

# The GALAH survey: unresolved triple Sun-like stars discovered by the *Gaia* mission

Klemen Čotar<sup>1</sup>,<sup>1</sup>★ Tomaž Zwitter<sup>1</sup>, Gregor Traven<sup>2</sup>, Janez Kos<sup>1</sup>, Martin Asplund<sup>3,4</sup>, Joss Bland-Hawthorn<sup>3,5,6</sup>, Sven Buder<sup>7</sup>, Valentina D’Orazi<sup>8</sup>, Gayandhi M. De Silva<sup>9</sup>, Jane Lin<sup>3,4</sup>, Sarah L. Martell<sup>3,10</sup>, Sanjib Sharma<sup>3,5</sup>, Jeffrey D. Simpson<sup>10</sup>, Daniel B. Zucker<sup>9</sup>, Jonathan Horner<sup>11</sup>, Geraint F. Lewis<sup>5</sup>, Thomas Nordlander<sup>3,4</sup>, Yuan-Sen Ting<sup>12,13,14</sup>, Rob A. Wittenmyer<sup>11</sup>  
and the GALAH collaboration

<sup>1</sup>Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia

<sup>2</sup>Lund Observatory, Department of Astronomy and Theoretical Physics, Box 43, SE-221 00 Lund, Sweden

<sup>3</sup>ARC Center of Excellence for All-Sky Astrophysics in Three Dimensions (ASTRO-3D), Australia

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

<sup>5</sup>Sydney Institute for Astronomy, School of Physics, A28, The University of Sydney, NSW 2006, Australia

<sup>6</sup>Miller Professor, Miller Institute, UC Berkeley, Berkeley, CA 94720, USA

<sup>7</sup>Max Planck Institute for Astronomy (MPIA), Koenigstuhl 17, D-69117 Heidelberg, Germany

<sup>8</sup>INAF Osservatorio Astronomico di Padova, vicolo dell’Osservatorio 5, I-35122 Padova, Italy

<sup>9</sup>Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia

<sup>10</sup>School of Physics, UNSW, Sydney, NSW 2052, Australia

<sup>11</sup>Centre for Astrophysics, University of Southern Queensland, Toowoomba, QLD 4350, Australia

<sup>12</sup>Institute for Advanced Study, Princeton, NJ 08540, USA

<sup>13</sup>Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101, USA

<sup>14</sup>Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

Accepted 2019 May 16. Received 2019 May 16; in original form 2019 April 9

## ABSTRACT

The latest *Gaia* data release enables us to accurately identify stars that are more luminous than would be expected on the basis of their spectral type and distance. During an investigation of the 329 best solar twin candidates uncovered among the spectra acquired by the GALAH survey, we identified 64 such overluminous stars. In order to investigate their exact composition, we developed a data-driven methodology that can generate a synthetic photometric signature and spectrum of a single star. By combining multiple such synthetic stars into an unresolved binary or triple system and comparing the results to the actual photometric and spectroscopic observations, we uncovered 6 definitive triple stellar system candidates and an additional 14 potential candidates whose combined spectrum mimics the solar spectrum. Considering the volume correction factor for a magnitude-limited survey, the fraction of probable unresolved triple stars with long orbital periods is  $\sim 2$  per cent. Possible orbital configurations of the candidates were investigated using the selection and observational limits. To validate the discovered multiplicity fraction, the same procedure was used to evaluate the multiplicity fraction of other stellar types.

**Key words:** methods: data analysis – binaries: general – stars: solar-type – Galaxy: stellar content.

## 1 INTRODUCTION

The investigation of solar-type stars in the solar neighbourhood has revealed that around half of them are found in binary or more

complex stellar systems (Raghavan et al. 2010; Duchêne & Kraus 2013; Moe & Di Stefano 2017). Of all multiple systems, about 13 per cent are part of higher-order hierarchical systems (Raghavan et al. 2010; Tokovinin 2014). Beyond the solar neighbourhood, the angular separation between members of such multiple star systems becomes too small for their components to be spatially resolvable in the sky. As a result, all stellar components contribute

\* E-mail: klemen.cotar@mf.uni-lj.si

to the light observed by spectroscopic, photometric, and astrometric surveys. It has been suggested that the population of binaries in the field could be even higher than in the solar neighbourhood (Quist & Lindgren 2000); therefore, a combination of multiple complementary approaches must be used to detect and analyse multiple stellar systems with different properties (Moe & Di Stefano 2017).

If the orbital period of such a system is relatively short, with high orbital velocities, it can be spectroscopically identified as a multiple system in two ways. When the components are of comparable luminosities, the effect of multiple absorption lines can be observed in the system’s spectrum. Such an object is also known as a double-lined binary (SB2) system (Pourbaix et al. 2004; Matijević et al. 2010; Fernandez et al. 2017; Merle et al. 2017). By contrast, a single-lined binary (SB1) system does not show the same effect, as the secondary component is too faint to significantly contribute to the observables. Short-period SB1 systems can be identified from the periodic radial velocity (RV) variations if multi-epoch spectroscopy is available (Duquennoy & Mayor 1991; Nordström et al. 2004; Matijević et al. 2011; Troup et al. 2016; Badenes et al. 2018). Other extrema are very wide binaries (Garnavich 1988; Close, Richer & Crabtree 1990; Gould et al. 1995; Kraus & Hillenbrand 2009; Shaya & Olling 2011; Andrews, Chanamé & Agüeros 2017; Coronado et al. 2018) and co-moving pairs (Oh et al. 2017; Jiménez-Esteban, Solano & Rodrigo 2019) that can only be identified by their spatial proximity and common velocity vector.

Duchêne & Kraus (2013) summarized that the majority of solar-type stars are part of binary systems with periods of hundreds to thousands of years, whose period distribution reaches a maximum at  $\log(P) \approx 5$ , for  $P$  measured in days. Because of the wide separation and long orbital periods of the components in such a scenario, the RV variations will have both low amplitude and a long period. This makes them challenging to detect in large-scale spectroscopic surveys, which typically last for less than a decade, and have a low number of revisits. A spectrum of such a binary or triple still contains a spectroscopic signature of all members, and those contributions can be disentangled into individual components (El-Badry et al. 2018a,b). Such a decomposition is easier when a binary consists of spectrally different stars (Skopal 2005; Rebassa-Mansergas et al. 2007, 2012; Ren et al. 2013; Rebassa-Mansergas et al. 2016), but becomes much harder or near-impossible when the composite spectrum consists of contribution from near-identical stars whose individual radial velocities are almost identical (Bergmann et al. 2015). In that case, additional photometric and distance measurement have to be used to constrain possible combinations (Widmark, Leistedt & Hogg 2018). If spectroscopic data are not available, determination of multiples can be attempted purely on the basis of photometric information (Ferraro et al. 1997; Hurley & Tout 1998; Milone et al. 2016; Widmark et al. 2018), but such approaches are limited to certain stellar types, and yield results that might be polluted with young pre-main-sequence stars (Žerjal et al. 2019).

In the scope of the GALactic Archaeology with HERMES (GALAH) survey, most stars are observed at only one epoch, limiting the number of available data points that can be used to identify and characterize unresolved spectroscopic binaries. For this study, GALAH observations were complemented by multiple photometric surveys presented in Section 2. Both spectroscopic (Section 4.1) and photometric (Section 4.2) observations were used to create data-driven models of a single star. Those models were used to analyse solar twin candidates detected amongst the GALAH spectra (Section 3). Sections 5 and 6 describe the employed fitting procedure that was used to determine multiplicity and physical pa-

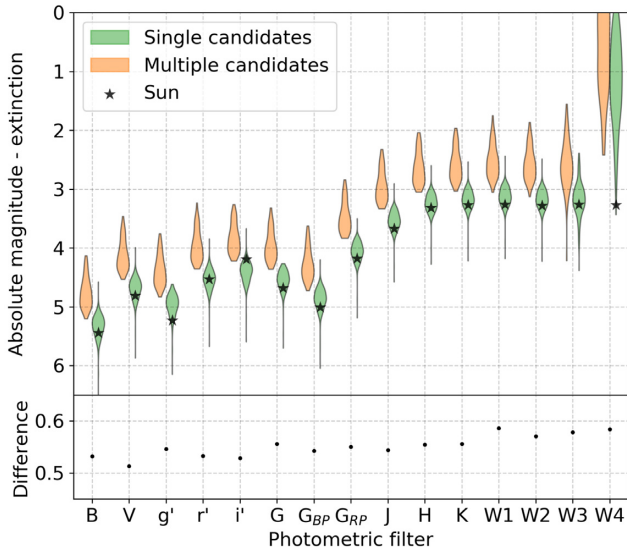
rameters of the twin systems. In Section 7 we determine constraints on orbital periods of the detected triple candidates. Constraints are imposed using the observational limits and simulations described in Section 8, where a population bias introduced by the magnitude-limited survey is assessed in Section 8.4. The concluding Section 9 summarizes the results and introduces additional approaches that could be used to verify the results.

## 2 DATA

In this work, we analyse a set of high-resolution stellar spectra that were acquired by the High Efficiency and Resolution Multi-Element Spectrograph (HERMES; Barden et al. 2010; Sheinis et al. 2015), a multifibre spectrograph mounted on the 3.9-metre Anglo-Australian Telescope (AAT). The spectrograph has a resolving power of  $R \sim 28\,000$  and covers four wavelength ranges (4713–4903, 5648–5873, 6478–6737, and 7585–7887 Å), frequently also referred to as ‘spectral arms’. Observations used in this study have been taken from multiple observing programmes that make use of the HERMES spectrograph: the main GALAH survey (De Silva et al. 2015), the K2-HERMES survey (Wittenmyer et al. 2018), and the TESS-HERMES survey (Sharma et al. 2018). Those observing programmes mostly observe stars at higher Galactic latitudes ( $|b| > 10^\circ$ ) and employ different selection functions, but share the same observing procedures, reduction, and analysis pipeline (internal version 5.3; Kos et al. 2017). All three programmes are magnitude-limited, with no colour cuts (except the K2-HERMES survey), and observations predominantly fall in the  $V$  magnitude range between 12 and 14. This leads to an unbiased sample of mostly southern stars that can be used for different population studies, such as multiple stellar systems in our case. Additionally, stellar atmospheric parameters and individual abundances derived from spectra acquired during different observing programmes are analysed with the same *The Cannon* procedure (internal version 182112 that uses parallax information to infer  $\log g$  of a star; Ness et al. 2015; Buder et al. 2018), so they are intercomparable. *The Cannon* is a data-driven interpolation approach trained on a set of stellar spectra that span the majority of the stellar parameter space (for details, see Buder et al. 2018). Whenever we refer to valid or unflagged stellar parameters in the text, only stars with the quality flag `flag_cannon` equal to 0 were selected.

