



# The GALAH survey: co-orbiting stars and chemical tagging

Jeffrey D. Simpson<sup>1,2★</sup>, Sarah L. Martell<sup>1,3</sup>, Gary Da Costa<sup>4</sup>,  
Andrew R. Casey<sup>5,6</sup>, Ken C. Freeman<sup>4</sup>, Jonathan Horner<sup>7</sup>, Yuan-Sen Ting<sup>8,9,10</sup>,  
David M. Nataf<sup>11</sup>, Geraint F. Lewis<sup>12</sup>, Melissa K. Ness<sup>13</sup>, Daniel B. Zucker<sup>14</sup>,  
Peter L. Cottrell<sup>15,16</sup>, Klemen Čotar<sup>17</sup>, Martin Asplund<sup>3,4</sup>, Joss Bland-Hawthorn<sup>3,12</sup>,  
Sven Buder<sup>18</sup>, Valentina D’Orazi<sup>19</sup>, Gayandhi M. De Silva<sup>2,3,20</sup>, Ly Duong<sup>3,4</sup>,  
Janez Kos<sup>12</sup>, Jane Lin<sup>3,4</sup>, Karin Lind<sup>18,21</sup>, Katharine J. Schlesinger<sup>4</sup>, Sanjib Sharma<sup>12</sup>,  
Tomaž Zwitter<sup>17</sup>, Prajwal R. Kafle<sup>22</sup> and Thomas Nordlander<sup>3,4</sup>

*Affiliations are listed at the end of the paper*

Accepted 2018 November 6. Received 2018 October 10; in original form 2018 June 19

## ABSTRACT

We present a study using the second data release of the GALAH survey of stellar parameters and elemental abundances of 15 pairs of stars identified by Oh et al. They identified these pairs as potentially co-moving pairs using proper motions and parallaxes from *Gaia* DR1. We find that 11 very wide ( $>1$  pc) pairs of stars do in fact have similar Galactic orbits, while a further four claimed co-moving pairs are not truly co-orbiting. Eight of the 11 co-orbiting pairs have reliable stellar parameters and abundances, and we find that three of those are quite similar in their abundance patterns, while five have significant [Fe/H] differences. For the latter, this indicates that they could be co-orbiting because of the general dynamical coldness of the thin disc, or perhaps resonances induced by the Galaxy, rather than a shared formation site. Stars such as these, wide binaries, debris of past star formation episodes, and coincidental co-orbiters, are crucial for exploring the limits of chemical tagging in the Milky Way.

**Key words:** stars: abundances – stars: formation – galaxy: disc.

## 1 INTRODUCTION

The GALactic Archaeology with HERMES (GALAH) survey is a large and ambitious spectroscopic investigation of the local stellar environment (De Silva et al. 2015). One of its principal aims is to determine precise abundances of nearly 30 elements<sup>1</sup> from one million stars and to use chemical tagging to identify dispersed stellar clusters in the field of the disc and halo (for the initial motivating papers, see Freeman & Bland-Hawthorn 2002; Bland-Hawthorn, Krumholz & Freeman 2010). It relies on the assumption that although they may disperse into different regions of kinematic phase space, the stars that form within a single cluster will continue to possess a common and unique pattern of chemical abundances. Chemically tagging the stars from many formation sites would enable us to unravel the formation and evolutionary history of the Galaxy in a way that it is not possible from their spatial, photometric, or kinematic properties.

\* E-mail: [jeffrey.simpson@aao.gov.au](mailto:jeffrey.simpson@aao.gov.au)

<sup>1</sup>While measurements of nearly 30 elements are possible from spectra obtained with HERMES, in GALAH DR2 we report abundances for 23 elements, and in this work we consider only the abundances of 19 elements that were present in our stars of interest.

Chemical tagging solely in abundance space is a challenging task, and there is much discussion in the literature about the prospects of the technique being successful (e.g. Ting, Conroy & Goodman 2015; Bovy 2016; Hogg et al. 2016). Any tagged group identified solely in abundance space would still need to have stellar parameters consistent with a single age in order to be believable, and it would be quite unlikely for their orbits to belong to entirely different Galactic components. The aim of GALAH is to carry out chemical tagging using all of this available information together. Abundance data forms the foundation of this group identification, with stellar parameters and kinematics acting as a confirmation rather than a primary tool. Cases in which we can test for coherence in both kinematics and chemical composition are an important step towards that goal. Stars in streams and moving groups fall between the extremes of stars still in their formation clusters and the majority of disc stars that have lost their original spatial and kinematic coherence. They are a critical test set for chemical tagging since their orbital similarities can provide a confirmation of the shared formation history that we would infer from their compositions.

Spectra obtained in the GALAH survey provide the radial velocities of the stars, but we require full 6D (position, velocity) phase-space information about the stars to place these stars in streams and moving groups. The ESA *Gaia* mission (Prusti et al. 2016)

provides us with this. The first data release (Brown et al. 2016) was utilized by a number of authors to identify potentially co-moving pairs and groups of stars (Andrews, Chanamé & Agüeros 2017; Oelkers, Stassun & Dhital 2017; Oh et al. 2017, hereafter O17). These studies have each adopted different methods and goals for their searches. The work for this paper was primarily performed prior to the release of *Gaia* DR2 (Brown et al. 2018), but we use its parallaxes and proper motions.

We will focus on the pairs of stars identified by O17. These stars are all found within  $\sim 600$  pc of the Sun, which is where the errors in parallax found by the Tycho–*Gaia* astrometric solution (TGAS; Michalik, Lindegren & Hobbs 2015; Lindegren et al. 2016) are small enough<sup>2</sup> to permit a reliable determination of distances and orbits. O17 used the TGAS data to identify over 13 000 pairs of co-moving stars with separations less than 10 pc. Because *Gaia* DR1 did not contain radial velocity information for the stars, they had to marginalize over the unknown 3D velocities of the stars. In their method, each star can be paired with multiple other stars, and many of the pairs they identified were parts of larger networks. Their analysis recovered several known clusters, including the Pleiades, the Sco–Cen young stellar association, the Hyades, and NGC 2632. However, most of their groups do not have a known counterpart in the literature, and many were isolated pairs of stars.

Interestingly, by requiring that the proper motions of the stars be highly similar, O17 might reject close binary star systems as potential co-moving pairs. From the calculations of Andrews et al. (2017), the semi-amplitude of the orbital velocity in the systems would be  $>5 \text{ km s}^{-1}$  for systems with separations  $<15$  au. If a significant component of this motion were oriented in the plane of the sky, it is easy to imagine that the two stars would appear to have a relative proper motion too large to allow them to orbit together, even though in truth they follow their common barycentre around its orbit.

Further investigations of some of the pairs identified by O17 has been done with low-resolution spectroscopy (Price-Whelan, Oh & Spergel 2017) and infrared photometry (Bochanski et al. 2018) and for one pair, high-resolution spectra (Oh et al. 2018, using results from Brewer et al. 2016). Andrews, Chanamé & Agüeros (2018b) found that their candidate wide binary stars typically had very similar metallicity, using the public catalogues from the RAVE and LAMOST spectroscopic surveys. However, LAMOST does not publish detailed abundances based on their low-resolution spectra, and the abundance precision of the RAVE catalogue was not high enough to pursue further chemical tagging.

Using data from the GALAH survey, we can expand on these studies by adding critical information. Not only does high-resolution spectroscopy provide radial velocities that allow complete orbital calculations, GALAH also derives stellar parameters and elemental abundances:  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ , and abundances for up to 23 elements. For these candidate co-moving pairs, which may be from dissolving and disrupting clusters, we can evaluate whether they have common origins with the aid of kinematic and chemical information.

Spectroscopic stellar parameters and abundances from the GALAH survey make it possible for us to distinguish between different types of moving groups. There are known groups that are ‘true’ moving groups of stars, consisting of the disrupted remnants of old clusters: e.g. HR 1614, Wolf 630, and the Argus moving

groups (De Silva et al. 2007; Bubar & King 2010; De Silva et al. 2013). However, there are other groups (e.g. the Hercules group; Bensby et al. 2007; Quillen et al. 2018) that have distinctly different chemical abundances and are on similar orbits as a result of dynamical resonances within the Galaxy. Simply relying on kinematics would identify that these are true groups of co-moving stars but would not provide a full picture of the chemodynamical history of the Galaxy.

This work is structured as follows: data reduction and abundance analysis (Section 2); kinematic evaluation of the groups observed by GALAH (Section 3); investigation of the abundance patterns of the co-orbiting pairs (Section 4); and a discussion of the intrinsic limits of and future prospects for chemical tagging (Section 5).

## 2 OBSERVATIONS AND SPECTRUM ANALYSIS

We make use of an internally released catalogue of a similar size and composition to the GALAH survey’s second data release (GALAH DR2; see the release paper: Buder et al. 2018) that maximized the overlap between GALAH and O17. It is based upon spectra obtained between 2014 January and 2018 January using the 3.9-m Anglo-Australian Telescope with the HERMES spectrograph (Sheinis et al. 2015) and the Two-Degree Field (2dF) top-end (Lewis et al. 2002). The 2dF allows for the concurrent acquisition of up to  $\sim 360$  science targets per exposure. HERMES simultaneously acquires spectra using four independent cameras with non-contiguous wavelength coverage totalling  $\sim 1000 \text{ \AA}$  at a spectral resolving power of  $R \approx 28000$ . Its fixed wavelength bands are 4715–4900, 5649–5873, 6478–6737, and 7585–7887  $\text{ \AA}$ . For details on the observing procedures see Martell et al. (2017) and Buder et al. (2018). The spectra were reduced using an IRAF-based pipeline that was developed specifically for GALAH and optimized for speed, accuracy, and consistency. We direct the reader to Kos et al. (2017) for a detailed description of the reduction procedure.

The GALAH stellar parameter and abundance pipeline description can be found in Buder et al. (2018). Briefly, the pipeline uses a two-step process. In the first step, spectra with high signal to noise are identified and analysed with the spectrum synthesis code Spectroscopy Made Easy (SME; Valenti & Piskunov 1996; Piskunov & Valenti 2017) to determine the stellar labels ( $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ ,  $v_{\text{mic}}$ ,  $v \sin i$ ,  $v_{\text{rad}}$ , and  $[\text{X}/\text{Fe}]$ ). This training set includes the *Gaia* benchmark stars, globular and open cluster stars, and stars with accurate asteroseismic surface gravity from *K2* Campaign 1 (Stello et al. 2017). In the second step, THE CANNON (Ness et al. 2015) learns the training set labels from SME and builds a quadratic model at each pixel of the normalized spectrum as a function of the labels. Abundance estimates are then generated from THE CANNON model.

Overall, the GALAH release used in this work contains a total of 365 516 stars with up to 23 elemental abundances per star. For a minority of stars, the label results from THE CANNON are not reliable: The label result could be too far from the training set, the  $\chi^2$  between the observed spectrum and the spectrum calculated by THE CANNON could be too large, or the spectra could have been classified by t-SNE (for details on the application of t-SNE to GALAH spectra see Traven et al. 2017) as having problems. In addition, the individual elemental abundance can be flagged for similar reasons via `flag_x_fe`. In this work, we only use abundance values for which `flag_x_fe` is zero, which means that the particular abundance is likely to be reliable.

<sup>2</sup>Using the O17 definition of requiring the parallax signal-to-noise ratio  $\varpi/\sigma_{\varpi} > 8$ .

