczynski B., 1971, *ARA&A*, 9, 183  
 Pawlak M., et al., 2019, *MNRAS*, 487, 5932  
 Paxton B., Bildsten L., Dotter A., Herwig F., Lesaffre P., Timmes F., 2011, *ApJS*, 192, 3  
 Paxton B., et al., 2013, *ApJS*, 208, 4  
 Paxton B., et al., 2015, *ApJS*, 220, 15  
 Perez F., Granger B. E., 2007, *CSE*, 9, 21  
 Pinsonneault M. H., et al., 2018, *ApJS*, 239, 32  
 Price-Whelan A. M., et al., 2018, *AJ*, 156, 18  
 Price-Whelan A. M., et al., 2020, *ApJ*, 895, 2  
 Rafikov R. R., 2005, *ApJ*, 621, L69  
 Raghavan D., et al., 2010, *ApJS*, 190, 1  
 Robitaille T. P., et al., 2013, *A&A*, 558, A33  
 Sanders J. L., Das P., 2018, *MNRAS*, 481, 4093



- Semenov D., Henning T., Helling C., Ilgner M., Sedlmayr E., 2003, *A&A*, 410, 611
- Stanway E. R., Eldridge J. J., Chrimes A. A., 2020, *MNRAS*, 497, 2201
- Suda T., et al., 2013, *MNRAS*, 432, L46
- Tanaka K. E. I., Omukai K., 2014, *MNRAS*, 439, 1884
- Ting Y.-S., Conroy C., Rix H.-W., Cargile P., 2019, *ApJ*, 879, 69
- Tokovinin A., 2017, *MNRAS*, 468, 3461
- Tokovinin A., Moe M., 2020, *MNRAS*, 491, 5158
- Vera-Ciro C., D’Onghia E., Navarro J., Abadi M., 2014, *ApJ*, 794, 173
- Weinberg D. H., et al., 2019, *ApJ*, 874, 102
- Wilson J. C., et al., 2019, *PASP*, 131, 055001
- Winters J. G., et al., 2019, *AJ*, 157, 216
- Yuan H., Liu X., Xiang M., Huang Y., Chen B., Wu Y., Hou Y., Zhang Y., 2015, *ApJ*, 799, 135
- Zasowski G., et al., 2013, *AJ*, 146, 81
- Zasowski G., et al., 2017, *AJ*, 154, 198
- de Mink S. E., Belczynski K., 2015, *ApJ*, 814, 58

## APPENDIX A: CHOICES FOR THE MONTE CARLO SIMULATION

We altered several of the choices listed in Sec. 3.1 to gauge their effects on the completeness estimate. The first is our choice of  $\log(g)$ ; it only affects the calculations for the critical period  $P_{\text{crit}}$  and circularisation period  $P_{\text{circ}}$  (Badenes et al. 2018). The critical period is calculated using

$$P_{\text{crit}} = \frac{2\pi}{\sqrt{R^3(q)(1+q)}} \left( \frac{GM}{g^3} \right)^{1/4} \quad (\text{A1})$$

where  $M$  is the mass of the primary in grams,  $g$  is the primary’s surface gravity in  $\text{cm s}^{-2}$ ,  $q$  is the system’s mass ratio, and  $R(q)$  is the ratio between the radius of the Roche Lobe and the orbital separation (Eggleton 1983). The circularisation period is calculated for  $\log(g/\text{cm s}^{-2})=4.25$ ,  $M = 1.0M_{\odot}$ , and  $[\text{Fe}/\text{H}]=0.0$  dex. For  $1M_{\odot}$  and  $\log(g/\text{cm s}^{-2})=4.25$ , we calculate  $\log(P_{\text{circ}}/\text{days}) = 0.888$  and  $\log(P_{\text{crit}}/\text{days}) = -0.313$ . Running the MC with 2x and 0.5x these values resulted in no significant change to the completeness estimate, which is expected; given the Raghavan et al. (2010) period distribution, very few stars will be at short enough periods to be affected by shifts of this magnitude in either of these parameters.

We implemented a 25 per cent increase in twins ( $0.95 \geq q \geq 1.0$ ) motivated by the results of Moe & Di Stefano (2017). Reducing this fraction, even to 0 per cent, did not significantly change our completeness estimate.