With the latest *Gaia* DR2 release (Gaia Collaboration 2016, 2018), the determination of distance for stars within a few kpc of the Sun becomes straightforward, and the derived distances are accurate to a few per cent. Along with the measurements of parallax and proper motion, stellar magnitudes in up to three photometric bands ( $G$ ,  $G_{BP}$ , and  $G_{RP}$ ) are provided. Coupling those two measurements together, we can determine the absolute magnitude and luminosity of a star. This can be done for the majority of stars observed in the scope of the GALAH survey as more than 99 per cent of them can be matched with sources in *Gaia* DR2. Although the uncertainty in *Gaia* mean radial velocities are much larger than those from GALAH (Zwitter et al. 2018), especially at its faint limit, they can still be used in the multi-epoch analysis (Section 7.3) to determine the lower boundary of its variability.

Because observations of both surveys were acquired at approximately the same epoch ranges, wide binaries with long orbital periods would show little variability in their projected velocities. Observational time-series were therefore extended into the past with the latest data release from The Radial Velocity Experiment (RAVE DR5; Kunder et al. 2017) which also surveyed the southern part of the sky. More recent spectra of accessible (stars’ declination



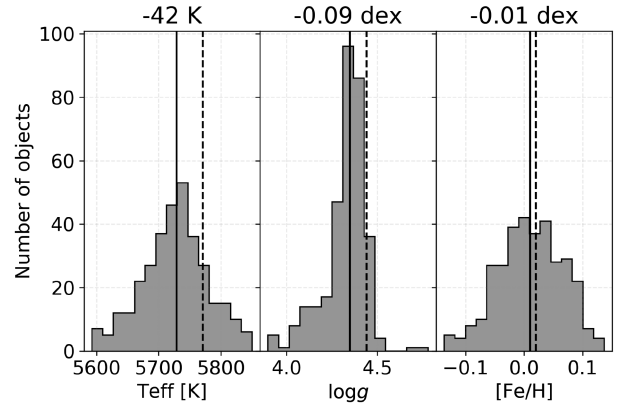
**Figure 1.** Violin density plots showing the distribution of extinction-corrected absolute magnitudes in multiple photometric systems. Separate distributions are given for stars that we considered as single and multiple in our analysis. Star symbols indicate the absolute magnitude of the Sun (Willmer 2018). The bottom panel shows a difference between the median magnitudes of both distributions.

$> -25^\circ$  and  $G$  magnitude  $< 13.5$ ) multiple candidates were acquired by the high-resolution Echelle spectrograph (with the resolving power  $R \sim 20\,000$ ) mounted on the 1.82-m Copernico telescope located at Cima Ekar in Asiago, Italy. As this telescope is located in a northern observatory, we were able to observe only a limited number of stars. Every observation consisted of two half-hour-long exposures that were fully reduced, shifted to the heliocentric frame, combined together, and normalized. The system’s radial velocity was determined by cross-correlating the observed spectrum with that of the Sun. Data from those two additional spectroscopic surveys enabled us to further constrain the variability of the analysed systems.

For a broader wavelength coverage, additional photometric data were taken from three large all-sky surveys. In the visual part of the spectrum we rely on the AAVSO Photometric All-Sky Survey (APASS; Henden et al. 2015)  $B$ ,  $V$ ,  $g'$ ,  $r'$ ,  $i'$  bands, which are supplemented by the Two Micron All-Sky Survey (2MASS; Skrutskie et al. 2006)  $J$ ,  $H$ ,  $K_s$  bands and the *Wide-field Infrared Survey Explorer* (*WISE*; Wright et al. 2010)  $W1$ ,  $W2$  bands. All of those surveys were matched with the GALAH targets, which resulted in up to 13 photometric observations per star. Photometric values and their uncertainties were taken as published in these catalogues, ignoring any specific quality flags. During the initial investigation of their usefulness, the *WISE*  $W3$  and  $W4$  bands proved to be unreliable for our application and were therefore removed from further use. The main reason for their removal is a large scatter in magnitude measurements of similar stars and a strong overlap between single and multiple stars evident in Fig. 1.

### 3 SOLAR-LIKE SPECTRA

The selection of the most probable solar twin candidates observed in the scope of the GALAH survey was performed purely on the basis of spectral comparison using the Canberra distance metric



**Figure 2.** Distribution of physical parameters for discovered solar twin candidates. Values above the plots represent the difference between the median of the distribution (solid vertical line) and actual solar values (dashed vertical line).

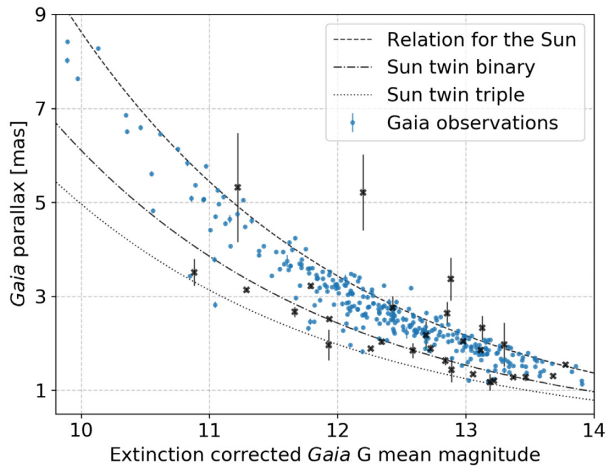
(Lance & Williams 1967). It is defined as

$$\text{distance}_{\text{Canberra}}(f_{\odot}, f_{\text{obs}}) = \sum_{\lambda} \frac{|f_{\odot,\lambda} - f_{\text{obs},\lambda}|}{|f_{\odot,\lambda}| + |f_{\text{obs},\lambda}|}, \quad (1)$$

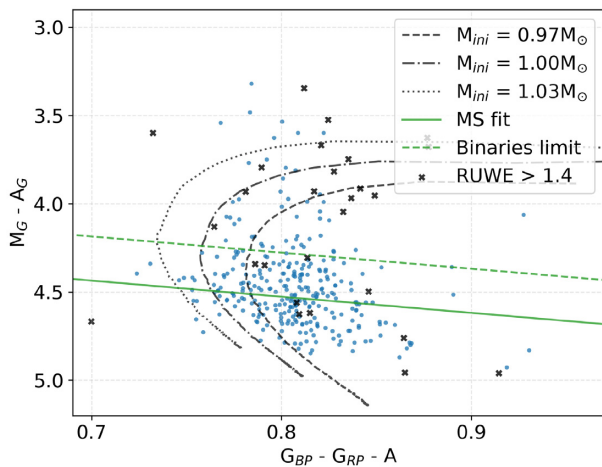
where  $f_{\odot}$  is the reference solar spectrum and  $f_{\text{obs}}$  the observed spectrum. This avoided the need for prior knowledge about the physical parameters of the considered stars. Observed spectra were compared with a reference twilight solar spectrum that was acquired by the same spectrograph. The comparison was done in two steps. First, the observed spectra were compared arm by arm, where only 7 per cent of the best-matching candidates were selected for every HERMES arm. The selection does not consist of only one hard threshold on spectral similarity, but it varies as a power law of spectral signal-to-noise ratio (SNR). Larger discrepancies between spectra are therefore allowed for spectra with low SNR. The final selection consisted of a set of stars whose spectra were adequately similar to the solar spectrum in every arm. Further processing consisted of modelling the spectral noise and the comparison of individual chemical abundances. For details about spectral comparison and solar twin determination see Čotar et al. (in preparation). The parameter distribution for the selected set of objects is presented in Fig. 2. As the computation of  $\log g$  uses prior knowledge about the absolute magnitude of a star, lower  $\log g$  values are attributed to objects that exhibit excess luminosity.

#### 3.1 Candidate multiple systems

Among the determined solar twin candidates, we noticed a photometric trend that is inconsistent with the distance of an object that resembles the Sun. Solar twins, mimicking the observed solar spectrum, should also be similar to it in all other observables such as luminosity, effective temperature, surface gravity, chemical composition, and absolute magnitude. Plotting their apparent magnitude against *Gaia* parallax measurement (Fig. 3), all of the detected stars should lie near or on the theoretical line, describing the same relation for the Sun observed at different parallaxes. As the magnitude of the Sun is not directly measured by the *Gaia* or determined by the *Gaia* team, we computed its absolute magnitude using the relations published by Evans et al. (2018) that connect the *Gaia* photometric system with other photometric systems. The reference solar magnitudes (in multiple filters) that were used in the computation were taken from Willmer (2018). The



**Figure 3.** Parallax versus measured apparent *Gaia* *G* magnitude for our solar twin candidates. Stars marked with black crosses have large normalized astrometric uncertainties ( $\text{RUWE} > 1.4$ ), which may lead to wrongly determined distance and consequently multiplicity results. The dashed line represents a theoretical absolute magnitude of the Sun as it would be observed at different parallaxes. Similarly, the relation for a binary and a triple system composed of multiple solar twins is plotted with the dash-dotted and dotted lines.



**Figure 4.** *Gaia* extinction-corrected absolute *G* magnitude and colour index computed from the  $G_{BP}$  and  $G_{RP}$  bands. Stars marked with black crosses have large normalized astrometric uncertainty ( $\text{RUWE} > 1.4$ ). The green dashed line represents a threshold that was used as a delimiter between objects treated as multiple and single stars. Overlaid evolutionary tracks, constructed from the PARSEC isochrones (Marigo et al. 2017), represent an evolution of stars with solar-like initial mass  $M_{\text{ini}}$  and metallicity  $[\text{M}/\text{H}] = 0$  for stellar ages between 0.1 and 12 Gyr.

resulting absolute *G* magnitude of the Sun is  $4.68 \pm 0.02$ , where the uncertainty comes from the use of multiple relations. This value also coincides with the synthetic *Gaia* photometry produced by Casagrande & VandenBerg (2018), who determined the magnitude of the Sun to be  $M_{G,\odot} = 4.67$ .