**Table 1.** The stellar parameters and orbital characteristics of the 15 pairs. The stellar parameters are given for stars with `flag_cannon == 0` (i.e. they are not believed to be unreliable). For each star, 1000 random samples of the 6D information of the stars taking into account their uncertainties and covariances were created and then GALPY was used to integrate the orbit to find the median eccentricity,  $z_{\max}$ , perigalacticon, and apogalacticon of the orbit. The uncertainties are the 5th and 95th percentiles of the distributions. The ordering is the same as in Table 2, namely increasing  $\Delta(U, V, W)$ . The online version contains columns for the elemental abundances for each star.

Group ID	subject_id	$T_{\text{eff}}$ (K)	log g	[Fe/H]	$G$	$G_{\text{BP}} - G_{\text{RP}}$	e	$z_{\max}$ (pc)	Peri (kpc)	Apo (kpc)
3410	160813001601030	6106 ± 55	4.28 ± 0.15	-0.47 ± 0.07	10.98	0.73	0.07 ± 0.00	29 ± 1	7.15 ± 0.01	8.31 ± 0.01
3410	160813001601029	5664 ± 73	4.49 ± 0.18	-0.36 ± 0.08	11.86	0.86	0.07 ± 0.00	28 ± 1	7.17 ± 0.02	8.32 ± 0.01
3612	150211003701379	5533 ± 67	4.55 ± 0.17	+0.23 ± 0.08	12.02	0.95	0.12 ± 0.00	87 ± 1	6.81 ± 0.01	8.70 ± 0.02
3612	150211003701380	5371 ± 61	4.51 ± 0.16	+0.23 ± 0.07	12.36	1.01	0.12 ± 0.00	89 ± 1	6.79 ± 0.02	8.66 ± 0.03
3	160125004501147				10.26	1.13	0.05 ± 0.00	48 ± 0	7.12 ± 0.02	7.95 ± 0.00
3	160130006301220				9.79	1.01	0.06 ± 0.00	52 ± 1	7.10 ± 0.03	7.95 ± 0.00
237	170615003401085				9.78	0.44	0.05 ± 0.00	123 ± 2	7.16 ± 0.02	7.89 ± 0.01
237	160817001601245				9.22	0.27	0.05 ± 0.02	136 ± 17	7.21 ± 0.25	7.90 ± 0.05
987	170516000601281	6056 ± 51	4.16 ± 0.14	-0.31 ± 0.06	10.50	0.69	0.13 ± 0.00	284 ± 2	6.78 ± 0.03	8.89 ± 0.03
987	170516000601016	6130 ± 52	4.18 ± 0.14	-0.22 ± 0.06	10.49	0.68	0.12 ± 0.00	287 ± 10	7.14 ± 0.06	9.06 ± 0.04
40	170711001501145	6387 ± 18	4.13 ± 0.06	-0.19 ± 0.02	9.09	0.60	0.06 ± 0.00	57 ± 1	7.11 ± 0.02	7.95 ± 0.01
40	150706001601135				9.66	1.01	0.04 ± 0.01	59 ± 4	7.41 ± 0.16	8.01 ± 0.02
1313	170615003401071	5568 ± 63	4.07 ± 0.17	+0.30 ± 0.08	11.62	0.96	0.11 ± 0.00	206 ± 4	6.42 ± 0.03	8.00 ± 0.00
1313	160815002101306	6219 ± 61	4.09 ± 0.16	-0.26 ± 0.07	10.94	0.71	0.10 ± 0.00	168 ± 6	6.44 ± 0.03	7.89 ± 0.01
3496	160611003101049	6137 ± 54	4.29 ± 0.15	-0.02 ± 0.07	9.74	0.68	0.04 ± 0.00	128 ± 1	7.36 ± 0.02	8.05 ± 0.00
3496	160611003101279	5190 ± 90	4.54 ± 0.20	+0.29 ± 0.09	11.60	1.08	0.05 ± 0.00	124 ± 1	7.15 ± 0.02	7.91 ± 0.00
1220	150703001601389	5733 ± 60	4.49 ± 0.16	+0.11 ± 0.07	10.95	0.93	0.12 ± 0.00	42 ± 2	6.79 ± 0.02	8.72 ± 0.01
1220	170712001601389	5988 ± 57	4.42 ± 0.15	-0.17 ± 0.07	10.73	0.82	0.15 ± 0.00	34 ± 2	6.65 ± 0.04	9.06 ± 0.02
1223	170712001601319	5848 ± 55	4.21 ± 0.15	-0.19 ± 0.07	10.28	0.80	0.18 ± 0.00	56 ± 2	6.41 ± 0.03	9.18 ± 0.02
1223	150703001601348	5791 ± 65	4.34 ± 0.17	+0.02 ± 0.08	10.64	0.96	0.16 ± 0.00	79 ± 3	7.16 ± 0.05	9.83 ± 0.02
4512	161009002601246	5770 ± 44	4.30 ± 0.12	+0.09 ± 0.05	12.18	0.82	0.20 ± 0.01	283 ± 56	5.64 ± 0.06	8.44 ± 0.01
4512	161009002601314	5599 ± 49	4.09 ± 0.13	+0.29 ± 0.06	11.92	0.89	0.18 ± 0.00	254 ± 20	5.99 ± 0.04	8.55 ± 0.01
3959	170531001901267	5899 ± 58	4.22 ± 0.16	+0.16 ± 0.07	11.02	0.85	0.19 ± 0.00	137 ± 1	6.23 ± 0.03	9.16 ± 0.02
3959	160524002101209	5814 ± 59	4.47 ± 0.16	-0.11 ± 0.07	11.45	0.89	0.21 ± 0.00	208 ± 2	5.60 ± 0.03	8.52 ± 0.02
3560	170513004901374	5293 ± 61	4.37 ± 0.16	-0.09 ± 0.07	11.12	1.00	0.21 ± 0.01	18 ± 4	5.60 ± 0.10	8.53 ± 0.01
3560	170615003901348	5138 ± 67	4.48 ± 0.17	+0.35 ± 0.08	11.31	1.09	0.24 ± 0.00	75 ± 2	5.69 ± 0.02	9.33 ± 0.02
271	160423002201186	6053 ± 43	4.23 ± 0.12	-0.04 ± 0.05	10.77	0.72	0.19 ± 0.00	73 ± 1	5.73 ± 0.02	8.41 ± 0.01
271	160522002101256	5581 ± 61	4.36 ± 0.16	+0.14 ± 0.07	11.64	0.93	0.13 ± 0.00	116 ± 3	7.39 ± 0.04	9.50 ± 0.03
3027	160513001101131	5798 ± 52	4.50 ± 0.14	+0.08 ± 0.06	11.25	0.85	0.09 ± 0.00	129 ± 1	7.23 ± 0.04	8.68 ± 0.01
3027	160513001101351	5358 ± 63	4.52 ± 0.17	+0.02 ± 0.08	11.84	1.00	0.14 ± 0.00	252 ± 1	5.96 ± 0.01	7.93 ± 0.00

### 3 CO-MOVING GROUPS IN GALAH

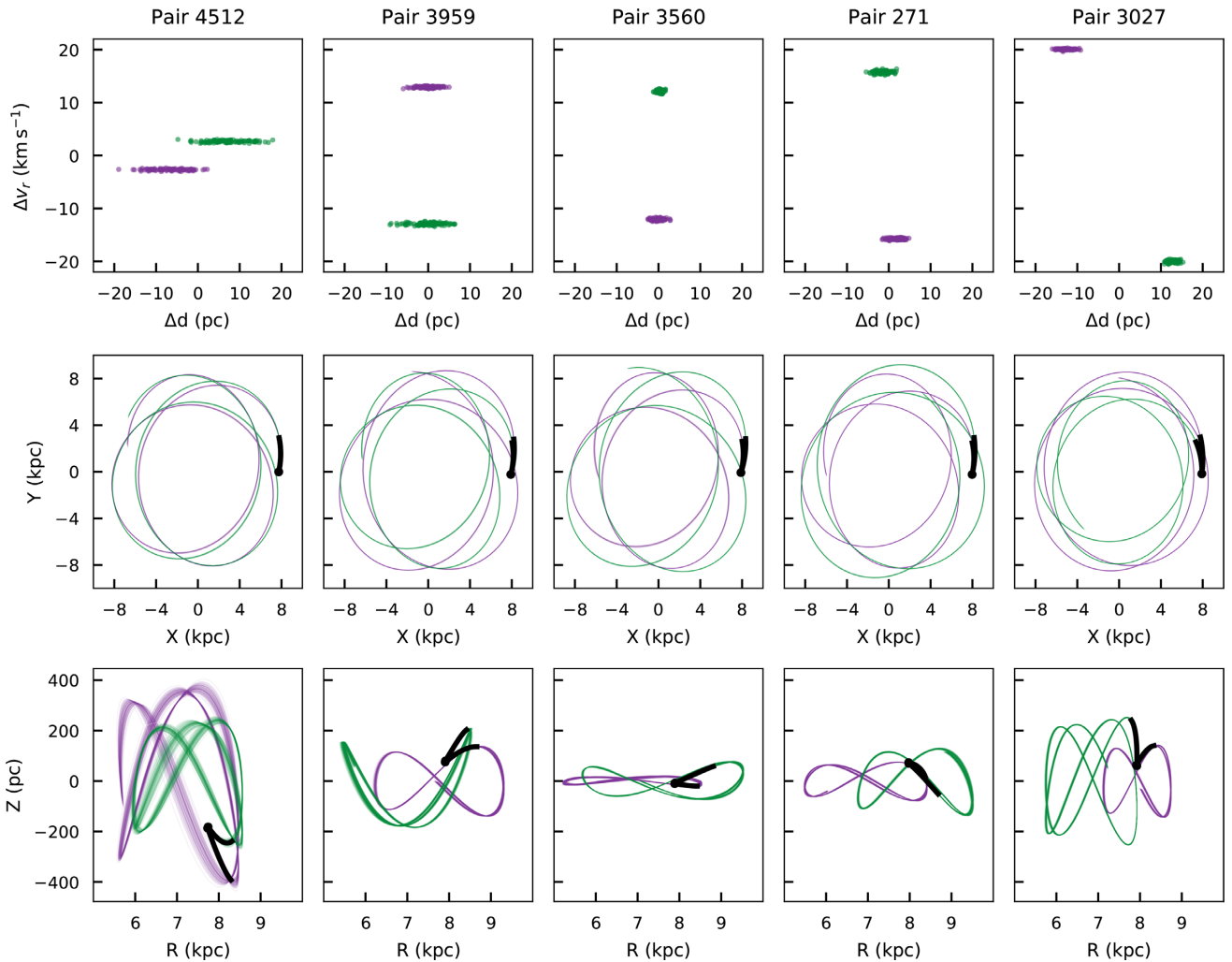
**O17** identified 10 606 stars to be in non-exclusive co-moving pairs or groups.<sup>3</sup> Unfortunately, only 117 of these stars are found in the GALAH catalogue, and only 15 pairs had both stars observed (i.e. 30 stars). Table 1 lists the **O17** group ID, the GALAH subject\_id, stellar parameters, photometry, and orbital parameters for the 15 pairs. This very small overlap is the result of two selection effects within GALAH. First, the majority of stars in GALAH are found in the magnitude range  $12 < V < 14$ , with a smaller number of stars up to  $V = 9$ , while TGAS (used by **O17**) is predominantly  $G < 11$ . This means that most of the **O17** stars are brighter than GALAH's magnitude range. Second, GALAH only observes stars at declinations  $-80^\circ < \delta < +10^\circ$  and Galactic latitudes  $10^\circ < |b| \lesssim 50^\circ$ .