Fully characterising the observed RV uncertainty distribution is a difficult task, especially given the diverse nature of our sample. Consequently, our choice for a simulated RV error distribution is not motivated by physical intuition; rather, we have tried a variety of possible distributions and distribution parameters in order to roughly reproduce the shape of the observed  $\Delta\text{RV}_{\text{max}}$  distribution. We have found the Student’s t distribution to reproduce the relative shape of the core and tail for a variety of  $\log(g)$  values, and so we chose to use it in the completeness estimate described in Section 3.1. This choice does not significantly affect our completeness estimates, however. Table A1 gives completeness fractions given two  $\Delta\text{RV}_{\text{max}}$  thresholds for three RV error distributions:

- (i) a Student’s t (scipy.stats.t) with degrees of freedom 3.5,  $\mu = 0.0$ ,  $\sigma = 0.25 \text{ km s}^{-1}$  (the same used in Section 3)
- (ii) a Gaussian (scipy.stats.norm) with  $\mu = 0.0$ ,  $\sigma = 0.25 \text{ km s}^{-1}$
- (iii) a constant RV error of  $\sigma_{\text{RV}} = 0.25 \text{ km s}^{-1}$  applied to every simulated RV

The simulated  $\Delta\text{RV}_{\text{max}}$  distribution for each choice are shown alongside the observed distribution in Fig. A1. The completeness fractions are nearly identical between the Student’s t and the Gaussian, and while the Student’s t has a larger percentage of false positives, it is still a modest increase. The constant RV uncertainty has very similar detection efficiencies to the other two at the low- $P$  end, where the RV variability is the largest. As expected, it diverges for larger periods, where it fails to distinguish between a core and tail in the  $\Delta\text{RV}_{\text{max}}$  distribution, apparent in Fig. A1. Between the Student’s t and Gaussian, the  $\Delta\text{RV}_{\text{max}}$  distributions are qualitatively similar. The Gaussian is slightly narrower with a sharper transition between the core and tail, whereas the Student’s t has a slightly better match to the overall shape of our observed  $\Delta\text{RV}_{\text{max}}$  distribution. For this reason, we chose the Student’s t, though we note that this reason is purely qualitative and our choice does not significantly affect the completeness fractions. Future work to better understand the RV uncertainties may favour one distribution over another, but that is beyond the scope of this work.

From theoretical predictions (see Section 4.2) and recent observations of eclipsing binaries (Jayasinghe et al. 2020), the period distribution for solar-type binaries may depend on chemistry. We briefly explored this scenario by implementing a three-component period distribution. Motivated by Fig. 19 of Moe et al. (2019), we simulated a sample of  $N = 50,000$  stars using three chemistry-dependant log-normal period distributions:

- (i)  $-1 \leq [\text{Fe}/\text{H}] \leq -0.2$ :  $\{\mu_{\log(P/\text{days})} = 4.0, \sigma_{\log(P/\text{days})} = 1.5\}$
- (ii)  $-0.2 \leq [\text{Fe}/\text{H}] \leq 0.2$ :  $\{\mu = 5.0, \sigma = 2.0\}$
- (iii)  $[\text{Fe}/\text{H}] \geq 0.2$ :  $\{\mu = 6.0, \sigma = 2.5\}$

where  $[\text{Fe}/\text{H}]$  values are given in dex. We simulated  $N_1 = 6,250$ ,  $N_2 = 37,500$ ,  $N_3 = 6,250$ , which is proportional to the number of stars in our sample within those  $[\text{Fe}/\text{H}]$  ranges. The calculated completeness fractions for each MC subsample do not vary significantly from those shown in Table 1, and when we combine the subsamples into a single  $N = 50,000$  sample, the calculated completeness fractions also do not vary significantly. Changing the subsample sizes to (15,000; 20,000; 15,000) also did not result in significantly different completeness fractions.

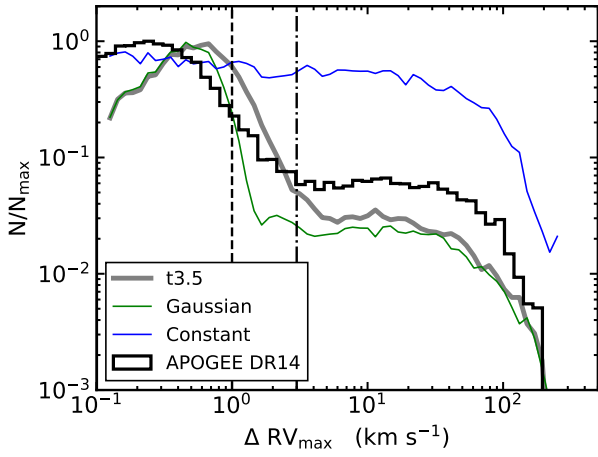
## APPENDIX B: ADDITIONAL FIGURES

## APPENDIX C: FORMAT OF DATA PRODUCTS

Here, we make available the likely SB2s identified by two methods as discussed in Sec. 2. Table C1 describes the column structure, and Tables C2-C3 provide the results from each method. Each entry in Tables C2-C3 is for an individual visit spectrum; the table from re-analysing APOGEE CCFs has 13,970 total entries, and the table from re-calculating the