Within our sample of probable solar twins, we identified 64 stars that show signs of being too bright at a given parallax. In Fig. 3 they are noticeable as a sequence of data points that lie below the theoretical line and are parallel to it. Another even more obvious indication of their excess luminosity is given by the colour–magnitude diagram in Fig. 4. There, the same group of stars is

brighter by  $\sim 0.7$  mag when comparing stars with the same colour index. As both groups of stars are visually separable, the multiple stellar candidates can easily be isolated by selecting objects with an extinction-corrected absolute *G* magnitude above the binary limit line shown in Fig. 4. To compute the absolute magnitudes, we used the distance to stars inferred by the Bayesian approach that takes into account the distribution of stars in the Galaxy (Bailer-Jones et al. 2018). As the reddening published along the *Gaia* DR2 (Andrae et al. 2018) could be wrong for stars located away from the used set of isochrones, we took the information about the reddening at specific sky locations and distances from the all-sky three-dimensional dust map produced by Capitanio et al. (2017). To infer a band-dependent extinction from the acquired reddening, a reddening coefficient ( $R$ ) was used. The values of  $R$ , considering the extinction law  $R_V = 3.1$ , were taken from the tabulated results published in Schlafly & Finkbeiner (2011).

To determine the limiting threshold between single and multiple candidates, we first fit a linear representation of the main sequence to the median of the absolute magnitudes distribution in the 0.02-mag-wide colour bins. The lower limit for the binaries was placed 0.25 mag above the fitted line.

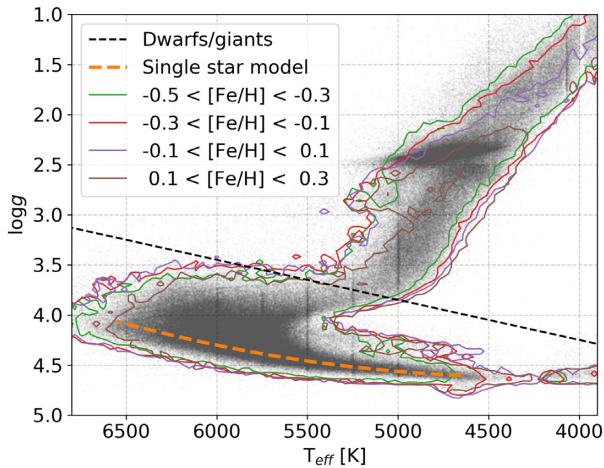
Confirmation that this extra flux could be contributed by the unseen companion also comes from other photometric systems, where the distribution of absolute magnitudes for both groups is shown in Fig. 1. On average, multiple candidates are brighter by  $\sim 0.55$  mag in every considered band. For an identical binary system, a measured magnitude excess would be 0.75, and 1.2 for a triple system. As the observed difference is not constant in every band, as would be expected for a system composed of identical stars, we expect some differences in parameters between the components of the system. This can be said under the assumption that all considered photometric measurements were performed at the same time. Of course, that is not exactly true in our case as the acquisition time between different photometric surveys can vary by a few to 10 yr. As the Sun-like stars are normal, low-activity stars, this effect is most probably negligible, but events like occultations between the stars in the system can still occur.

## 4 SINGLE-STAR MODELS

Once the selection of interesting stars was performed, we began with the analysis of their possible multiplicity. Our procedure for the analysis of suspected multiple stellar systems is based on spectroscopic and photometric data-driven single-star models that were constructed from observations taken from multiple large sky surveys. With this approach, we exclude assumptions about stellar properties and populations that are usually used to generate synthetic data. In this section, we describe approaches that were used to create those models.

### 4.1 Spectroscopic model

Every stellar spectrum can be largely described using four basic physical stellar parameters:  $T_{\text{eff}}$ ,  $\log g$ , chemical composition, and  $v \sin i$ . To construct a model that would be able to recreate a spectrum corresponding to any conceivable combination of those parameters, we used a data-driven approach named *The Cannon*. The model was trained on a set of normalized GALAH spectra that meet the following criteria: The spectrum must not be flagged as peculiar (Traven et al. 2017), have an SNR per resolution element in the green arm  $> 20$ , does not contain any monitored reduction problems and have valid *The Cannon* stellar parameters. Additionally, we limited



**Figure 5.** Complete observational set of valid stellar parameters shown as a varied density of grey dots. Contours around the diagram illustrate a coverage of parameter space in different  $[\text{Fe}/\text{H}]$  bins. The overlaid orange dashed curve represents a relation on the main sequence from where single stars considered in the fit are drawn. Additionally, the black dashed linear line represents an arbitrarily defined limit between giant and dwarf stars that were used in the process of training a spectroscopic single-star model.

our set to main-sequence dwarf stars (below the arbitrarily defined line shown in Fig. 5) as giants are not considered for our analysis. Additionally, the decision not to consider giants was taken as a result of the fact that accurate modelling of their spectra requires information about their luminosity. It should be noted that the application of these limits does not ensure that our training set is completely free from spectra of unresolved (or even clearly resolved SB2 binaries), as would be desired in the case of an ideal training set.

In order to train the model, all spectra were first shifted to the rest frame by the reduction pipeline (Kos et al. 2017), and then linearly interpolated on to a common wavelength grid. The training procedure consists of minimizing a loss function between an internal model of *The Cannon* and observations for every pixel of a spectrum (Ness et al. 2015).

The result of this training procedure are quadratic relations that take the desired stellar parameters  $T_{\text{eff}}$ ,  $\log g$ ,  $[\text{Fe}/\text{H}]$ , and  $v \sin i$  to reconstruct a target spectrum. Spectra generated in this manner are trustworthy only within the parameter space defined by the training set, where the main limitation is the effective temperature, which ranges from  $\sim 4600$  to  $\sim 6700$  K on the main sequence. Spectra of hotter stars are easy to reproduce, but they lack the elemental absorption lines that we would like to analyse. On the other hand, spectra of colder stars are packed with molecular absorption lines and are therefore harder to reproduce and analyse. The model itself can be used to extrapolate spectra outside the initial training set, but as they cannot be verified, they were not considered to be useful for the analysis.

#### 4.2 Photometric model

With the use of a model that produces normalized spectra for every given set of stellar parameters, we lose all information about the stars' colour, luminosity, and spectral energy distribution. This can be overcome using another model that generates the photometric signature of a desired star. To create this kind of a model, we first collected up to 13 apparent magnitudes from

the selected photometric surveys (*Gaia*, APASS, 2MASS, and *WISE*) for every star in the GALAH survey. Whenever possible, these values were converted to absolute magnitudes using the distance to stars inferred by the Bayesian approach (Bailer-Jones et al. 2018). Before using the pre-computed published distances, we removed all sources whose computed astrometric renormalized unit-weight error (RUWE; Lindgren 2018) was greater than 1.4. The magnitudes of every individual star were also corrected for the reddening effect, except for the *WISE* photometric bands W1 and W2, which were considered to be extinction free.

Using the valid *The Cannon* stellar parameters, the inferred and corrected absolute magnitudes in multiple photometric bands were grouped into bins that contain stars within  $\Delta T_{\text{eff}} = \pm 80$  K,  $\Delta \log g = \pm 0.05$  dex, and  $\Delta [\text{Fe}/\text{H}] = \pm 0.1$  dex around the bin centre. As spectroscopically unresolved multiple stellar systems could still be present in this bin, median photometric values are computed per bin to minimize their effect. Extrapolation outside this grid, which covers the complete observational stellar parameter space, is not desired or implemented as it may produce erroneous values. When a photometric signature of a star with parameters between the grid points is requested, it is recovered by linear interpolation between the neighbouring grid points. The median photometric signature for the requested stellar parameters could also be computed on-the-fly using the same binning, but we found out that this produced insignificant difference and increased the processing time by more than a factor of 2.

#### 4.3 Limitations in the parameter space

As already emphasized, spectroscopic and photometric models were built on real observations and are therefore limited by the training set coverage. The limitations are visually illustrated by Fig. 5, where the dashed curve, from which single stars are drawn, is plotted over the observations. This arbitrary quadratic function is defined as

$$\log g = 2.576 + 9.48 \times 10^{-4} T_{\text{eff}} - 1.10 \times 10^{-7} T_{\text{eff}}^2, \quad (2)$$

where values of  $T_{\text{eff}}$  and  $\log g$  are given in units of K and  $\text{cm s}^{-2}$ , respectively. The polynomial coefficients were determined by fitting a quadratic function to manually defined points that represent regions with the highest density of stars on the shown Kiel diagram. From equation (2), it follows that the  $\log g$  of a selected single star is not varied freely, but computed from the selected  $T_{\text{eff}}$  whose range is limited within the values  $6550 > T_{\text{eff}} > 4650$  K.

Focusing on solar-like stars gives us an advantage in their modelling as the whole observational diagram in Fig. 5 is sufficiently populated with stars of solar-like iron abundance. When going towards more extreme  $[\text{Fe}/\text{H}]$  values (high or low), coverage of the main sequence starts to decrease. For cooler stars, this happens at low iron abundance ( $[\text{Fe}/\text{H}] < -0.3$ ) and for hot stars at high abundance ( $[\text{Fe}/\text{H}] > 0.3$ ). Those limits pose no problems for our analysis, unless the wrong stellar configuration is used to describe the observations. At that point, the spectroscopic fitting procedure would try to compensate for too deep or too shallow spectral lines (effect of wrongly selected  $T_{\text{eff}}$ ) by decreasing or increasing  $[\text{Fe}/\text{H}]$  beyond values reasonable for solar-like objects.

## 5 CHARACTERIZATION OF MULTIPLE-SYSTEM CANDIDATES

For the detailed characterization of solar twin candidates that show excess luminosity, we used their complete available photometric

and spectroscopic information. The excess luminosity can only be explained by the presence of an additional stellar component or a star that is hotter or larger than the Sun. Both of those cases can be investigated and confirmed by the data and models described in Sections 2 and 4. In the scope of our comparative methodology, we constructed a broad collection of synthetic single, double, and triple stellar systems that were compared and fitted to the observations.

As the measured *Gaia* DR2 parallaxes, and therefore inferred distances, of some objects are badly fitted or highly uncertain, the distance results provided by Bailer-Jones et al. (2018) actually yield three distinct distance estimates – the mode of an inferred distance distribution (`r_est`) and a near and distant distance (`r_lo` and `r_hi`) estimation, between which 68 percent of the distance estimations are distributed. As the actual shape of the distribution is not known and could be highly skewed, we did not draw multiple possible distances from the distribution, but only used its mode value.

### 5.1 Fitting procedure

A complete characterization and exploration fitting procedure for every stellar configuration (single, binary, and triple) consists of four consecutive steps, which are detailed in the following sections. As we are investigating solar twin spectra, the initial assumption for the iron abundance of the system is set to  $[\text{Fe}/\text{H}] = 0$ . This also includes the assumption that stars in a system are of the same age, at similar evolutionary stages, and were formed from a similarly enriched material. If that is true, we can set the iron abundance to be equal for all stars in the system. This notion is supported by the simulations (Kamdar et al. 2019a) and studies (Kamdar et al. 2019b) of field stars showing that close stars are very likely to be co-natal if their velocity separation is small.