Our first step in evaluating whether the possible **O17** pairs are truly co-moving was to integrate their orbits around the Galaxy. For each star the covariance matrix was constructed from the reported errors and covariances in *Gaia* DR2, and then 100 samples were drawn using `numpy.random.multivariate_normal` to give the RA, Dec., inverse parallax, proper motions in RA and Dec., and radial velocity ( $\alpha$ ,  $\delta$ ,  $r_\odot$ ,  $\mu_\alpha \cos \delta$ ,  $\mu_\delta$ ,  $v_r$ ). It is important to consider the uncertainties of these values as it is

not intuitive how a large uncertainty in one parameter will impact the orbit, especially as we are taking the projected velocities on the sky. An orbit was computed for each sample using GALPY (version 1.4; <http://github.com/jobovy/galpy>; Bovy 2015) with the recommended Milky-Way-like `MWPotential2014` potential and the solar motion defined by Schönrich, Binney & Dehnen (2010). The orbits were integrated forward in time for 500 Myr with 0.5 Myr resolution. Note that these orbital integrations do not take into account the mutual gravitational attraction of the pairs.

We show projections of 100 orbit samples for each star of each pair in Figs 1, 2, and 3. Each star in each pair is plotted in a different colour. For each star we show the radial velocities of the stars in the pair relative to the mean versus the distances of the stars in the group relative to the mean; the integrated orbits projected onto the Galactic  $X$ - $Y$  plane; and the orbits in the  $R$ - $Z$  plane. In the orbit panels, the current position of each star is shown with a black dot, and the first 15 Myr of its orbit is shown with a black line, to indicate the direction of motion. In Table 1 we give the eccentricity, maximum vertical height, perigalacticon, and apogalacticon for these orbits. In most cases the uncertainties and covariances of the input parameters are fairly small. The major component of the uncertainties of the orbital parameters is the error in the parallax ( $\sim 50$  percent of the uncertainty). Pair 4512 has a large uncertainty in the stellar distances, which causes a range of possible future orbits. It is important to consider the uncertainties

<sup>3</sup>Group numbers referred to in this work are the Group column of **O17**.



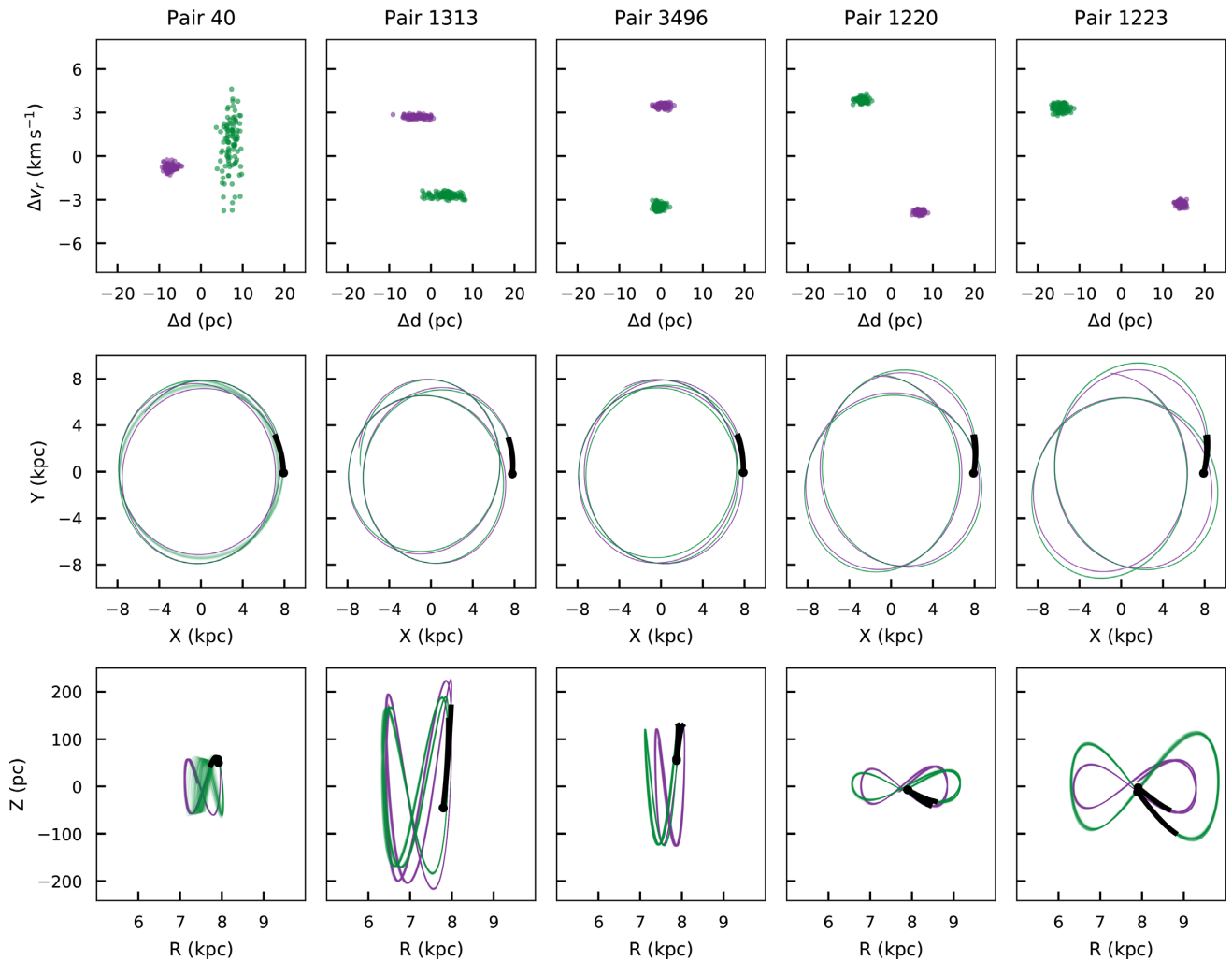
**Figure 1.** Projections of the orbits of the five pairs with the largest radial velocity difference. For each star, 100 versions of its orbit are shown where the input parameters were drawn from a multivariate normal distribution that took into account the covariances between the parameters and their uncertainties. The top row shows the relative radial velocities of and separations between the stars in each [O17](#) pair; the second row is the projections of their Galactic orbits on the  $X$ - $Y$ ; and the third row is their  $R$ - $Z$  planes. On the orbit plots, a black dot indicates the current median position of the stars and the black lines show the median direction of motion. In all cases, the orbit integrations are consistent with none of these pairs being co-moving about the Galaxy.

and covariances of the input parameters when calculating the errors in the orbital parameters, as it is not necessarily intuitive how a large uncertainty in a given parameter will impact the orbit.

The thin disc is dynamically fairly cold; that is, the velocity dispersion in the  $(U, V, W)$  velocity space is not very large, and the orbits of individual stars tend to have low eccentricity and be confined to the plane of the disc (e.g. Edvardsson et al. 1993; Pasetto et al. 2012). This introduces the possibility that stars might be close to each other and co-orbiting without having formed together. To better understand how random but spatially close pairs of stars are distributed in the  $(U, V, W)$  space, we carried out a simple experiment within the GALAH data set. We identified 5755 pairs of stars in GALAH within 600 pc of the Sun, for which each star has only one other star within 10 pc (i.e. not in clusters or large associations), and calculated  $\Delta(U, V, W)$ , the Cartesian distance between the velocities of the two stars. Fig. 4 shows the distribution of  $\Delta(U, V, W)$  with the values for the co-orbiting stars overlaid. In Table 2 we give the spatial and observed kinematic differences between each of the co-orbiting pairs.

For the random pairs, the peak of the  $\Delta(U, V, W)$  is about  $40 \text{ km s}^{-1}$ . The pairs of stars we consider from [O17](#) tend to have lower  $\Delta(U, V, W)$  than the bulk of the random pairs. This is not surprising; our selection of the random pairs only required them to be spatially close, while the [O17](#) pairs were both spatially close and similar in proper motion. Four of the 15 pairs have large  $\Delta(U, V, W)$ , at values like the peak of the distribution of  $\Delta(U, V, W)$  for the random pairs. These four pairs (271, 3027, 3560, and 3959, shown in Fig. 1) have the largest radial velocity differences, with  $|\Delta RV| > 24 \text{ km s}^{-1}$ . Their integrated orbits show that within these pairs, the stars have widely different inclinations or eccentricities. We, therefore, conclude that the apparent association in proper motion of the stars in these four groups is coincidental.

For the other 11 pairs (Figs 1, 2, and 3), we find that some do appear very likely to be co-moving. The difference in radial velocity for all of these pairs is  $< 7.7 \text{ km s}^{-1}$ , and in four cases it is  $< 1.0 \text{ km s}^{-1}$ . These four pairs (Fig. 3) we consider to be the most likely to be co-moving. The orbits of the stars in the pairs with the largest relative radial velocity do diverge in the  $R$ -



**Figure 2.** Same as Fig. 1 but for five pairs of stars with smaller radial velocity differences. One star of Pair 40 has a large uncertainty in its radial velocity. This manifests in the orbit integrations as a larger uncertainty in its perigalacticon than for most other stars considered (Table 1).