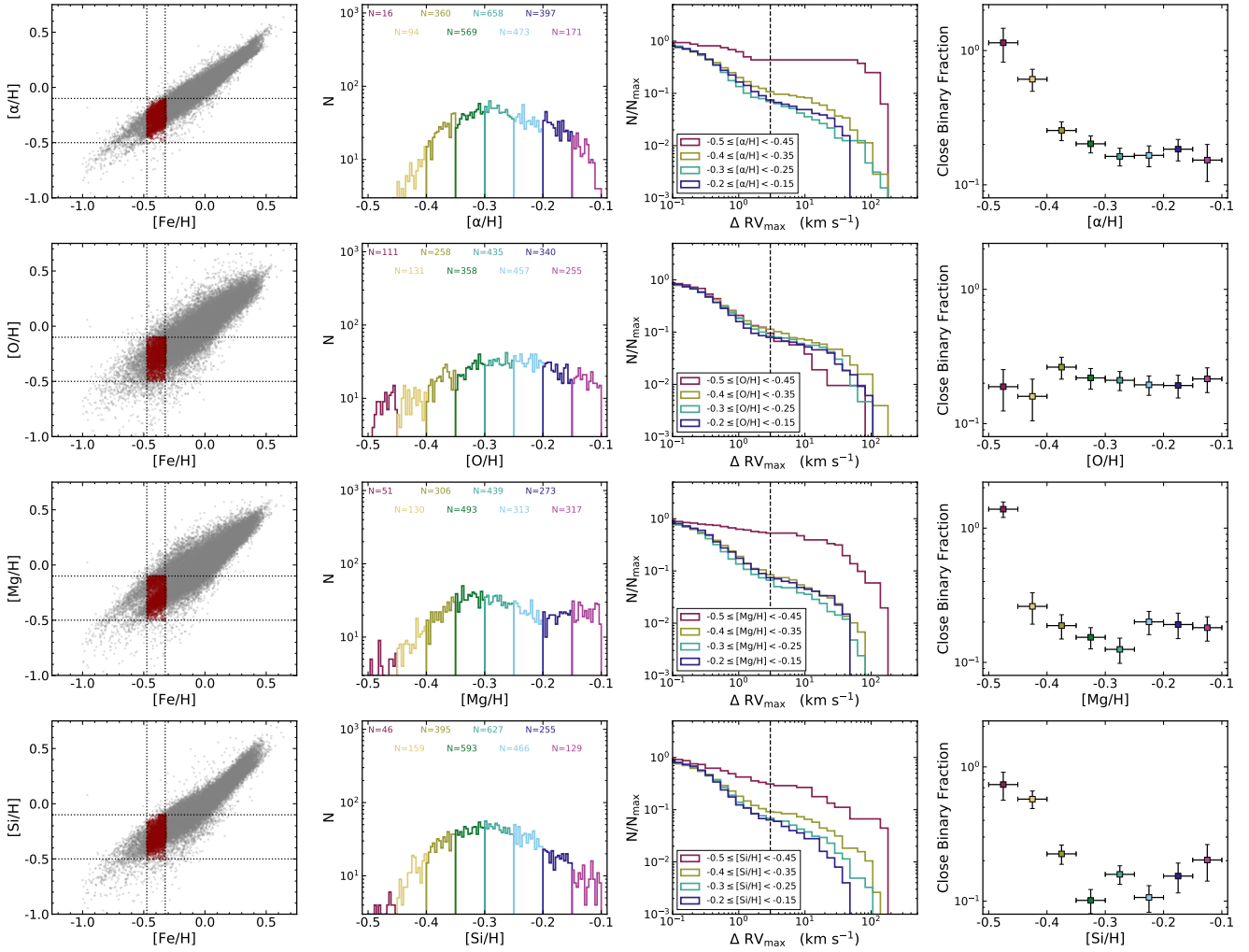
**Table A1.** Completeness and false positive fractions for selected  $\log(P/\text{days})$  and  $\Delta RV_{\text{max}}$  thresholds, given three RV error distributions all with  $\sigma = 0.25 \text{ km s}^{-1}$ .