The observed systems must be composed of multiple main-sequence stars; otherwise the giant companion would dominate the observables, and the system would not be a spectroscopic match to the Sun. Therefore, their parameters  $T_{\text{eff}}$  and  $\log g$  are drawn from the middle of the main sequence determined by *The Cannon* parameters in the scope of the GALAH survey. The Kiel diagram of the stars with valid parameters and model of the main-sequence isochrone used in the fitting procedure are shown in Fig. 5.

### 5.2 Photometric fitting – first step

With those initial assumptions in mind, we begin with the construction of the photometric signature of the selected stellar configuration. To find the best model that describes the observations, we employ a Bayesian Markov chain Monte Carlo (MCMC) fitting approach (Foreman-Mackey et al. 2013), where the varied parameter is the effective temperature of the components. The selected  $T_{\text{eff}}$  values, and inferred  $\log g$  (equation 2), are fed to the photometric model (Section 4.2) to predict a photometric signature of an individual component. Multiple stellar signatures are combined together into a single unresolved stellar source using the following equation:

$$M_{\text{model}} = -2.5 \log_{10} \left( \sum_{i=1}^{n_s} 10^{-0.4M_i} \right); \quad n_s = [1, 2, 3], \quad (3)$$

where  $M_i$  denotes the absolute magnitude of a star in one of the used photometric bands, and  $n_s$  number of components in a system. The newly constructed photometric signature at selected  $T_{\text{eff}}$  values is compared to the observations using the photometric log-likelihood function  $\ln p_P$  defined as

$$\ln p_P(T_{\text{eff}}|M, \sigma) = -\frac{1}{2} \sum_{i=1}^{n_p} \left[ \frac{(M_i - M_{\text{model},i})^2}{\sigma_i^2} + \ln(2\pi\sigma_i^2) \right], \quad (4)$$

where  $M$  and  $\sigma$  represent extinction-corrected absolute magnitudes and their measured uncertainties that were taken for multiple published catalogues presented in Section 2. The constructed photometric model of a multiple system is represented by the variable  $M_{\text{model}}$  and the number of photometric bands by  $n_p$ . The maximum, and most common, value for  $n_p$  is 13, but in some cases, it can drop to as low as 8. The MCMC procedure is employed to maximize this log-likelihood and find the best-fitting stellar components.

To determine the best possible combination of  $T_{\text{eff}}$  values, we initiate the fit with 1200 uniformly distributed random combinations of initial temperature values that span the parameter space shown in Fig. 5. The number of initial combinations is intentionally high in order to sufficiently explore the temperature space. Excessive or repeated variations of initial parameters are rearranged by a prior limitation that the temperatures of components must be decreasing; therefore,  $T_{\text{eff}1} \geq T_{\text{eff}2} \geq T_{\text{eff}3}$  in the case of a triple system (example of used initial walker parameters is shown in Fig. B1). The initial conditions are run for 200 steps. The number of steps was selected in such a way to ensure a convergence of all considered cases (example of the walker convergence is shown in Fig. B2). The distribution of priors for such a run is shown as a corner plot in Fig. 6. After that, only the best 150 walkers are kept, their values perturbed by 2 per cent, and run for another 200 steps to determine the posterior distribution of the parameters varied during the MCMC fit. This two-step run is needed to speed up the process and discard solutions with lower  $\ln p_P$ . During the initial tests, we found that the investigated parameter space can have multiple local minima that attract walkers, especially in the case of a triple system.

After the completion of all MCMC steps, the considered parameter combinations are ordered by their log-likelihood in descending order. Of those, only 10 per cent of the best-fitting combinations are used to compute final  $T_{\text{eff}[1-3]}$  values. They are computed as the median of the selected best combinations.

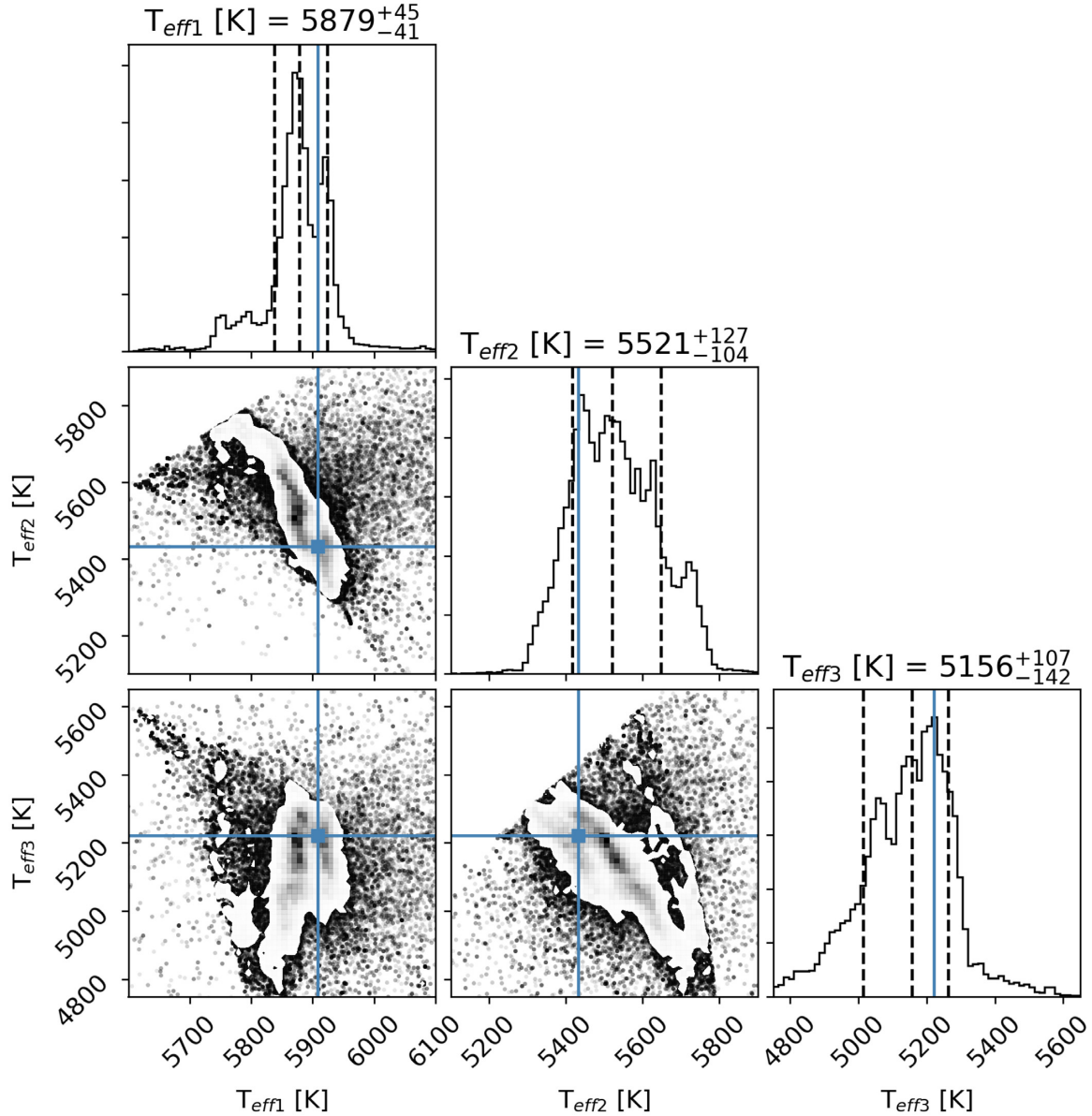
### 5.3 Spectroscopic fitting – second step

After an effective temperature of the components has been determined, we proceed with the evaluation of how well they reproduce the observed spectrum. As a majority of the fitted systems do not consist of multiple components with a  $T_{\text{eff}}$  equal to solar, the  $[\text{Fe}/\text{H}]$  of the system must be slightly changed to equalize absorption strength of the simulated and observed spectral lines.

To determine the  $[\text{Fe}/\text{H}]$  of the system, we first compute a simulated spectrum for every component using a spectroscopic model described in Section 4.1. Individual spectra are afterwards combined using equation (5) in the case of a binary system or equation (6) in the case of a triple system.

$$r_{12} = \frac{L_{2,\lambda}}{L_{1,\lambda}}; \quad \lambda = [1, 2, 3, 4] \quad (5)$$

$$f_{\text{model},\lambda} = \frac{f_{1,\lambda}}{1 + r_{12}} + \frac{f_{2,\lambda}}{1 + 1/r_{12}}$$



**Figure 6.** Distribution of considered posteriors during the initial MCMC photometric fit for one of the objects that were at the end classified as triple candidates. As the plots show the first step in the fitting procedure that explores the complete parameter space, the distributions are not expected to be smooth, because of possible multiple local minima. Values indicated on the scatter plots represent medians for the 10 per cent of the best fitting solutions.

$$r_{12} = \frac{L_{2,\lambda}}{L_{1,\lambda}}, \quad r_{13} = \frac{L_{3,\lambda}}{L_{1,\lambda}} \quad (6)$$

$$f_{\text{model},\lambda} = \frac{f_{1,\lambda}}{1 + r_{12} + r_{13}} + \frac{r_{12}f_{2,\lambda}}{1 + r_{12} + r_{13}} + \frac{r_{13}f_{3,\lambda}}{1 + r_{12} + r_{13}}$$

Individual normalized spectra, denoted by  $f_{n,\lambda}$  in equations (5) and (6), are weighted by the luminosity ratios between the components ( $r_{xy}$ ) and then summed together.

As the HERMES spectrum covers four spectral ranges, whose distribution of spectral energy depends on stellar  $T_{\text{eff}}$ , different luminosity ratios also have to be used for every spectral arm. Of all of the used photometric systems, the APASS filters  $B$ ,  $g'$ ,  $r'$ , and  $i'$  have the best spectral match with the blue, green, red, and infrared HERMES arm. The modelled APASS magnitudes of the same stars are used to compute luminosity ratios between them.

The described summation of the spectra introduces an additional assumption about the analysed object. With this step, we assume that components have a negligible internal spread of projected radial velocities that could otherwise introduce asymmetries in the shape of observed spectral lines. The assumption allows us to combine individual components without any wavelength corrections.