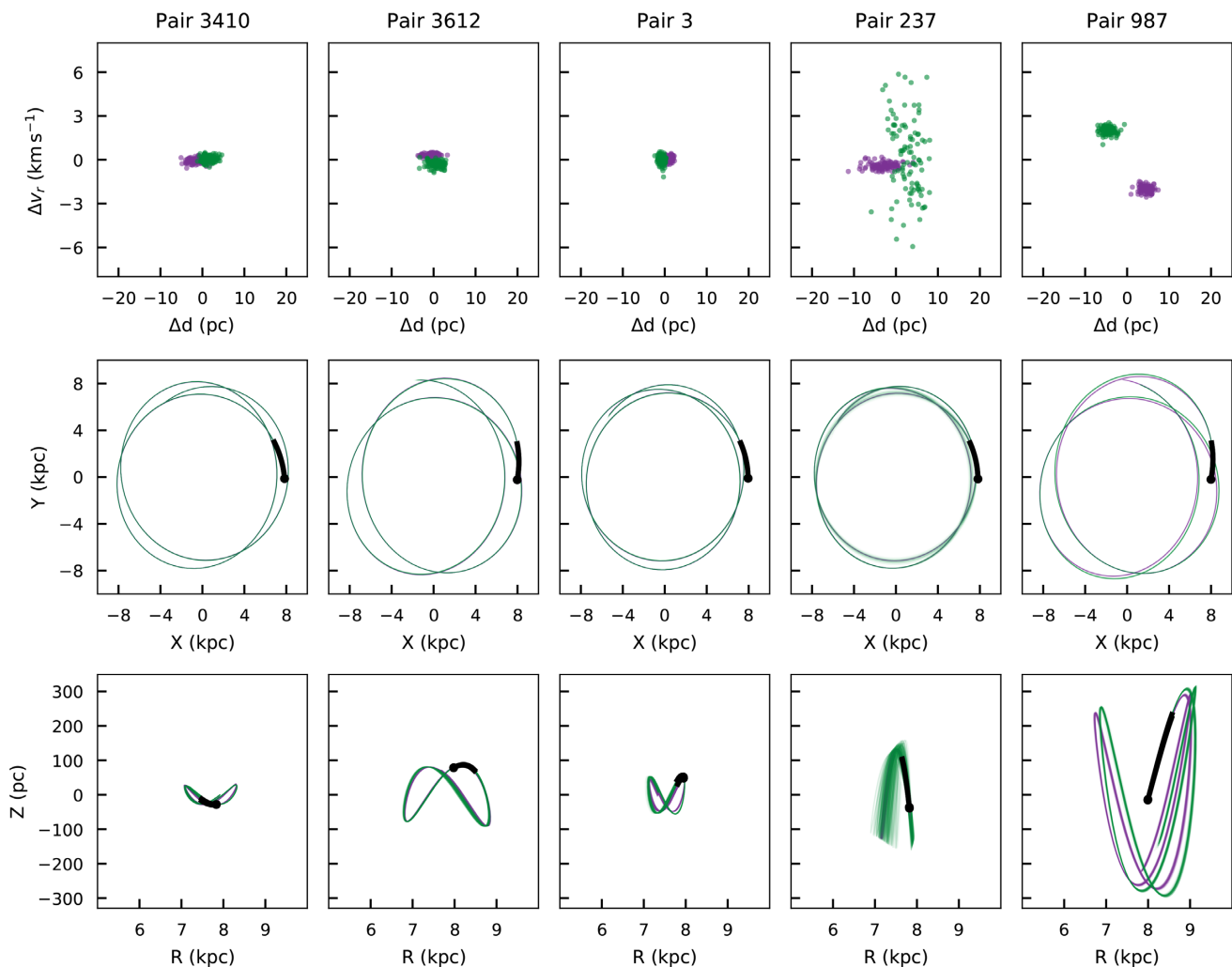
Z plane (Fig. 2), especially in pairs 4512 and 1223, so although they have similar orbits now they may not remain associated in the future. Andrews et al. (2018b) predict, and Andrews, Chanamé & Agüeros (2018a) demonstrate, that the majority of apparent co-moving pairs in TGAS with separations larger than 0.2 pc are not truly co-orbiting. All of the potentially co-moving pairs in this study have separations larger than that limit. While we find that a minority of the 15 pairs we consider are not co-orbiting, this is a small enough data set that we cannot strongly support or contradict the 0.2 pc limit described by Andrews et al. (2018b). In all cases, more detailed orbit integrations (i.e. more realistic potential, considering the gravitational interaction between the stars) would provide a more conclusive answer to whether these stars will continue to orbit together. While we are considering only a small subset of the O17 sample, it is clear that full 6D velocity confirmation is necessary before drawing larger conclusions about cluster dissolution or disc substructure from the reported co-moving pairs.

### 3.1 Similarity in stellar parameters, photometry, and spectra

In several of the co-orbiting pairs, the stars have similar magnitudes and colours (which are given in Table 1). Since they were selected to have similar distances from the Sun, we expect that their stellar

parameters should also be close. Fig. 5 shows the (absolute  $G$ ,  $G_{BP} - G_{RP}$ ) colour–magnitude diagram for all GALAH stars within 600 pc of the Sun. The low  $\Delta(U, V, W)$  pairs are highlighted as larger orange dots and the four kinematically dissimilar pairs as larger blue dots, and each pair is connected with a line. Unsurprisingly for apparently bright stars located relatively nearby, the stars in the pairs tend to be on the main sequence. There is one pair that contains a potential subgiant (Pair 40). Rather than stellar twins, almost all of the pairs consist of a brighter, hotter star with a fainter, cooler ‘companion’. This is consistent with the pairs of stars being of the same age and differing slightly in mass, but it is not conclusive, since (i) the majority of stars within 600 pc of the Sun are main-sequence stars, so randomly assigned pairs would tend to behave in this same way, and (ii) the kinematically dissimilar pairs show similar arrangements.

There are three pairs for which we can evaluate ages slightly more precisely: groups 237, 3, and 40. Group 237 is the most luminous main-sequence pair in our data set, and given the shorter main-sequence lifetimes for higher-mass stars, the stars in this group must be closer in age than the other pairs. Comparison with a solar metallicity MIST isochrone (Paxton et al. 2011, 2013, 2015; Choi et al. 2016; Dotter 2016) returns a maximum age for the brighter star in group 237 of about 800 Myr.



**Figure 3.** Same as Fig. 1 but for five pairs of stars with the smallest radial velocity differences. One star of Pair 237 has a large radial velocity (RV) uncertainty (as for one star of Pair 40; Fig. 2). For Pairs 3410, 3612, 3, and 237, these orbit integrations provide good evidence that they are truly co-orbiting the Galaxy.

Meanwhile, the stars from group 3 are distinctly redder than most of the other pairs for their luminosity. These stars are members of the Sco-Cen OB association and their main sequence is brighter at redder colours than that of the bulk of the stars in our sample because of their young age. Isochrone fits yield an age of about 10 Myr for this pair, which is consistent with age estimates for the association (e.g. Pecaat, Mamajek & Bubar 2012). It is likely that they are pre-main-sequence stars, so it is unsurprising that THE CANNON, which was not trained for that evolutionary stage, has not returned reliable parameters for either star.

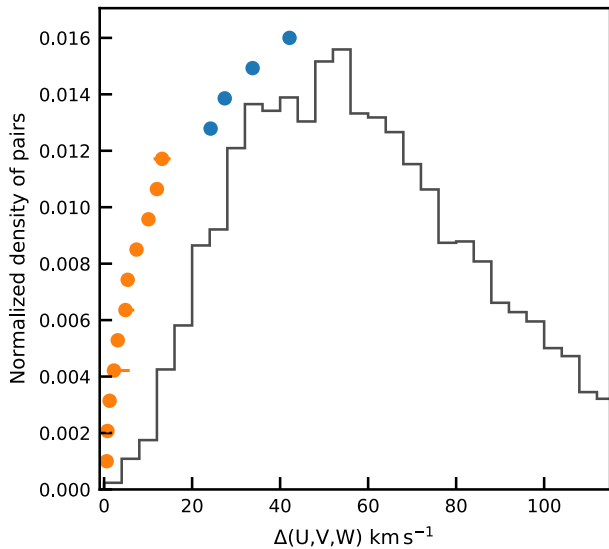
Group 40 has one main-sequence member and one apparent subgiant member, but it was not possible to fit a single isochrone to both stars. The subgiant star is faint enough that it must have a lower mass and higher age than the main-sequence star, even though they are likely co-orbiting. Unfortunately, we only have reliable label flags for the main-sequence star, so we cannot confirm this with stellar parameters and abundances.

In some cases, the stars are similar enough that a direct comparison of the spectra is a sufficient demonstration of highly similar abundance patterns (similar to what was shown in Bovy 2016). The stellar parameters derived for the two stars in group 3612 are practically identical. Their full spectra are shown in Fig. 6 and small

cut-outs of spectral regions used for the abundance determination in Fig. 7. The spectra of the two stars are remarkably alike, which, together with the matching parameters, indicates that the two stars must have quite similar abundance patterns. Deriving stellar parameters and abundances for two stars that are clearly so similar in their observational properties is a good basic verification of the GALAH analysis process. This is explored in the next section.

#### 4 ABUNDANCE BEHAVIOUR IN CO-ORBITING PAIRS

The chemical homogeneity of stars that form in the same environment at the same time is an axiom of star formation; considering gas clouds that will collapse to form stars, Feng & Krumholz (2014) showed that turbulent mixing is highly effective in homogenizing the composition of higher-mass clouds. Lower-mass clouds and star-forming clouds in regions with a lower star formation efficiency may potentially be less well mixed, but they are not as well studied. However, when abundances of stars in open clusters are measured at very high precision, there are clear inhomogeneities. Using high-precision differential abundances measured from very high-quality spectra ( $R = 50\,000$ , signal-to-noise ratio, SNR > 270 per pixel) of stars in the Hyades, Liu et al. (2016) found that there are star-to-star



**Figure 4.** The distribution of  $\Delta(U, V, W)$  of the 5755 random GALAH-observed pairs (black histogram) compared to that of the GALAH DR2 pairs in common with O17 (orange and blue dots with errorbars; values from Table 2). The O17 pairs have been vertically offset for visual clarity. The background distribution uses pairs that were selected solely to be spatially close ( $d < 10$  pc), while the O17 pairs are spatially close and have similar proper motions. We find that 4 of the 15 pairs have  $\Delta(U, V, W)$  at the peak of the background distribution. The other 11 all have  $\Delta(U, V, W) < 14$  km s $^{-1}$ , which was only the case for 2 per cent (111/5755) of the random pairs.

abundance variations on the order of 0.02 dex and that most of the elemental abundances in each star are correlated to each other.

We expect, therefore, to see a baseline of homogeneity for stars that formed together, to within the precision of the GALAH abundances ( $\approx 0.1$  dex), and we do not expect this for stars co-orbiting due to dynamical effects (like stars in the Hercules stream, which are co-orbiting as a result of resonance with the Galactic bar; e.g. Fux 2001; Famaey et al. 2005). However, it is important to remember that there are reasons that co-natal, co-eval stars might have mismatched abundances, e.g. atomic diffusion (e.g. Dotter et al. 2017), planet formation (e.g. Meléndez et al. 2009), and binary mass transfer (e.g. Hansen et al. 2015), all of which affect a particular set of elemental abundances during particular evolutionary phases. As one example of mismatched but potentially co-natal stars, Oh et al. (2018) used archival Keck/HIRES spectra for pair 1199 from O17 and found that the two stars have very similar 3D velocities and are consistent with having similar ages. However, they found that one star is enhanced by 0.2 dex in refractory elements and by 0.05–0.10 dex in volatile elements relative to the other star. They interpret this pair of stars as having formed together, with the relatively enhanced star having later accreted rocky material, presumably following the formation of a planetary system.

In 12 of the 15 pairs, both stars have `flag_cannon` = 0, meaning that their stellar labels have no evidence for being untrustworthy (unfortunately the missing three are some of the most kinematically similar pairs). We plot their metallicity,  $[\alpha/\text{Fe}]$ ,<sup>4</sup> and elemental abundances in Figs 8 and 9, with one panel per element and the same horizontal scale in all panels. Fig. 8 plots the  $[\text{X}/\text{Fe}]$  values for

<sup>4</sup>In GALAH,  $[\alpha/\text{Fe}]$  is the error-weighted combination of Mg, Si, Ca, and Ti abundances. See Buder et al. (2018).

each star of each pair, while Fig. 9 plots the abundance differences  $-\Delta([\text{X}/\text{Fe}])$  of each pair. Large circles are used for small  $\Delta(U, V, W)$  pairs and small symbols for the non-co-moving pairs. The colour of each pair is merely to aid the eye and it has the same vertical arrangement in each panel, sorted by  $\Delta(U, V, W)$ . The group numbers from O17 are listed to the right of each group in the panel showing  $[\text{Fe}/\text{H}]$ . In some cases, individual element abundances in one or both stars of a pair were flagged as unreliable, so some pairs are missing from some panels.

The elements are arranged by their dominant nucleosynthetic groups, represented by the colour of the background histogram, with the group name (‘p’ for proton capture, ‘odd Z’ for light odd-Z elements, ‘ $\alpha$ ’ for alpha elements, ‘s’ for the slow neutron capture process, ‘r’ for the rapid neutron capture process, and ‘Fe peak’ for the iron-peak elements) next to the name of the element. In Fig. 8 the distribution of each element’s abundance for all GALAH targets within about 600 pc of the Sun is shown as a smoothed histogram in the background of each panel. In Fig. 9 we show the distribution of differences for the 5755 random, spatially close pairs drawn from GALAH within 600 pc of the Sun (as described in Section 3). The number of stars or pairs in the background histogram is given in the upper left of each panel, since not all elements are measured in all stars.