log( $P/\text{days}$ ) threshold	$\Delta RV_{\text{max}} \geq 1 \text{ km s}^{-1}$			$\Delta RV_{\text{max}} \geq 3 \text{ km s}^{-1}$		
	t3.5	Gaussian	Constant	t3.5	Gaussian	Constant
$\log P \leq 0.0$	1.00	1.00	1.00	0.96	0.97	1.00
$\log P \leq 2.0$	0.93	0.94	0.93	0.84	0.84	0.85
$\log P \leq 4.0$	0.55	0.50	0.64	0.34	0.34	0.47
False Positive Fraction	2.68%	1.20%	0.0%	0.10%	0.0%	0.0%



**Figure A1.** Distributions of  $\Delta RV_{\text{max}}$  from our APOGEE DR14 sample (black) and simulated by our MC with three RV error distributions. All three RV error distributions have  $\sigma = 0.25 \text{ km s}^{-1}$ , and the Gaussian and Student's  $t$  with degrees of freedom 3.5 (t3.5) both have  $\mu = 0$ .

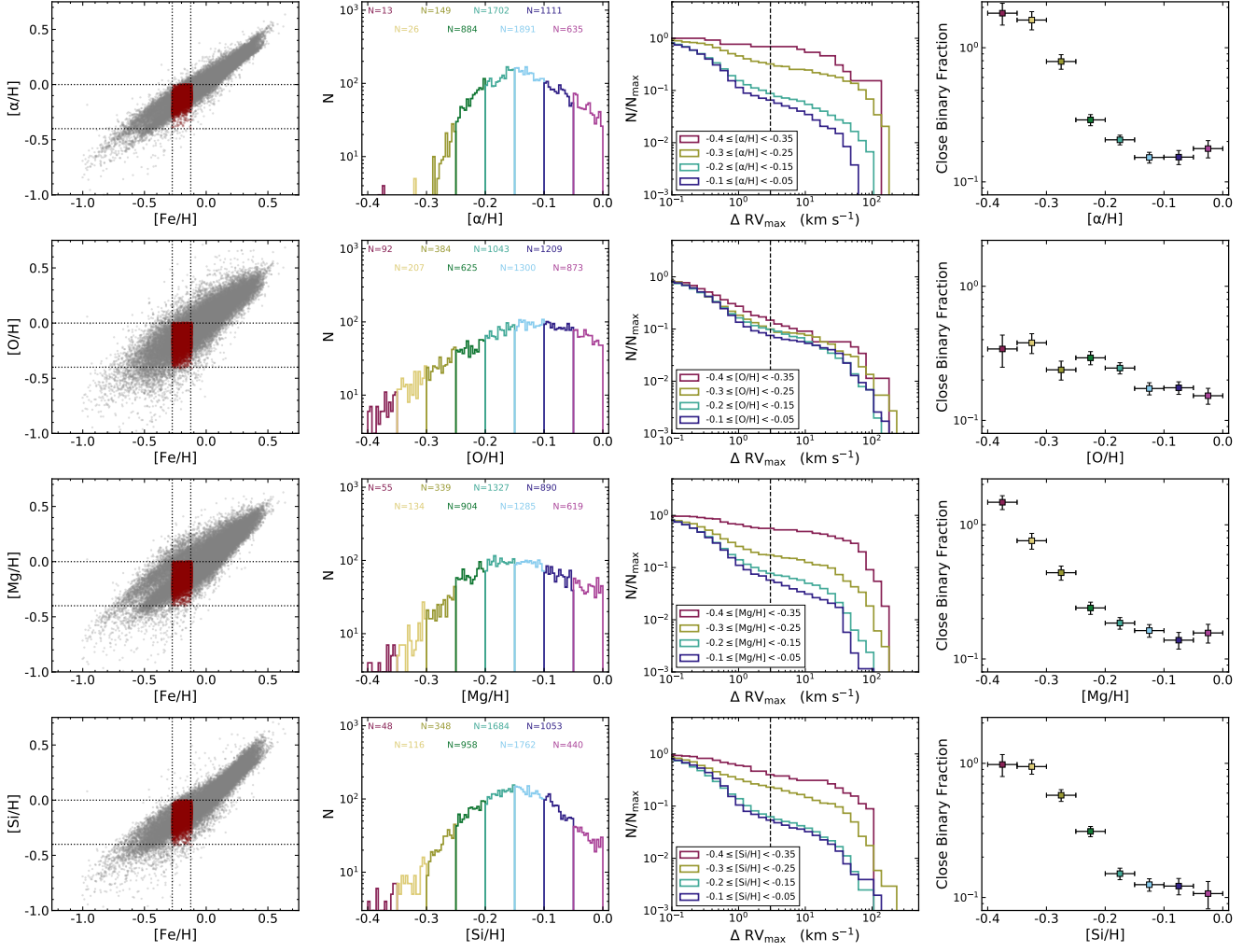
CCFs has 12,044. There are 2832 and 2238 unique APOGEE IDs in each table, respectively.