Similarly, as in the previous case, a Bayesian MCMC fitting procedure was used to maximize the spectroscopic log-likelihood  $\ln p_S$  defined as

$$\ln p_S([\text{Fe}/\text{H}]|f, \sigma) = -\frac{1}{2} \sum_{\lambda} \left[ \frac{(f_{\lambda} - f_{\text{model},\lambda})^2}{\sigma_{\lambda}^2} + \ln(2\pi\sigma_{\lambda}^2) \right], \quad (7)$$

where  $f$  and  $\sigma$  represent the observed spectrum and its per-pixel uncertainty, respectively. A modelled spectrum of the system, at

selected  $[\text{Fe}/\text{H}]$ , is represented by the variable  $f_{\text{model}}$  and the number of wavelength pixels in that model by  $\lambda$ . Combined, all four spectral bands consist of almost 16 000 pixels.  $T_{\text{eff}}$  values of the components are fixed for all considered cases.

The MCMC fit is initiated with 150 randomly selected  $[\text{Fe}/\text{H}]$  values, whose uniform distribution is centred at the initial  $[\text{Fe}/\text{H}]$  value of the system and has a span of 0.4 dex. All of the initiated walkers are run for 100 steps. At every  $[\text{Fe}/\text{H}]$  level, a new simulated spectrum composite is generated and compared to the observed spectrum by computing the log-likelihood  $\ln p_S$  of a selected  $[\text{Fe}/\text{H}]$  value. The range of possible  $[\text{Fe}/\text{H}]$  values considered in the fit is limited by a flat prior between  $-0.5$  and  $0.4$ .

By the definition of  $[\text{Fe}/\text{H}]$  in the scope of the GALAH *The Cannon* analysis, the parameter describes stellar iron abundance and not its metallicity as commonly used in the literature. Therefore, only spectral absorption regions of unblended Fe atomic lines are used to compute the spectral log-likelihood. Having to fit only one variable at a time, the solution is easily found and computed as a median value of all posteriors considered in the fit.

### 5.4 Final fit – third step

A changed value of  $[\text{Fe}/\text{H}]$  for the system will introduce subtle changes to its photometric signature; therefore, we reinitiate the photometric fitting procedure. It is equivalent to the procedure described in Section 5.3, but with much narrower initial conditions. These new initial conditions are uniformly drawn from the distribution centred at  $T_{\text{eff}}$  values determined in the first step of the fitting procedure. The width of the uniform distribution is equal to 100 K. Drawn initial conditions are afterwards run through the same procedure as described before.

At this point, the second and third steps in the fitting procedure can be repeatedly run several times to further pinpoint the best solution. We found out that further refinement was not needed in our case as it did not influence the determined number of stars in the system.

### 5.5 Number of stellar components – final classification

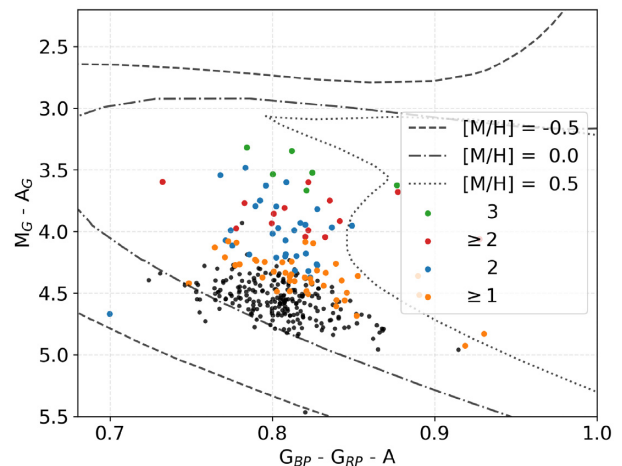
The fitting procedure described above was used to evaluate observations of every multiple stellar candidate to determine whether they belong to a single, binary, or triple stellar system. This resulted in the following set of results for every configuration: predicted  $T_{\text{eff}}$  of the components,  $[\text{Fe}/\text{H}]$  of the system, simulated spectrum, and simulated photometric signature of the system.

As the photometric and spectroscopic fits do not always agree on the best configuration, the following set of steps and rules was applied to classify results in one of six classes presented in Table 1.

- (i) Compute  $\chi^2$  between the simulated photometric signature of the modelled system and extinction-corrected absolute photometric observations for every considered stellar configuration.
- (ii) Compute  $\chi^2$  between the simulated spectrum of the modelled system and the complete GALAH observed spectrum for every considered stellar configuration.
- (iii) Independently select the best-fitting configuration with the lowest  $\chi^2$  for photometric and spectroscopic fit.
- (iv) If the best photometric and spectroscopic fits point to the same configuration, then the system is classified as having a number of stars defined by both fitting procedures.
- (v) If the best photometric and spectroscopic fits do not point to the same configuration, then the system is classified as having at

**Table 1.** Number of different systems discovered by the fitting procedure performed on possible multiple stars that exhibit excess luminosity.

Configuration classification	Number of systems
1 star	2
$\geq 1$ star	14
2 stars	27
$\geq 2$ stars	14
3 stars	6
Inconclusive	1
Total objects	64



**Figure 7.** Showing the same data points as Fig. 4 with indicated definite triple (green dots), possible triple (red dots), binary (blue dots), and possible binary stellar systems (orange dots). All other classes are shown with black dots. Overplotted are PARSEC isochrones (Marigo et al. 2017) for stars with the age of 4.5 Gyr and different metallicities.

least as many stars as determined by the prediction with a lower number of stars (e.g.  $\geq 2$  stars).

- (vi) If the difference between those two predictions is greater than 1 (e.g. photometric fit points to a single star and spectroscopic fit to a triple star), then the system is classified as inconclusive.

The classification produced using these rules is shown as a colour-coded *Gaia* colour–magnitude diagram in Fig. 7.

### 5.6 Quality flags

In addition to our final classification, we also provide additional quality checks that might help to identify cases for which our method might return questionable determination of a stellar configuration. Every one of those checks, listed in Table 2, is represented by one bit of a parameter `flag` in the final published table (Table A1). The first bit gives us an indication of whether the object could have an uncertain astrometric solution, whereas the second and third bits indicate if the final fitted solution has a worse match with the observations than the parameters produced by the *The Cannon* pipeline. To evaluate this, we used the original stellar parameters reported for the object to construct their photometric and spectroscopic synthetic model that was compared to the observations by computing their  $\chi^2$  similarity (`m_sim_p` and `m_sim_f` in Table A1). The resulting fitted spectrum or photometric signature is marked as deviating if its similarity towards observations is worse than

**Table 2.** Explanation of the used binary quality flags in the final classification of the stellar configuration. A raised bit could indicate possible problems or mismatches in the determined configuration. The symbol X in the last two descriptions represents the best-fitting configuration; therefore,  $X = [1,2,3]$ .

Raised bit	Description
0	None of the flags was raised
1st bit	High astrometric uncertainty ( $RUWE > 1.4$ )
2nd bit	Deviating photometric fit ( $sX_{sim.p} > m_{sim.p}$ )
3rd bit	Deviating spectroscopic fit ( $sX_{sim.s} > m_{sim.s}$ )

**Table 3.** Number of different systems discovered by the fitting procedure performed on stars that do not exhibit excess luminosity. Classes and their description are the same as used in Table 1. In addition, the number of stars without parallactic measurements is added for completeness.

Configuration classification	Number of systems
1 star	230
$\geq 1$ star	31
2 stars	4
$\geq 2$ stars	0
3 stars	0
Inconclusive	0
Unknown parallax	0
Total objects	265

for the reported one-star parameters. This might not be the best indication of possible mismatch as it is common that *The Cannon* parameters of the analysed multiple candidates deviate from the main sequence in the Kiel diagram (Fig. 5) and therefore fall into less populated parameter space, where they can skew the single-star models (Sections 4.1 and 4.2).

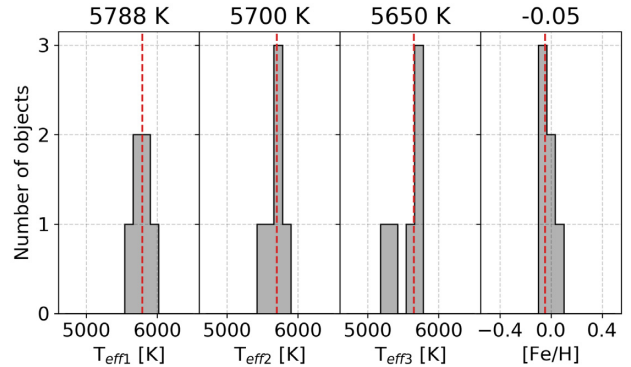
## 6 CHARACTERIZATION OF SINGLE-STAR CANDIDATES

Once we concluded our analysis of the set of 64 objects that showed obvious signs of excess luminosity, we then proceeded to study the remaining 265 objects that are most probably not part of a complex stellar system. To explore their composition, they were analysed with the same procedure as multiple candidates (Section 5). Before running the procedure, we omitted the option to fit for a triple system as they clearly do not possess enough excess luminosity for that kind of a system.

The obtained results were analysed and classified using the same set of rules introduced in Section 5.5. Retrieved classes are summarized in Table 3 and Fig. 7. In the latter we see that the potential binaries are located at the top of the colour–magnitude diagram, which is consistent with the potential presence of an additional stellar source.

## 7 ORBITAL PERIOD CONSTRAINTS

The observational data we have gathered on possible multiple systems point to configurations that change slowly as a function of time. Using the observational constraints and models that describe the formation of the observed spectra, we try to set limiting values



**Figure 8.** Parameters of triple stellar systems discovered and characterized by our analysis. Histograms represent the distribution of all fitted results for six objects that were classified as triple stellar systems and observationally mimic the solar spectrum. Median values of the distributions are given above individual histograms and indicated with a dashed vertical line.

on the orbital parameters of the detected triple stellar system candidates. In order to have a greater sample size, we use both definite (class 3) and probable (class  $\geq 2$ ) triple stars.

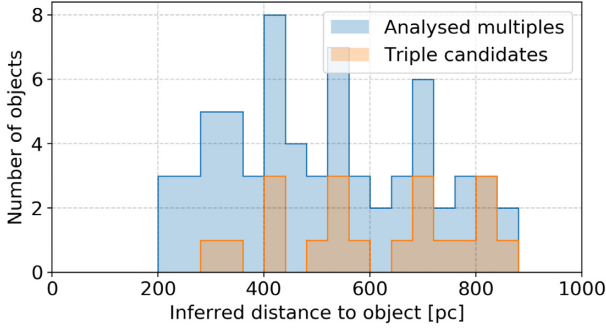
With the limited set of observations, we have to set assumptions about the constitution of those systems. For a hierarchical  $2 + 1$  system to be dynamically stable on long time-scales, its ratio between the orbital period of an inner pair  $P_S$  and outer pair  $P_L$  must be above a certain limit. Eggleton (2006) showed that  $P_L/P_S$  must be higher than 5. The same lower limit is noticeable when the correlation between periods of the known triple stars is plotted (Tokovinin 2008, 2018).