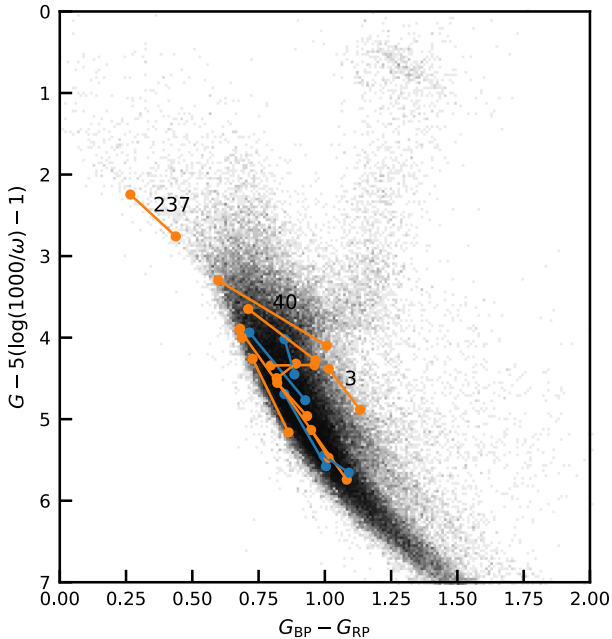
In Fig. 8, the background distributions of  $\alpha$  and iron-peak abundances are quite narrow, indicating that Type Ia supernovae played a strong role in enriching the gas that formed these stars. This is not a surprise since the stars are all within about 600 pc of the Sun, and (based on our orbital integrations) typically associated with the thin disc. The slightly broader ranges in the proton capture and s-process elements suggests that feedback from low- and intermediate-mass stars also contributed to the chemical evolution of the material from which the stars formed. This is also consistent with these stars being relatively young thin-disc stars since their late formation allows time for (as an example) asymptotic giant branch stars to produce and eject significant amounts of CNO-cycle and s-process elements. These broader distributions may also reflect their larger abundance uncertainties, as there are typically fewer spectral lines being used for these elements (Buder et al. 2018).

Looking at our pairs of stars in the  $[\text{Fe}/\text{H}]$  panel of Fig. 8, we find that they are evenly distributed across the background distribution. In most cases, for both the kinematically similar and dissimilar pairs, there are large metallicity differences between the stars in the pair: Typically  $\Delta[\text{Fe}/\text{H}] > 0.15$  (Fig. 9). This is a much larger difference than would ordinarily be explained by inhomogeneities in the pre-stellar nebula, planet formation (e.g. Meléndez et al. 2017), or differential effects of diffusion on main-sequence stars with different masses (e.g. Michaud et al. 2004). The kinds of dynamical processes that could result in non-co-natal stars being gravitationally bound to each other, such as stellar captures or partner swapping in binaries/triple systems, are unlikely to explain the pairs with very similar orbits but differing metallicity. Given their large separation, the co-moving pairs could only be weakly gravitationally bound, if at all. The similarity in their orbits is likely due to the constrained range of orbital parameters available in the thin disc.

Pairs 3410, 3612, and 987 are the only kinematically similar pairs that have  $\Delta[\text{Fe}/\text{H}] < 0.11$ . Looking at the other abundance-difference panels of Fig. 9, these are the only pairs that are consistently similar within the uncertainties. In general, the abundance difference between the stars in each pair is smaller than the spread of the background distribution, reinforcing the claim from the orbital properties that the stars in each pair are intrinsically associated. Considering Pair 3612, which we highlighted in Fig. 6, the results

**Table 2.** Differences in position and velocities of 15 pairs of stars found in GALAH that were identified in O17. The values are the median, and the 5th and 95th percentiles of 10000 random samples of the 6D information of the stars taking account their uncertainties and covariances in *Gaia* DR2. The pairs are sorted by the  $\Delta(U, V, W)$  difference of the pair. The pairs above the line have lower  $\Delta(U, V, W)$  and are potentially co-orbiting pairs, while those below the line are likely not. All but one have a physical separation greater than 3.1 pc.

Group ID	Ang sep (arcmin)	Separation (pc)	$\Delta(v_r)$ (km s <sup>-1</sup> )	$\Delta(U)$ (km s <sup>-1</sup> )	$\Delta(V)$ (km s <sup>-1</sup> )	$\Delta(W)$ (km s <sup>-1</sup> )	$\Delta(U, V, W)$
3410	1	3.1 <sup>+3.3</sup> <sub>-2.7</sub>	0.1 ± 0.4	-0.1 ± 0.4	-0.2 ± 0.3	-0.5 ± 0.2	0.6 ± 0.2
3612	3	1.5 <sup>+2.8</sup> <sub>-1.2</sub>	0.6 ± 0.4	-0.1 ± 0.8	0.5 ± 0.4	-0.3 ± 0.2	0.8 ± 0.5
3	110	4.2 <sup>+0.8</sup> <sub>-0.4</sub>	0.1 ± 0.6	-0.9 ± 0.3	0.1 ± 0.5	-0.8 ± 0.2	1.3 ± 0.4
237	88	8.6 <sup>+5.0</sup> <sub>-2.1</sub>	0.8 ± 4.9	-0.3 ± 3.6	-0.8 ± 3.1	-0.4 ± 1.2	2.2 ± 3.6
987	51	9.6 <sup>+2.8</sup> <sub>-2.7</sub>	3.9 ± 0.5	0.1 ± 0.9	-2.9 ± 0.5	-0.9 ± 0.2	3.1 ± 0.5
40	185	16.4 <sup>+2.2</sup> <sub>-2.1</sub>	1.7 ± 2.8	-1.0 ± 1.8	-4.7 ± 2.1	-0.7 ± 0.6	4.9 ± 2.0
1313	44	7.9 <sup>+4.5</sup> <sub>-3.7</sub>	5.4 ± 0.3	-4.4 ± 0.6	1.9 ± 0.6	2.3 ± 0.2	5.4 ± 0.4
3496	69	3.2 <sup>+1.1</sup> <sub>-0.2</sub>	6.9 ± 0.3	-5.3 ± 0.3	5.2 ± 0.4	0.0 ± 0.1	7.4 ± 0.3
1220	112	14.8 <sup>+1.8</sup> <sub>-1.7</sub>	7.7 ± 0.3	10.0 ± 0.6	0.4 ± 0.4	-0.7 ± 0.2	10.1 ± 0.6
1223	111	29.2 <sup>+2.1</sup> <sub>-2.0</sub>	6.6 ± 0.4	10.9 ± 0.8	-3.5 ± 0.6	3.7 ± 0.3	12.0 ± 0.7
4512	46	13.8 <sup>+10.6</sup> <sub>-8.6</sub>	5.3 ± 0.4	-1.4 ± 1.2	-6.9 ± 1.1	-11.1 ± 1.5	13.1 ± 2.0
3959	101	7.8 <sup>+2.3</sup> <sub>-0.4</sub>	25.8 ± 0.4	-11.0 ± 1.7	20.9 ± 0.8	-5.1 ± 0.2	24.2 ± 0.5
3560	172	6.9 <sup>+0.7</sup> <sub>-0.1</sub>	24.2 ± 0.4	21.1 ± 0.8	-16.3 ± 1.1	-6.3 ± 0.3	27.4 ± 0.5
271	86	6.9 <sup>+2.3</sup> <sub>-1.0</sub>	31.5 ± 0.4	11.1 ± 0.9	-31.4 ± 0.5	5.6 ± 0.3	33.8 ± 0.5
3027	48	25.8 <sup>+3.1</sup> <sub>-3.0</sub>	40.1 ± 0.4	-28.2 ± 0.6	30.1 ± 0.5	-8.4 ± 0.1	42.1 ± 0.4



**Figure 5.** Colour–magnitude diagram of all stars observed by GALAH within about 600 pc of the Sun. Also plotted are the 15 stellar pairs from O17 that were serendipitously observed as part of the GALAH survey. Those pairs that GALAH data show are kinematically similar are shown in orange, and those that are dissimilar are shown in blue. Each pair is connected by a line. The three numbered pairs are discussed in Section 3.1.

shown in Fig. 9 find that only in  $[Y/Fe]$  does the difference fall outside of the  $\pm 0.15$  dex range.

Pair 3027 has the largest velocity difference, but its stars have similar  $[Fe/H]$ . This is likely a coincidence as both stars in this pair sit at the most common value of the underlying metallicity distribution – slightly supersolar. It has abundance differences for several of the elements determined by GALAH. All of the other kinematically dissimilar pairs have large metallicity differences, confirming the results from the radial velocities that these are neither co-orbiting nor co-natal pairs.

We warn the reader against overinterpreting these abundance data. The background distributions of the alpha and iron-peak elements are narrow enough that it can be difficult for those abundances to be significantly different, even in randomly chosen pairs. They are not that informative for chemical tagging to distinguish between different co-orbiting pairs, then, but since the background distributions are so narrow, they would be very effective for distinguishing the thin disc from any thick-disc (e.g. Bensby, Feltzing & Oey 2014) or halo (e.g. Roederer et al. 2010) stars in the data set.