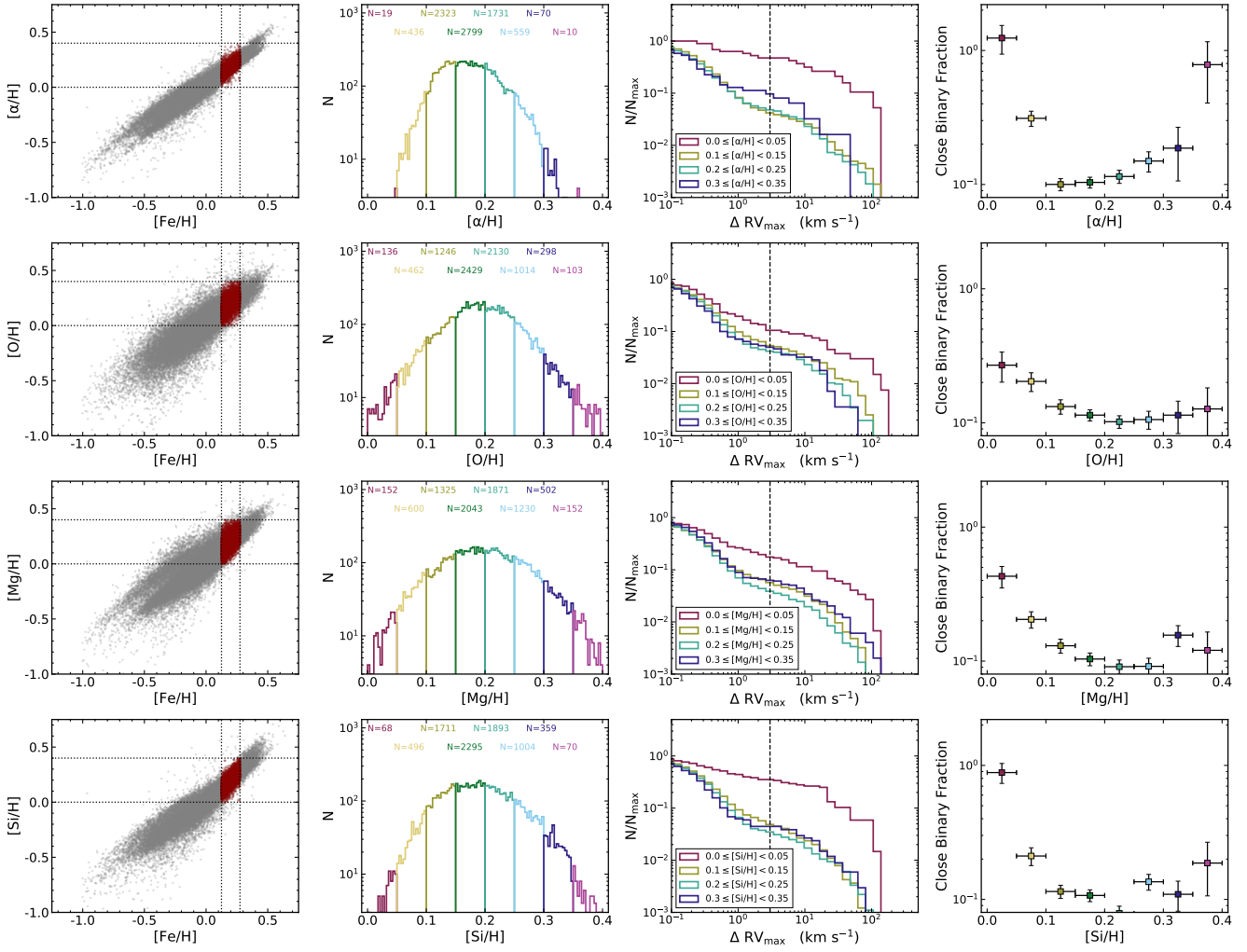
**Figure B1.** Identical to Fig. 10 but with boundaries  $-0.475 \leq [Fe/H] \leq -0.325$  and  $-0.5 \leq X \leq -0.1$ .