In order to estimate the periods, we have to know the masses and distances between the stars in a system. Without complete information about the projected velocity variation in a system, masses can also be inferred from the spectral type. As we are looking at the solar twin triples, whose effective temperatures are all very similar and close to solar values (see Fig. 8), our rough estimate is that all stars also have a solar-like mass  $\sim M_\odot$ . From this, we can set the inner mass ratio  $q_S$  to be close to 1 and the outer mass ratio  $q_L \sim 0.5$ . When we are dealing with the outermost star, the inner pair is combined into one object with twice the mass of the Sun. The likelihood of such a configuration is also supported by the observations (Tokovinin 2008) where a higher concentration of triple systems is present around those mass ratios. Contrary to our systems, twin binaries with equal masses usually have shorter orbital periods (Tokovinin 2000).

With the initial assumption about the configuration of the triple systems and masses of the stars, the periods of the inner and outer pair can be constrained to some degree as other orbital elements (inclination, ellipticity, phase ...) are not known.

### 7.1 Outer pair and *Gaia* angular resolution

Limited by its design, the *Gaia* spacecraft and its on-ground data processing are in theory unable to resolve stars with angular separation below  $\sim 0.1$  arcsec. Above that limit, their separability is governed by the flux ratio of the pair. The validation report of the latest data release (Arenou et al. 2018) shows that the currently achieved angular resolution is approximately 0.4 arcsec as none of the source pairs is found closer than this separation. We used this reported angular limit to assess possible orbital configurations



**Figure 9.** Histogram showing the distribution of radial distances for the analysed set of stars. Distances were inferred by the Bayesian approach that takes into account the distribution of stars in the Galaxy (Bailer-Jones et al. 2018).

that are consistent with both the spectroscopic and astrometric observations.

Along with the final astrometric parameters, the *Gaia* data set also contains information about the goodness of the astrometric fit. Evans (2018) used those parameters to confirm old and find new candidates for unresolved exoplanet-hosting binaries in the data set. As we are looking at a system of multiple stars that orbit around their common centre of mass, we expect their photo-centre to slightly shift during such orbital motion. The observed wobble of the photo-centre also depends on the mass of stars in the observed system. In the case of a binary system containing two identical stars, the wobble would not be observed, but its prominence would increase as the difference between their luminosities becomes more pronounced. This subtle change in the position of a photo-centre adds additional stellar movement to the astrometric fit, consequently degrading its quality. Improvements for such a motion will be added in later *Gaia* data releases. If a system has an orbital period much longer than the time-span of *Gaia* DR2 measurements (22 months), its movement has not yet affected the astrometric solution. This puts an upper limit on an orbital period as it should not be longer than a few years in order to already affect the astrometric fit results.

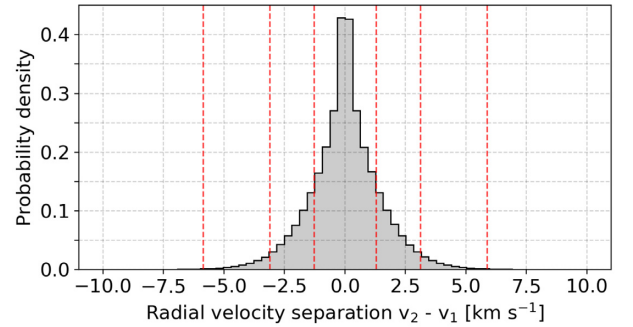
Setting limits to the parameters of astrometric fit quality (`astrometric_excess_noise` > 5 and `astrometric_gof_all` > 20 as proposed by Evans 2018), none of our 329 stars meets those requirements. This suggests that all of them are most likely well below the *Gaia* separability limit and/or have long orbital periods. Another indicator for a lower-quality astrometric fit that we can use is RUWE. Fig. 4 shows the distribution of potentially problematic large RUWE among single and multiple candidates. The latter, on average, have a much poorer fit quality that might indicate a presence of an additional parameter that needs to be considered in the astrometric fit.

Distances to triple stars, shown by the histogram in Fig. 9 range from around 0.3 to 0.9 kpc. From this we can assume the maximal allowable distance between components of an outer pair to be of the order of 100–350 au, pointing to outer orbital periods larger than 500 yr. To test if such systems would meet our detection constraints, we created 100 000 synthetic binary systems whose orbital parameters were uniformly distributed within the parameter ranges given in Table 4. Observable radial velocities of both components were computed using the following equations:

$$v_1 = \frac{2\pi \sin i}{P\sqrt{1-e^2}} \frac{aq}{1+q} (\cos(\theta + \omega) + e \cos \omega) \quad (8)$$

**Table 4.** Ranges of the orbital parameters used for the prediction of observable radial velocity separation between stars in an outer binary pair. The range of the semimajor axis length  $a$  is set between the *Gaia* separability limit for the closest and farthest triple candidate. The uniform distribution of  $a$  is a good approximation of the real periodicity distribution published by Raghavan et al. (2010) as we are sampling a narrow range of it. Use of the real observed distribution would in our case introduce insignificant changes in radial velocity separation as we are simulating wide, slowly rotating systems.

Parameter	Considered range
$M_1$	$2 M_\odot$
$a$	100 ... 350 au
$q$	0.45 ... 0.55
$\sin i$	0 ... 1 ( $i = 0 \dots 90$ deg)
$e$	0.1 ... 0.8
Phase	0 ... 1, used for calculation of $\theta$
$\omega$	0 ... 360 deg

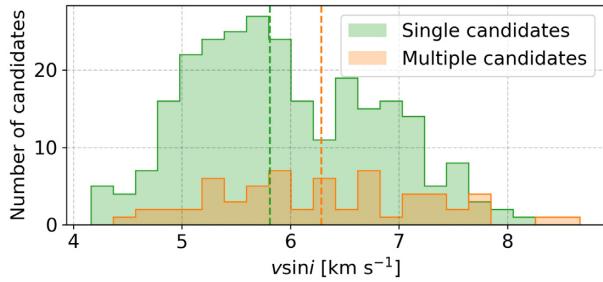


**Figure 10.** Distribution of computed RV separations between a primary and secondary component of the simulated binary systems defined by the orbital parameters given in Table 4. Red vertical lines show 1, 2, and  $3\sigma$  probabilities of the given distribution.

$$v_2 = \frac{2\pi a \sin i}{P\sqrt{1-e^2}} \frac{a}{1+q} (\cos(\theta + \omega + \pi) + e \cos(\omega + \pi)) \quad (9)$$

$$P = \sqrt{a^3 \frac{4\pi^2}{GM_1(1+q)}}. \quad (10)$$

The distribution of velocity separations ( $\Delta v = v_2 - v_1$ ) for synthetic systems is given by Fig. 10, where we can see that more than 99.7 per cent of generated configurations would produce a spectrum that would still be considered as a solar twin ( $\Delta v < 6 \text{ km s}^{-1}$ ; see Section 8.1 and Fig. B3 for further clarification). If we set the semimajor axis  $a$  (in Table 4) to a single value in the same simulation, we can find the closest separation that would still meet the same observational criteria in at least 68 per cent of the cases. In our simulation this happens at the mutual separation of 10 au (and at 50 au for 95 per cent of the cases). The orbital period of such an outer pair is about 18 yr (and 200 yr for 50 au) and is most probably way too long to have a significant effect on the quality of the astrometric fit. Considering observationally favourable orbital configurations with face-on orbits, the actual system could be much more compact than estimated.



**Figure 11.** Distribution of determined  $v\sin i$  for analysed single and multiple candidates. Median values of the distributions are marked with dashed vertical lines.

## 7.2 Inner binary pair and formation of double lines in a spectrum

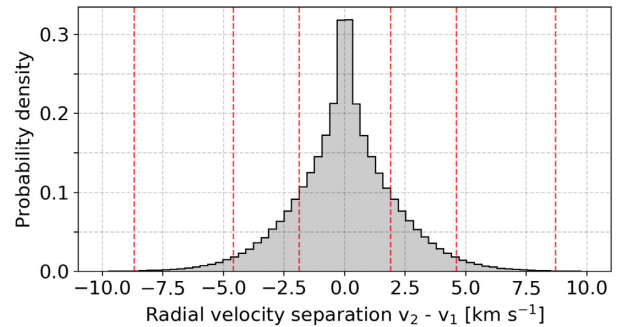
An upper estimate of orbital sizes for an outer pair gives us a confirmation that almost every considered random orbital configuration would satisfy the observational and selection constraints. To ensure the long-term stability of such a system, the inner pair must have a period that is at least five times shorter (Eggleton 2006). At such short orbital periods, the inner stars could potentially move sufficiently rapidly in their orbits as to produce noticeable absorption line splitting for edge-on orbits. To be confident that none of the analysed objects produces such an SB2 spectrum, we visually checked all considered spectra and found no noticeable line splitting in any of the acquired GALAH or Asiago spectra. A subtle hint about a possible broadening of the spectral lines comes from the determined  $v\sin i$ . The median of its distribution in Fig. 11 is higher for multiple candidates with excess projected velocity of  $\sim 0.5 \text{ km s}^{-1}$ .

Accounting for the GALAH resolving power and spectral sampling, we can estimate a minimal RV separation between components of a spectrum to show clear visual signs of duplicated spectral lines. To determine a lower limit, we combined two solar spectra of different flux ratios and visually evaluated when the line splitting becomes easily noticeable. With equally bright sources, this happens at the separation of  $\sim 14 \text{ km s}^{-1}$ . When the secondary component contributes 1/3 of the total flux, the minimal separation is increased to  $\sim 20 \text{ km s}^{-1}$ . A similar separation is needed when the secondary contributes only 10 per cent of the flux. As no line splitting is observed in our analysed spectra, we can be confident that the velocity separation between the binary components was lower than that during the acquisition.