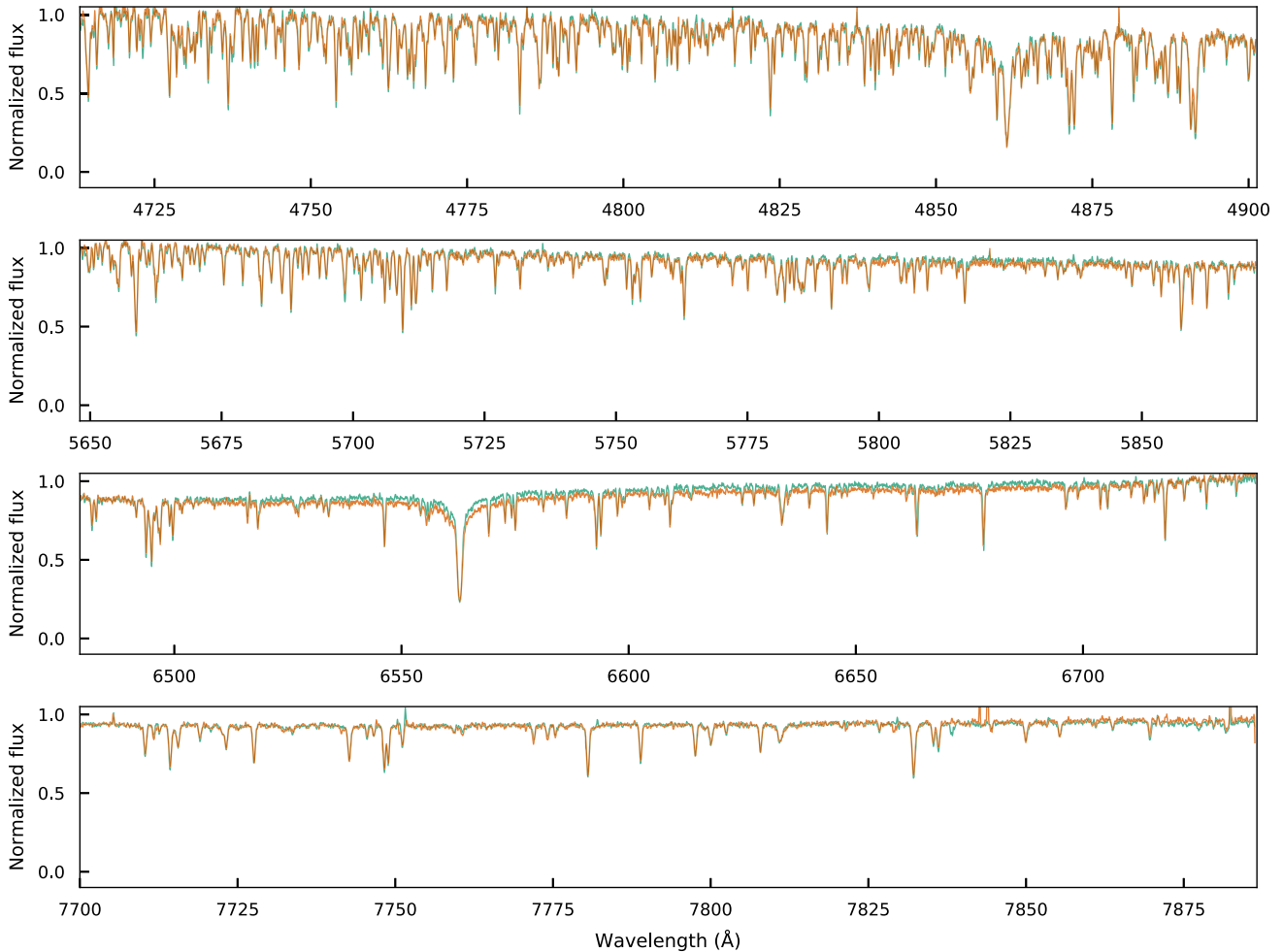
## 5 DISCUSSION

In this work, we have presented a fundamental step in the development of detailed chemical tagging and explored our ability to distinguish true co-orbiting stars from those that are coincidentally kinematically close using spectroscopic stellar parameters and abundances. We reiterate that the main driver of the GALAH survey is chemical tagging. As discussed in Section 1, we aim to identify co-eval and co-natal groups of stars first by their similarity in abundance space and use stellar parameters, ages, and kinematics as a confirmation of their association. In this work, in contrast, stars were initially selected by their phase-space information, and then the GALAH spectra and abundances were used to confirm (or refute) a similar common origin for the stars.

The small size of the data set considered in this study allows the star-by-star analysis we have chosen to use, but full chemical tagging in the Galactic disc, like many other goals of Galactic archaeology, will require a fairly high level of automation based on well-justified metrics and statistics. There have been several methods already proposed and tested, including the Manhattan distance metric (Mitschang et al. 2013), t-SNE dimensionality reduction (Anders et al. 2018; Kos et al. 2018), principal component analysis (Blanco-Cuaresma et al. 2015), k-means clustering (Hogg et al. 2016), deriving the chemical dimensionality from the spectra in a semi-model-independent method (Price-Jones & Bovy 2018), and unsupervised clustering (including a minimum spanning tree; Boesso & Rocha-Pinto 2018).

The strengths of these various methods will make them more or less suited for particular chemical tagging problems. For example,





**Figure 6.** Spectra of the two stars of Pair 3612 in the four HERMES bandpasses. The spectra are so similar that it is difficult to distinguish them. In Fig. 7, we show small cut-outs of some of the spectral lines used in the abundance determination.

k-means and extreme deconvolution (Bovy, Hogg & Roweis 2010) will assign all stars in a data set to a given number of groups and are therefore ideal for disentangling multiple known populations mixed together, while t-SNE and DBSCAN clustering (Traven et al. 2017) will identify all groups above a certain density threshold. Principal Component Analysis (PCA) is very effective at identifying baseline trends in abundance space but can be derailed by outliers.

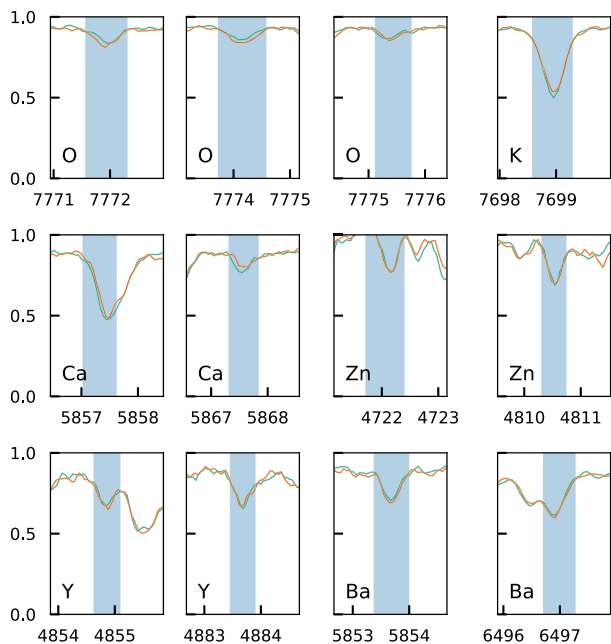
The challenges in applying these methods come both from physics and from data analysis. Galactic chemical evolution is extremely complex, with many sources of enrichment contributing differently, but that does not necessarily guarantee that each star formation site over its history has had a unique abundance pattern (e.g. Ness et al. 2018). A low level of intrinsic abundance scatter has been seen in Galactic open clusters (Liu et al. 2016), and the level of that scatter is theoretically expected to be a function of cluster mass (Feng & Krumholz 2014). Furthermore, even if stars begin with perfectly identical abundance patterns, atomic diffusion will deplete some of those elements in main-sequence stars, with larger effects in higher-mass stars (e.g. Gao et al. 2018). The first dredge-up then restores the abundances to their pre-diffusion values, creating a variable abundance offset between co-natal dwarfs and giants.

Spectroscopic analysis introduces a range of uncertainties to chemical tagging. Each elemental abundance value has an error bar

that depends on how precisely the observed absorption features can be fit by synthetic spectra, which is affected by the SNR and dispersion of the spectrum. The model atmospheres from which synthetic spectra are calculated are not perfect captures of the real physical properties of stellar atmospheres, with non-local thermodynamic equilibrium and 3D atmospheric effects making significant differences in the calculated abundances of some elements (e.g. Lind et al. 2017). There can also be systematic differences in the abundances determined for dwarf and giant stars (e.g. Korn 2010).

A complete chemical tagging method will need to account for these factors, interpreting the abundance data using stellar evolutionary models to account for factors like diffusion and abundance evolution, building on a probabilistic model of Galactic chemical evolution and the intrinsic abundance scatter among co-natal stars, and accounting appropriately for systematic and random uncertainties in the measured abundances. Future GALAH data releases will use the Bayesian isochrone matching code ELLI (Lin et al. 2018), which returns probability distributions for stellar age and mass, along with an estimate of the initial composition. Using the initial composition should alleviate some of the difficulties in chemical tagging caused by the evolution of surface abundances.

This method will also need to consider some fairly fundamental questions: Which elements need to be considered? How should their importance be weighted? Should multiple elements from each



**Figure 7.** Assorted cut-outs from Fig. 6 of some of the spectral regions used for the abundance determination, highlighting how very similar the spectra of Pair 3612 are. The blue shaded region in each panel is the mask used by THE CANNON for that line.

nucleosynthetic group be counted together or separately? Should we attempt to find similar stars in an  $[X/H]$  chemical space or an  $[X/Fe]$  chemical space? Just as different chemical tagging problems are better addressed by extreme deconvolution or unsupervised clustering or principal component analysis, they may also be more effectively answered with different baseline choices about which measurements are the most informative.

It is interesting to consider the dynamical history of co-natal stars that are spatially close and on very similar orbits at the present day. Studies of young stellar associations (e.g. Wright et al. 2016) often find diverging velocities, indicating that they will not remain gravitationally bound. The age distribution of open clusters (Friel 1995) is highly skewed towards young clusters, indicating that the typical dissolution time for star formation events with masses around  $10^3 M_{\odot}$  is less than 500 Myr. Lower-mass star formation sites will have a shallower gravitational potential, lowering the escape velocity and leading to a shorter dissolution time. Even if stars that formed together have small relative velocities, scattering interactions with giant molecular clouds or other stars, radial migration, and resonances can all dramatically alter their orbits.

Each star’s probability of undergoing some kind of interaction increases with time, such that pairs or groups of co-natal stars that start out on similar orbits will be consistently disrupted over time as one or more stars undergo some kind of scattering or interaction (e.g. Goodwin & Kroupa 2005; Dhital et al. 2010). Without a dramatic scattering event, but with even a slight difference in  $L_z$ , unbound but co-orbiting stars will drift apart slowly (Jiang & Tremaine 2010). As a result, we expect that truly co-natal co-moving stars will tend to have completed fewer orbits of the Galaxy than pairs that are coincidentally co-moving or randomly selected spatially close pairs. We can also infer that co-natal, co-moving stars are likely to still be near their initial orbits since interactions that would change the orbit of the pair would potentially also disrupt the coherence of the individual stars’ orbits.

The study of Price-Whelan et al. (2017) describes follow-up spectroscopy for 311 potential co-moving pairs from O17 and finds that 40 per cent of those pairs have highly similar 3D velocities (albeit with radial velocity errors of the order of  $5 \text{ km s}^{-1}$ ). For the 15 pairs of stars considered in this data set, we find that 60 per cent have similar 3D velocities (though with a much smaller sample of stars). However, it is not entirely clear from our orbit integrations that they are truly co-orbiting, or will continue to be so in the future. For example, the stars in group 1313, which have a  $\Delta(v_r)$  of  $5.4 \pm 0.3 \text{ km s}^{-1}$ , have essentially the same eccentricity and orbital period, but one star is slightly more bound, with a lower total energy and angular momentum and a  $Z_{\text{max}}$  that is 15 per cent lower than the other star.