**Table C1.** Format of provided SB2 catalogues. Listed below are the name of each column with a description and any applicable units. For each entry that is given by an array, the array has 8 elements.

Name	Description	Units
OBJID	identifier in the APOGEE catalogue	—
PLATE	APOGEE visit plate ID	—
FIBER	APOGEE visit fiber ID	—
MJD	Modified Julian Date of APOGEE visit	—
N	number of deconvolved components	—
FLAG	array of integer quality flags, from 1.0 to 4.0	—
POS	array of RVs for each component, ordered by amplitude	km s $^{-1}$
AMP	array of amplitudes	—
FWH	array of full widths at half maximum	km s $^{-1}$
EPOS	array of RV uncertainties	km s $^{-1}$
EAMP	array of amplitude uncertainties	—
EFWH	array of full width at half maximum uncertainties	km s $^{-1}$



**Figure B2.** Identical to Fig. 10 but with boundaries  $-0.275 \leq [\text{Fe}/\text{H}] \leq -0.125$  and  $-0.4 \leq X \leq 0.0$ .



**Figure B3.** Identical to Fig. 10 but with boundaries  $0.125 \leq [\text{Fe}/\text{H}] \leq 0.275$  and  $0.0 \leq X \leq 0.4$ .

**Table C2.** Table of likely SB2s identified by re-analysing the APOGEE CCFs. This table is available in its entirety (with parameters for 13,970 visits with 2832 unique APOGEE IDs) in machine-readable form.

OBJID	PLATE	FIBER	MJD	N	FLAG	POS	AMP	FWH	EPOS	EAMP	EFWH
2M00023036+8524194	7950	224	57295	1	4.0 ... 0.0	-4.5720005 ... nan	0.5707569 ... nan	49.42374 ... nan	0.57283086 ... nan	0.013487913 ... nan	1.3489138 ... nan
2M00023036+8524194	9084	224	57556	2	4.0 ... 0.0	-71.81711 ... nan	0.37271407 ... nan	34.15913 ... nan	0.546208 ... nan	0.012155449 ... nan	1.2862214 ... nan
2M00023179+1521164	6560	77	56584	2	4.0 ... 0.0	20.618723 ... nan	0.49165606 ... nan	54.235355 ... nan	1.8823811 ... nan	0.014511261 ... nan	3.0066836 ... nan
2M00023179+1521164	6560	89	56588	1	4.0 ... 0.0	3.9876366 ... nan	0.54505944 ... nan	62.63368 ... nan	0.6144675 ... nan	0.010904319 ... nan	1.4469602 ... nan
...	...	...	...	...	...	...	...	...	...	...	...

**Table C3.** Table of likely SB2s identified by calculating our own CCFs. This table is available in its entirety (with parameters for 12,044 visits with 2238 unique APOGEE IDs) in machine-readable form.

OBJID	PLATE	FIBER	MJD	N	FLAG	POS	AMP	FWH	EPOS	EAMP	EFWH
2M00023036+8524194	5095	233	55821	2	4.0 ... 0.0	50.38858 ... nan	0.3384734 ... nan	33.308804 ... nan	0.51698554 ... nan	0.010713303 ... nan	1.217408 ... nan
2M00023036+8524194	5095	230	55824	2	4.0 ... 0.0	-74.34177 ... nan	0.32860735 ... nan	31.727573 ... nan	0.54551804 ... nan	0.011521589 ... nan	1.2845968 ... nan
2M00023036+8524194	5095	230	55840	2	4.0 ... 0.0	66.387726 ... nan	0.29461184 ... nan	20.609297 ... nan	0.94517577 ... nan	0.027554285 ... nan	2.2257187 ... nan
2M00023036+8524194	5095	239	55844	2	4.0 ... 0.0	42.22613 ... nan	0.31082585 ... nan	24.475822 ... nan	0.5241617 ... nan	0.013574529 ... nan	1.2343063 ... nan
...	...	...	...	...	...	...	...	...	...	...	...

This paper has been typeset from a  $\text{T}_{\text{E}}\text{X}/\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  file prepared by the author.