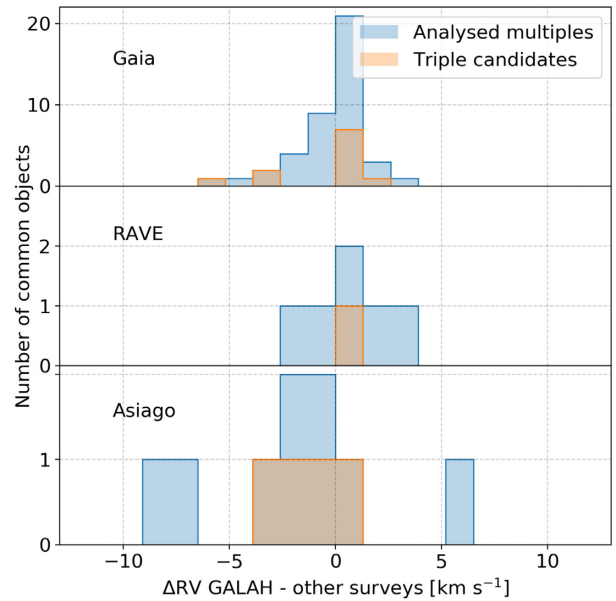
Considering the minimal ratio between the outer and inner binary periods, we can deduce an expected RV separation of an inner pair for the widest possible orbits. As explained in the previous section, we used equations (8)–(10) and possible ranges of inner orbital parameters (Table 5) to generate a set of synthetic binary systems. The distribution of their  $\Delta v$  is shown in Fig. 12 and represents inner binaries with orbital periods from 100 to 700 yr. At those orbital periods, more than 92 per cent of the considered configurations would satisfy the condition of  $\Delta v < 4 \text{ km s}^{-1}$ , ensuring that the observed composite of two equal solar spectra would still be considered as a solar twin (see Section 8.1 and Fig. 14 for further clarification). If we limit our synthetic inner binaries to only one orbital period, we can estimate the most compact system that would still meet the observational criteria in at least 68 per cent of the cases. This would happen at the orbital period of 40 yr with a semimajor axis of 14 au.

**Table 5.** Same as Table 4, but for an inner binary of a hierarchical triple stellar system.

Parameter	Considered range
$M_1$	$M_\odot$
$P_S$	$P_L / 5$ , used for calculation of $a$
$q$	0.9 ... 1.0
$\sin i$	0 ... 1 ( $i = 0 \dots 90 \text{ deg}$ )
$e$	0.1 ... 0.8
Phase	0 ... 1, used for calculation of $\theta$
$\omega$	0 ... 360 deg



**Figure 12.** Same as Fig. 10, but for an inner binary of a simulated triple stellar system.



**Figure 13.** RV difference between GALAH and other large spectroscopic surveys. The name of an individual survey or observatory is given in the upper left corner of every panel. Blue histograms represent the velocity difference for all investigated multiple candidates and orange histograms for discovered triple systems only.

## 7.3 Multi-epoch radial velocities

To support our claims about slowly changing low orbital speeds in detected triple candidates, we analysed changes in their measured radial velocities between GALAH and other comparable all-sky surveys. Distributions of changes are presented by three histograms in Fig. 13 where almost all velocity changes, except one, are

within  $5 \text{ km s}^{-1}$ . Differences between the GALAH and *Gaia* radial velocities were expected to be small as the latter reports median velocities in a time-frame that is similar to the acquisition span of the GALAH spectra. Extending the GALAH observations in the past with RAVE and in the future with Asiago observations did not produce any extreme changes. For this comparison, we also have to consider the uncertainty of the measurements that are of the order of  $\sim 2 \text{ km s}^{-1}$  for RAVE and  $\sim 1 \text{ km s}^{-1}$  for Asiago spectra.

With the data synergies, we could produce only three observational time-series that have observations at more than two sufficiently separated times. With only three data points in each graph, not much can be said about actual orbits.

## 8 SIMULATIONS AND TESTS

To determine the limitations of the employed algorithms, we used them to evaluate a set of synthetic photometric and spectroscopic sources. As the complete procedure depends on the criteria for the selection of solar twin candidates, we first investigated if the selection criteria allow for any broadening of the spectral lines or their multiplicity as they are both signs of a multiple system with components at different projected radial velocities.

In the second part, we generated a set of ideal synthetic systems that were analysed by the same fitting procedure as the observed data set. The results of this analysis are used to determine what kind of systems could be recognized with the fitting procedure and how the results could be used to spot suspicious combinations of fitted parameters.

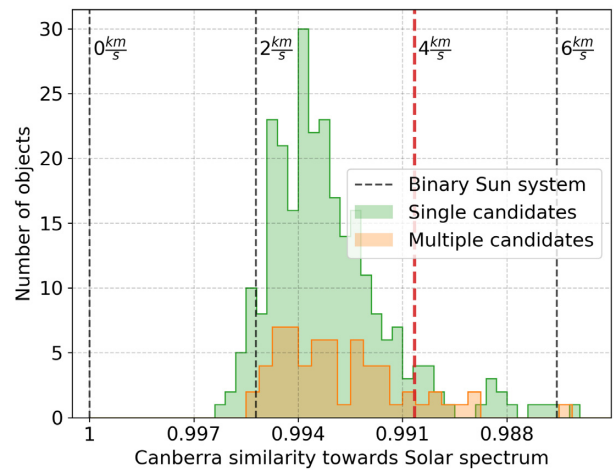
In the final part, we try to evaluate selection biases that might arise from the position of analysed stars on the *Gaia* colour–magnitude diagram and from the GALAH selection function that picks objects based on their apparent magnitude.

### 8.1 Radial velocity separation between components

To determine the minimal detectable radial velocity separation between components in a binary or a triple system, we constructed a synthetic spectrum resembling an observation of multiple Suns at a selected RV separation. The spectrum of a primary component was fixed at the rest wavelength with  $\text{RV} = 0 \text{ km s}^{-1}$  and a secondary spectrum shifted to a selected velocity. After the shift, these two spectra were added together based on their assumed flux ratio.

The generated synthetic spectrum was compared to the solar spectrum with exactly the same metric as described in Section 3. Computed spectral similarities at different separations were compared to the similarities of analysed solar twin candidates. The first RV separation that produces a spectrum that is more degraded than the majority of solar twin candidates was determined to be a minimal RV at which the observed spectrum would be degraded enough that it would no longer be recognized as a solar twin. The high SNR of the generated spectrum was not taken into account for this analysis, in contrast to the algorithm that was used to pinpoint solar twin candidates. Therefore, we also omitted candidates with a lower similarity, which in our case directly corresponds to their low SNR.

The result of this comparison is presented in Fig. 14, where we can see that the minimal detectable separation of two equally bright stars resembling the Sun is  $\sim 4 \text{ km s}^{-1}$ . In the case where a primary star contributes 2/3 of the total flux, the minimal RV increases to  $\sim 6 \text{ km s}^{-1}$  (see Fig. B3). Further increase in the ratio between their fluxes would also increase a minimal detectable separation, but only to a certain threshold from where on a secondary star would not contribute enough flux for it to be detectable, and its received



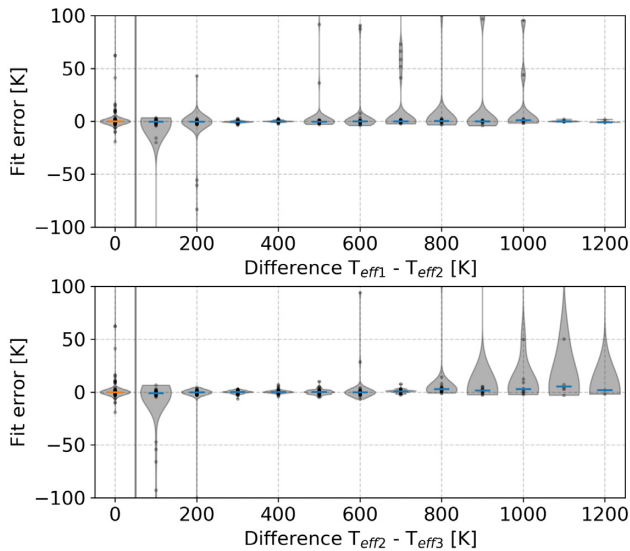
**Figure 14.** Histograms of Canberra similarities towards the solar spectrum for multiple stellar candidates (orange histogram) and single-star candidates (green histogram) in the red HERMES arm. Vertical dashed lines represent the same similarity measure, but for a binary system comprising two equally bright Suns, whose components are at different radial velocities. The separation between components is increasing in  $2 \text{ km s}^{-1}$  steps, where the leftmost vertical line represents the case where both stars are moving with the same projected velocity. The selected maximal velocity at which the composite spectrum would still be considered a solar twin is marked with the thicker red vertical line. The distribution of histograms also shows that spectra of multiple-system candidates are as (dis)similar as spectra of single stars.

flux would be comparable to the typical HERMES spectral noise. In our case, this happens when the secondary contributes less than 10 per cent of the total flux. These boundaries are only indicative as they also heavily depend on the quality of the acquired spectra. When a low level of noise with the Gaussian distribution ( $\sigma = 0.01$ ) is added to a secondary component with a comparable luminosity, the minimal RV decreases because the similarity between spectra also decreases. In that case, the similarity for  $\Delta v = 0$  is located near the mode value of the similarity distribution in Fig. 14.

### 8.2 Analysis of synthetic multiple systems

The limitations and borderline cases of the fitting procedure were tested by evaluating its performance on a set of synthetic binary and triple systems that were generated using the same models and equations as described in the fitting procedure. An ideal, virtually noiseless set comprised 95 binary and 445 triple systems whose components differ in temperature steps of 100 K. The hottest component in the system was set to values in the range between 5300 and 6200 K. The  $T_{\text{eff}}$  of the coldest component could go as low as 4800 K. The  $[\text{Fe}/\text{H}]$  of all synthetic systems was set to 0.0 to mimic solar-like conditions. Condensed results are presented in Figs 15 and B4.

As expected, the fitting procedure (Section 5) did not have any problem determining the correct configuration of the synthesized system. Temperatures of the components were also correctly recovered with a median error of  $0 \pm 13 \text{ K}$  for the hottest component and  $0 \pm 43 \text{ K}$  for the coldest component in a triple system. A more detailed analysis with the distribution of prediction errors, where the temperature difference between components is taken into account, is presented in Fig. 15. From that analysis, we can deduce that  $T_{\text{eff}}$  of the secondary component is successfully retrieved if it does not deviate by more than 1000 K from the primary. Results at



**Figure 15.** The accuracy of our analysis when the fitting procedure is applied to a synthetic triple system. The upper panel shows the distribution of  $T_{\text{eff}}$  prediction errors for a secondary star depending on the difference between selected temperatures of a primary and a secondary star. The lower panel shows the same relation but for a tertiary star in comparison to a secondary. As a reference, the prediction error of a primary star is shown on the left side of both panels. Labels  $T_{\text{eff}1}$ ,  $T_{\text{eff}2}$ , and  $T_{\text{eff}3}$  indicate decreasing effective temperatures of stars in a simulated system.