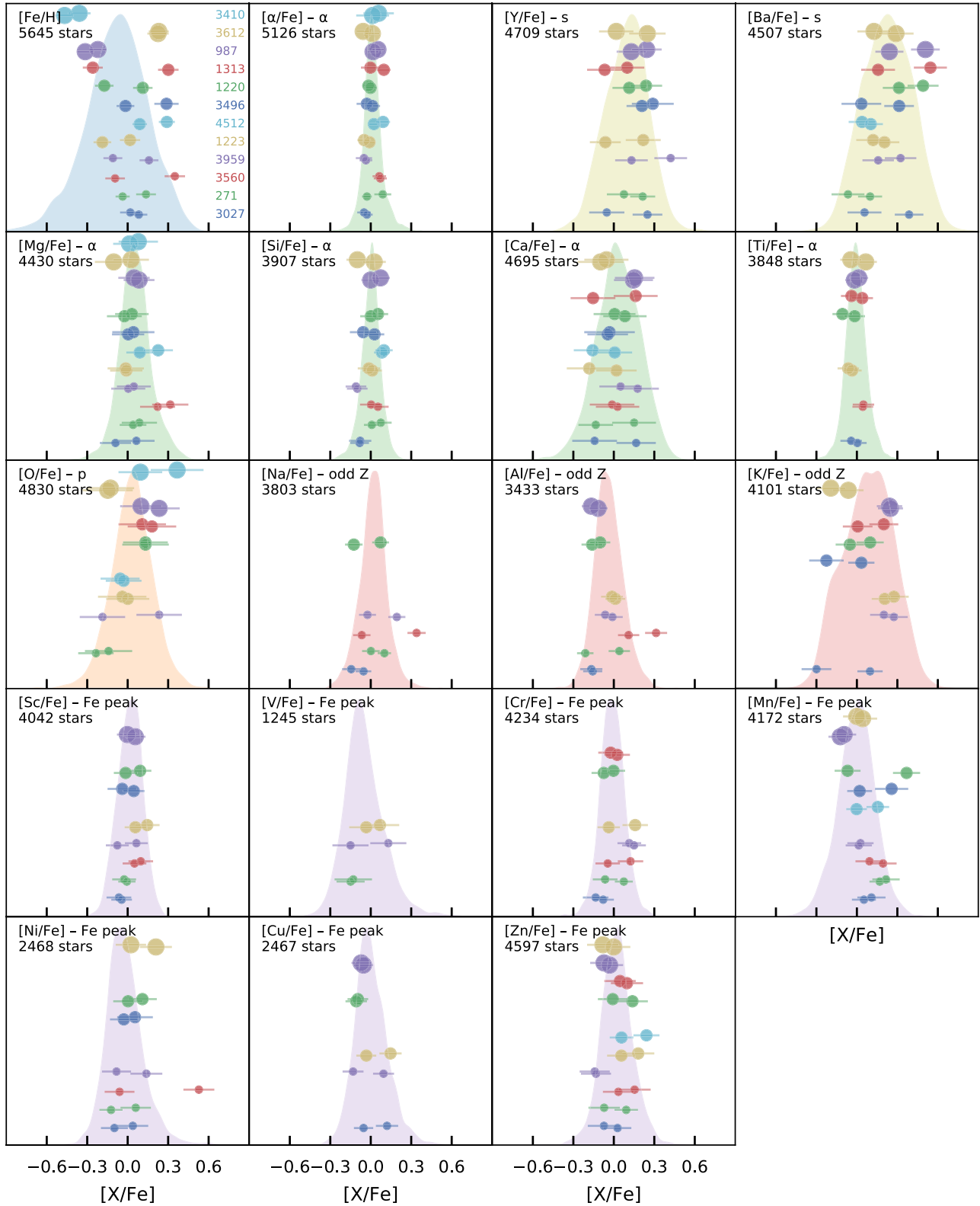
There is only one pair of stars in common between our work and Price-Whelan et al. (2017): O17 group 3. It is part of the Sco-Cen young stellar association. GALAH DR2 radial velocities are very similar to the RAVE radial velocities adopted for these stars by Price-Whelan et al. (2017), and since the stars are part of a known association, we agree with their kinematics-based assessment. However, considering stellar parameters and abundances, overall we tend to be more skeptical than Price-Whelan et al. (2017) about the likelihood that the pairs of stars are truly co-natal. Pair 40, for example, cannot originate in the same star formation event because one star must be distinctly older than the other based on their colour–magnitude positions.

Of the eight co-orbiting pairs for which we have stellar parameters and abundances, three have  $[Fe/H]$  differences of less than  $1.5\sigma$ , but the other five have  $[Fe/H]$  differences of up to  $8\sigma$ , or  $0.56 \text{ dex}$ , and their other abundance differences (as seen in Fig. 9) can also be quite large. The abundance differences in these five pairs do not correlate with the elements’ condensation temperature in the way that has been interpreted (by e.g. Meléndez et al. 2009) as a sign of rocky planet formation. For these five pairs, although their spatial locations and kinematics are presently quite similar, we conclude that they are not likely to be co-natal.

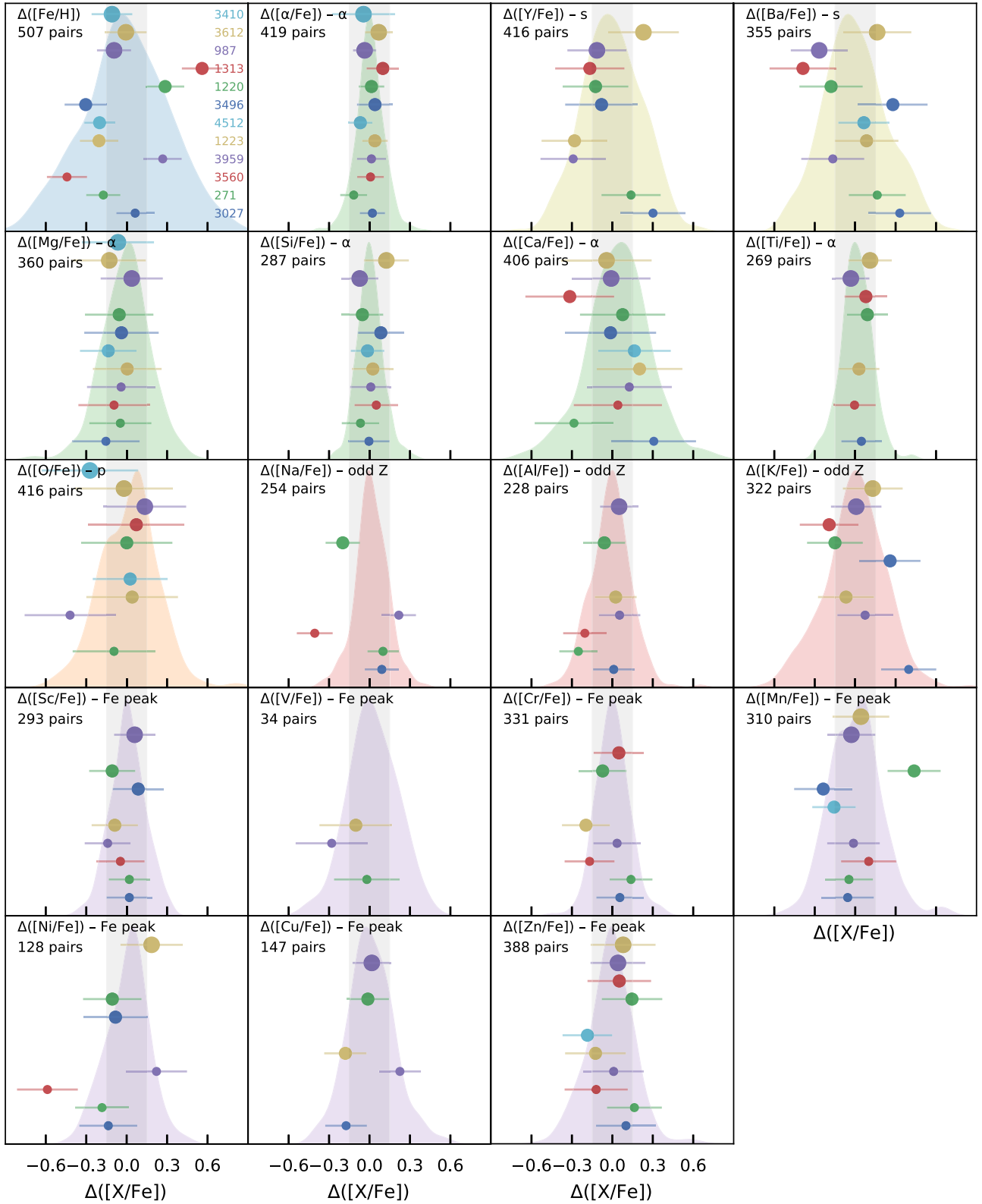
Looking ahead, the combination of spectroscopic abundance data from GALAH and other ongoing surveys with *Gaia* kinematics and distances will be extremely powerful. Even in this small sample of 15 reported co-moving groups, we have found three examples of pairs that are consistent with their being from unique formation sites. The large number of abundances available via the GALAH survey allows us to be much more confident with this chemical tagging result than if we had just radial velocities, metallicities, and perhaps an  $\alpha$ -abundance. Testing the assumption that each birth cluster will have a unique set of abundances will be very important. In this data set, comparing group 3612 with group 987, 3612 is more metal-rich but relatively depleted in Mg and O, while 987 is more metal-poor with a solar Mg abundance and a small enhancement in O. It is this type of unique abundance profile that we hope to exploit in chemical tagging.

Stars in open and globular clusters are a natural data set for investigating the prospects for chemical tagging and chemical homogeneity (i.e. it would be hoped that we can chemically tag 100 per cent of cluster members), but they do represent only a small number of star formation sites that are biased in some way; that is, they are unusually long-lived compared to most birth clusters. So, there is a great deal of information to be found in investigating the abundance behaviour of a large number of binary systems or co-moving stars (presumably co-natal systems in which 100 per cent of the stars can be resolved, unlike compact, dense star clusters).

Since there are so many more of these pairs than star clusters, they should provide a diverse sampling of the various processes that



**Figure 8.** Abundance patterns for the pairs in common between GALAH and O17. Large circles are used for the low  $\Delta(U, V, W)$  pairs [i.e.  $\Delta(U, V, W) < 20 \text{ km s}^{-1}$ ] and small symbols for the non-co-moving pairs. The colouring of the pairs is simply to aid the eye. The vertical ordering is by the difference in  $\Delta(U, V, W)$  velocity of the pair, from smallest to largest top to bottom, and is the same in all panels. Some pairs are missing from some panels because that particular element was flagged as unreliable in one or both of the stars. The background distribution is for all GALAH stars within about 600 pc of the Sun and the number of stars that make up the distribution is given in each panel.



**Figure 9.** Similar to Fig. 8 but showing the difference in the abundances for each pair. The background distributions are the abundance differences of the 5755 random, spatially close pairs selected from GALAH, with each panel giving the number of useful pairs for that element. The error bars are the sum of the errors of the pair. The shaded vertical region shows  $\Delta([X/Fe]) \pm 0.15$  dex. Only the top three pairs and the bottom pair show similar metallicities.

are sources of noise in the chemical tagging signal – for example, diffusion (Dotter et al. 2017); the possible abundance signature of exoplanets (Meléndez et al. 2009); binary interactions (e.g. Hansen et al. 2015); and any underlying abundance inhomogeneity in star-forming environments (Feng & Krumholz 2014; Liu et al. 2016). Any inference about chemical homogeneity based on data from a single cluster will depend on that cluster’s particular formation environment and history. We cannot observe that environment directly, and it is not necessarily appropriate to apply the inference from one star-forming site to the entire disc, with its wide range of star-forming conditions across the history of the Milky Way. Even with a potentially significant false positive rate, the large sample of co-moving pairs reported by O17 represents a large number of star formation sites, and they will give a more general picture of both the intrinsic chemical homogeneity of co-natal stars and the various processes that disturb that homogeneity.

With the release of *Gaia* DR2, over 92 per cent of GALAH targets have proper motion errors less than  $0.15 \text{ mas yr}^{-1}$  and parallaxes with a precision better than 0.1 mas. Nearly all of the dwarfs and many of the giants in the GALAH sample have precise parallax measurements from *Gaia* DR2. This will make it possible to more thoroughly investigate dissolving but still kinematically related clusters and looser star formation sites. This will be aided by cases where we can determine ages with greater precision than the ages in this study, such as turn-off stars, massive main-sequence stars, and stars with spectroscopic age indicators (e.g. Martig et al. 2016), since stars formed at the same place and time must have the same age.

We are on the cusp of a dramatic change in our ability to understand the chemodynamical history of the Milky Way Galaxy. With GALAH and other surveys delivering chemical information and *Gaia* delivering dynamical information, we will for the first time gain a statistically significant population of stars that we can chemically and dynamically tag together.

## ACKNOWLEDGEMENTS

The GALAH survey is based on observations made at the Australian Astronomical Observatory, under programmes A/2013B/13, A/2014A/25, A/2015A/19, A/2017A/18. We acknowledge the traditional owners of the land on which the AAT stands, the Gamilaraay people, and pay our respects to elders past and present.

The following software and programming languages made this research possible: CONFIGURE (Miszalski et al. 2006); IRAF (Tody 1986; Tody 1993); PYTHON (versions 3.6); ASTROPY (version 3.0; Robitaille et al. 2013; Price-Whelan et al. 2018), a community-developed core PYTHON package for astronomy; PANDAS (version 0.23.4; McKinney 2010); TOPCAT (version 4.4; Taylor 2005); and GALPY (version 1.4; Bovy 2015).

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 DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

Parts of this research were conducted by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. JDS and SLM acknowledge the support of the Australian Research Council through Discovery Project grant DP180101791. SB and KL acknowledge funds from the Alexander von Humboldt Foundation in the framework of the Sofja Kovalevskaja Award endowed

by the Federal Ministry of Education and Research. LD gratefully acknowledges a scholarship from Zonta International District 24 and support from Australian Research Council grant DP160103747. KL acknowledges funds from the Swedish Research Council (Grant nr. 2015-00415.3) and Marie Skłodowska Curie Actions (Cofund Project INCA 600398). TZ acknowledge the financial support from the Slovenian Research Agency (research core funding No. P1-0188). A. R. C. acknowledges support through the Australian Research Council through grant DP160100637. LD, KF, and Y-ST are grateful for support from the Australian Research Council grant DP160103747.

## REFERENCES

- Anders F., Chiappini C., Santiago B. X., Matijević G., Queiroz A. B., Steinmetz M., 2018, *A&A*, 619, A125
- Andrews J. J., Chanamé J., Agüeros M. A., 2017, *MNRAS*, 472, 675
- Andrews J. J., Chanamé J., Agüeros M. A., 2018a, *Res. Notes AAS*, 2, 29
- Andrews J. J., Chanamé J., Agüeros M. A., 2018b, *MNRAS*, 473, 5393
- Bensby T., Oey M. S., Feltzing S., Gustafsson B., 2007, *ApJ*, 655, L89
- Bensby T., Feltzing S., Oey M. S., 2014, *A&A*, 562, A71
- Blanco-Cuaresma S. et al., 2015, *A&A*, 577, A47
- Bland-Hawthorn J., Krumholz M. R., Freeman K., 2010, *ApJ*, 713, 166
- Bochanski J. J., Faherty J. K., Gagné J., Nelson O., Coker K., Smithka I., Desir D., Vasquez C., 2018, *AJ*, 155, 149
- Boesso R., Rocha-Pinto H. J., 2018, *MNRAS*, 474, 4010
- Bovy J., 2015, *ApJS*, 216, 29
- Bovy J., 2016, *ApJ*, 817, 49
- Bovy J., Hogg D., Roweis S., 2010, *Astrophysics Source Code Library*, record ascl: 1010.032
- Brewer J. M., Fischer D. A., Valenti J. A., Piskunov N., 2016, *ApJS*, 225, 32
- Brown A. G. A. et al., 2016, *A&A*, 595, A2
- Brown A. G. A. et al., 2018, *A&A*, 616, A1
- Bubar E. J., King J. R., 2010, *AJ*, 140, 293
- Buder S. et al., 2018, *MNRAS*, 478, 4513
- Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, *ApJ*, 823, 102
- De Silva G. M., Freeman K. C., Asplund M., Bland-Hawthorn J., Bessell M. S., Collet R., 2007, *AJ*, 133, 1161
- De Silva G. M., D’Orazi V., Melo C., Torres C. A. O., Gieles M., Quast G. R., Sterzik M., 2013, *MNRAS*, 431, 1005
- De Silva G. M. et al., 2015, *MNRAS*, 449, 2604
- Dhital S., West A. A., Stassun K. G., Bochanski J. J., 2010, *AJ*, 139, 2566
- Dotter A., 2016, *ApJS*, 222, 8
- Dotter A., Conroy C., Cargile P., Asplund M., 2017, *ApJ*, 840, 99
- Edvardsson B., Andersen J., Gustafsson B., Lambert D., Nissen P., Tomkin J., 1993, *A&AS*, 102, 603
- Famaey B., Jorissen A., Luri X., Mayor M., Udry S., Dejonghe H., Turon C., 2005, *A&A*, 430, 165
- Feng Y., Krumholz M. R., 2014, *Nature*, 513, 523
- Freeman K., Bland-Hawthorn J., 2002, *ARA&A*, 40, 487
- Friel E. D., 1995, *ARA&A*, 33, 381
- Fux R., 2001, *A&A*, 373, 511
- Gao X. et al., 2018, *MNRAS*, 481, 2666
- Goodwin S. P., Kroupa P., 2005, *A&A*, 439, 565
- Hansen T. T., Andersen J., Nordström B., Beers T. C., Yoon J., Buchhave L. A., 2015, *A&A*, 583, A49
- Hogg D. W. et al., 2016, *ApJ*, 833, 262
- Jiang Y.-F., Tremaine S., 2010, *MNRAS*, 401, 977
- Korn A. J., 2010, preprint ([arXiv:1010.2515](https://arxiv.org/abs/1010.2515))
- Kos J. et al., 2017, *MNRAS*, 464, 1259
- Kos J. et al., 2018, *MNRAS*, 473, 4612

Lewis I. J. et al., 2002, *MNRAS*, 333, 22  
 Lin J., Dotter A., Ting Y.-S., Asplund M., 2018, *MNRAS*, 477, 2966  
 Lind K. et al., 2017, *MNRAS*, 468, 4311  
 Lindegren L. et al., 2016, *A&A*, 4, A4  
 Liu F., Asplund M., Yong D., Meléndez J., Ramírez I., Karakas A. I., Carlos M., Marino A. F., 2016, *MNRAS*, 463, 696  
 Martell S. L. et al., 2017, *MNRAS*, 465, 3203  
 Martig M. et al., 2016, *MNRAS*, 456, 3655  
 McKinney W., 2010, in van der Walt S., Millman J., eds, Proceedings of the 9th Python in Science Conference, SciPy, Austin, p. 51  
 Meléndez J., Asplund M., Gustafsson B., Yong D., 2009, *ApJ*, 704, L66  
 Meléndez J. et al., 2017, *A&A*, 597, A34  
 Michalik D., Lindegren L., Hobbs D., 2015, *A&A*, 574, A115  
 Michaud G., Richard O., Richer J., VandenBerg D. A., 2004, *ApJ*, 606, 452  
 Miszalski B., Shortridge K., Saunders W., Parker Q. A., Croom S. M., 2006, *MNRAS*, 371, 1537  
 Mitschang A. W., De Silva G., Sharma S., Zucker D. B., 2013, *MNRAS*, 428, 2321  
 Ness M., Hogg D. W., Rix H.-W., Ho A. Y. Q., Zasowski G., 2015, *ApJ*, 808, 16  
 Ness M. et al., 2018, *ApJ*, 853, 198  
 Oelkers R. J., Stassun K. G., Dhital S., 2017, *AJ*, 153, 259  
 Oh S., Price-Whelan A. M., Hogg D. W., Morton T. D., Spergel D. N., 2017, *AJ*, 153, 257 (O17)  
 Oh S., Price-Whelan A. M., Brewer J. M., Hogg D. W., Spergel D. N., Myles J., 2018, *ApJ*, 854, 138  
 Pasetto S. et al., 2012, *A&A*, 547, A70  
 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  
 Pecaui M. J., Mamajek E. E., Bubar E. J., 2012, *ApJ*, 746, 154  
 Piskunov N., Valenti J. A., 2017, *A&A*, 597, A16  
 Price-Jones N., Bovy J., 2018, *MNRAS*, 475, 1410r  
 Price-Whelan A. M., Oh S., Spergel D. N., 2017, preprint ([arXiv:1709.03532](https://arxiv.org/abs/1709.03532))  
 Price-Whelan A. M. et al., 2018, *AJ*, 156, 123  
 Prusti T. et al., 2016, *A&A*, 595, A1  
 Quillen A. C. et al., 2018, *MNRAS*, 478, 228  
 Robitaille T. P. et al., 2013, *A&A*, 558, A33  
 Roederer I. U., Sneden C., Thompson I. B., Preston G. W., Shtetman S. A., 2010, *ApJ*, 711, 573  
 Schönrich R., Binney J., Dehnen W., 2010, *MNRAS*, 403, 1829  
 Sheinis A. et al., 2015, *J. Astron. Telesc. Inst. Syst.*, 1, 035002  
 Stello D. et al., 2017, *ApJ*, 835, 83  
 Taylor M., 2005, in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, ADASS XIV. Astron. Soc. Pac., San Francisco, p. 29  
 Ting Y.-S., Conroy C., Goodman A., 2015, *ApJ*, 807, 104  
 Tody D., 1986, in David L. Crawford, ed., Proc. SPIE., SPIE, Tucson, United States, pp. 0627–0627 - 16  
 Tody D., 1993, in Hanisch R., Brissenden R., Barnes J., eds, ASP Conf. Ser. Vol. 52, Astronomical Data Analysis Software and Systems II. Astron. Soc. Pac., San Francisco, p. 173  
 Traven G. et al., 2017, *ApJS*, 228, 24  
 Valenti J. A., Piskunov N., 1996, *A&AS*, 118, 595  
 Wright N. J., Bouy H., Drew J. E., Sarro L. M., Bertin E., Cuillandre J.-C., Barrado D., 2016, *MNRAS*, 460, 2593

## SUPPORTING INFORMATION

Supplementary data are available at [MNRAS](https://www.mnras.org) online.

**Table 1.** The stellar parameters and orbital characteristics of the 15 pairs.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

<sup>1</sup>*School of Physics, UNSW, Sydney, NSW 2052, Australia*

<sup>2</sup>*Australian Astronomical Observatory, 105 Delhi Rd, North Ryde, NSW 2113, Australia*

<sup>3</sup>*Centre of Excellence for Astrophysics in Three Dimensions (ASTRO-3D), Australia*

<sup>4</sup>*Research School of Astronomy & Astrophysics, Australian National University, ACT 2611, Australia*

<sup>5</sup>*School of Physics and Astronomy, Monash University, Clayton 3800, Victoria, Australia*

<sup>6</sup>*Faculty of Information Technology, Monash University, Clayton 3800, Victoria, Australia*

<sup>7</sup>*University of Southern Queensland, Toowoomba, Queensland 4350, Australia*

<sup>8</sup>*Institute for Advanced Study, Princeton, NJ 08540, USA*

<sup>9</sup>*Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA*

<sup>10</sup>*Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101, USA*

<sup>11</sup>*Center for Astrophysical Sciences and Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA*

<sup>12</sup>*Sydney Institute for Astronomy, School of Physics, A28, The University of Sydney, NSW, 2006, Australia*

<sup>13</sup>*Department of Astronomy, Columbia University, Pupin Physics Laboratories, New York, NY 10027, USA*

<sup>14</sup>*Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia*

<sup>15</sup>*School of Physical and Chemical Sciences, University of Canterbury, Private Bag 4800, Christchurch, New Zealand*

<sup>16</sup>*Monash Centre for Astrophysics, School of Physics and Astronomy, Monash University, Clayton 3800, Australia*

<sup>17</sup>*Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia*

<sup>18</sup>*Max Planck Institute for Astronomy (MPIA), Königstuhl 17, D-69117 Heidelberg, Germany*

<sup>19</sup>*Istituto Nazionale di Astrofisica, Osservatorio Astronomico di Padova, vicolo dell'Osservatorio 5, I-35122 Padova, Italy*

<sup>20</sup>*Australian Astronomical Optics, Macquarie University, 105 Delhi Rd, North Ryde, NSW 2113, Australia*

<sup>21</sup>*Department of Physics and Astronomy, Uppsala University, Box 517, SE-751 20 Uppsala, Sweden*

<sup>22</sup>*ICRAR, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.