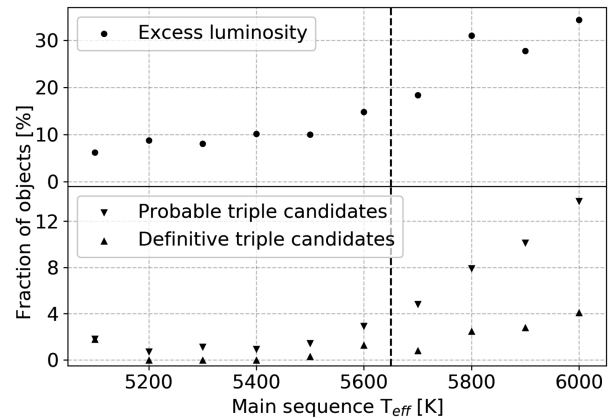
such large temperature differences are inconclusive as the number of simulated systems drops rapidly. The same can be said for the tertiary component, but with the limitation that it should not be colder by more than 700 K when compared to the secondary star. Beyond that point, the uncertainty of the fitted result increases and a star is determined to be hotter than it really is.

Another application of such an analysis is to identify signs that could point to a possibly faulty solution when it is comparably likely that a multiple system is comprised of two or three components. The results of such an analysis are presented in Fig. B4, where we tried to describe a binary system with a triple system fit. From the distribution of errors, we can observe that the effective temperature of a primary star with the largest flux is recovered with the smallest fit errors whose median is  $20 \pm 93$  K. As we are using too many components in that fit, the  $T_{\text{eff}}$  of a secondary is reduced in order to account for the redundant tertiary component in the fit. As imposed by the limit in the fitting procedure, the tertiary  $T_{\text{eff}3}$  is set as low as possible. If the same thing happens for a real observed system, this could be interpreted as a model overfitting.

### 8.3 Triple stars across the H–R diagram

In the current stage of Galactic evolution, binary stars with a solar-like  $T_{\text{eff}}$  are located near a region that is also occupied by main-sequence turnoff stars in the Kiel and colour–magnitude diagrams (Figs 5 and 7). This, combined with the fact that older stars with a comparable initial mass and metallicity would also pass a region occupied with binaries (Fig. 4), poses an additional challenge for the detection of unresolved multiples if their spectrum does not change sufficiently during the evolution.

To analyse the possible influence of the turnoff stars and older, more evolved stars, we ran the same detection procedure as described in Section 3 and analysis (Section 5) to determine



**Figure 16.** The upper panel shows the percentage of objects at different positions on the Kiel diagram shown in Fig. 5 that show excessive luminosity and are spectroscopically similar to main-sequence stars. Similarly, the lower panel shows the upper and lower boundary on a percentage of triple system candidates at the same positions. For a definitive candidate, both fits must agree on a triple configuration. The strong dashed vertical line represents a point on the Kiel diagram where the main sequence visually starts merging with the red-giant branch, the point where a region above the main sequence becomes polluted with more evolved stars.

the fraction of binaries and triples at different  $T_{\text{eff}}$ , ranging from 5100 to 6000 K, with a step of 100 K. A comparison median spectrum was computed from all spectra in the range  $\Delta T_{\text{eff}} \pm 60$  K,  $\Delta \log g \pm 0.05$  dex, and  $\Delta[\text{Fe}/\text{H}] \pm 0.05$  dex, where the main-sequence  $\log g$  was taken from the main-sequence curve shown in Fig. 5, and  $[\text{Fe}/\text{H}]$  was set to 0.0. The limiting threshold for multiple candidates was automatically determined in the same way as for solar twin candidates (Section 3.1).

Condensed results of the analysis are shown in Fig. 16, where we can see that both the fraction of stars with excessive luminosity and triple candidates starts increasing above  $T_{\text{eff}} \sim 5600$  K. This might indicate that the underlying distribution of more evolved and/or hotter stars might have some effect on our selection function. On the other hand, this increasing binarity fraction coincides with other surveys that show similar trends (Duchêne & Kraus 2013).

As the detected triple-star candidates encompass a fairly small region in a colour–magnitude space, we were interested in the degree to which our analysed spectra are similar to those of other stars in the same region. For this purpose, the GALAH objects with absolute *Gaia* corrected magnitudes in the range  $3.3 < M_G < 3.6$  and  $0.77 < G_{BP} - G_{RP} < 0.84$  were selected and plotted in Fig. B5, where spectra are arranged according to their similarity. In this 2D projection (details about the construction of which are given in Traven et al. 2017 and Buder et al. 2018) it is obvious that spectra considered in this paper clearly exhibit a far greater degree of mutual similarity than other spectra with a similar photometric signature. As expected, many of them lie inside the region of SB2 spectra. All of our solar twin candidates were visually checked for the presence of a resolvable binary component. Never the less, this plot gives us additional proof of their absence as none of the analysed twins is located inside that region.

### 8.4 Observational bias – GALAXIA model

Every magnitude-limited survey, such as GALAH, will introduce observational bias into the frequency of observed binary stars as their additional flux changes the volume of the Galaxy from where

they are sampled. Their distances and occupied volume of space are located further from the Sun and are therefore also greater than those for comparable single stars with the same apparent magnitude and colour.

To evaluate this bias for our type of analysed stars, we created a synthetic Milky Way population of single stars using the GALAXIA code (Sharma et al. 2011). First, the code was run to create a complete population of stars in the apparent  $V$  magnitude range between 10 and 16 without any colour cuts. In order for the distribution of synthesized stars to mimic the GALAH survey as closely as possible, only stars located inside the observed GALAH  $2^\circ$  fields were retained. After the spatial filtering, only stars with Sun-like absolute magnitude  $M_V = 4.81 \pm 0.05$  and colour index  $B - V = 0.63 \pm 0.05$  were kept in the set (reference magnitudes were taken from Willmer 2018).

From the filtered synthetic data set we took three different sub-sets representing a pool of observable single Solar twins ( $12.0 \leq V \leq 14.0$ ), binary twins ( $12.75 \leq V \leq 14.75$ ), and triple twins ( $13.2 \leq V \leq 15.2$ ) based on their apparent  $V$  magnitudes. Extinction and reddening were not used in this selection as their use did not significantly alter the relationship between the number of stars in sub-sets. Assuming that the frequency of multiple stars is constant and does not change in any of those sets, we can estimate the selection bias on the frequency of multiple stars. The number of stars in each sub-set was 15 862, 35 470, and 54 567, respectively. According to those star counts, the derived frequency of binaries would be too high by a factor of 2.2 and a factor 3.4 for triple stars. Considering those factors, the fraction of unresolved triple candidates with solar-like spectra is  $\sim 2$  per cent and  $\sim 11$  per cent for binary candidates.

Accidental visual binaries that lie along the same line of sight and have angular separation smaller than the field of view of the 2dF spectroscopic fiber (2arcsec) or smaller than the *Gaia* end-of-mission angular resolution (0.1arcsec) were not considered in this estimation.

## 9 CONCLUSIONS

Combining multiple photometric systems with spectroscopic data from the GALAH and astrometric measurements taken by *Gaia*, we showed the possible existence of triple stellar systems with long orbital periods whose combined spectrum mimics the solar spectrum. The average composition of such a system consists of three almost identical stars, where one of the stars is  $\sim 10$  K warmer than the Sun and the coldest has an effective temperature  $\sim 120$  K below the Sun. The derived percentage of such unresolved systems would be different for nearby/close stars as they become more/less spatially resolved. In the scope of our magnitude-limited survey, we sampled only a fraction of possible distances to the systems.

Without any obvious signs of the orbital periodicity in the measured radial velocities of the systems, orbital periods were loosely constrained based on the observational limits and a few assumptions. By the prior assumption that the *Gaia* spacecraft sees those systems as a single light source, we showed that they can be described by orbital periods where a difference between projected velocities of components does not sufficiently degrade an observed spectrum that it is no longer recognized as solar-like. The spectroscopic signature and radial velocity variations were further used to put a limit on the minimum orbital period of an inner pair to be at least 20 yr. Shorter periods are not completely excluded as it could happen that the spectrum was acquired in a specific orbital phase where the difference between projected orbital

speeds is negligible. From the fact that analysed objects are spatially unresolvable for the *Gaia* spacecraft, the orbital size of outer binary pair can extend up to 100–350 au and therefore have orbital periods of the order of a few hundred years.

To confirm their existence, detected systems are ideal candidates to be observed with precise interferometric measurements or high-time-resolution photometers if they happen to be occulted by the Moon. Simulation of the lunar motion showed that four of the analysed multiple candidates lie in its path if the observations are carried out from the Asiago Observatory, which has suitable photon-counting detectors.

The main drawback of the analysis was found to be its separate treatment of photometric and spectroscopic information in two independent fitting procedures. In future analyses, they should be combined to acquire even more precise results as different stellar physical parameters have a different degree of impact on these two types of measurements.

## ACKNOWLEDGEMENTS

This work is based on data acquired through the Australian Astronomical Observatory, under programmes: A/2014A/25, A/2015A/19, A2017A/18 (the GALAH survey); A/2015A/03, A/2015B/19, A/2016A/22, A/2016B/12, A/2017A/14 (the K2-HERMES K2-follow-up programme); A/2016B/10 (the HERMES-TESS programme); A/2015B/01 (Accurate physical parameters of *Kepler* K2 planet search targets); S/2015A/012 (planets in clusters with K2). 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.

KČ, TZ, and JK acknowledge financial support of the Slovenian Research Agency (research core funding No. P1-0188 and project N1-0040).

This research was partly supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013.

YST is supported by the NASA Hubble Fellowship grant HST-HF2-51425.001 awarded by the Space Telescope Science Institute.

Follow-up observations were collected at the Copernico telescope (Asiago, Italy) of the INAF – Osservatorio Astronomico di Padova.

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

## REFERENCES

- Andrae R. et al., 2018, *A&A*, 616, A8
- Andrews J. J., Chanamé J., Agüeros M. A., 2017, *MNRAS*, 472, 675
- Arenou F. et al., 2018, *A&A*, 616, A17
- Badenes C. et al., 2018, *ApJ*, 854, 147
- Bailer-Jones C. A. L., Rybizki J., Fousneau M., Mantelet G., Andrae R., 2018, *AJ*, 156, 58
- Barden S. C. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 773509
- Bergmann C., Endl M., Hearnshaw J. B., Wittenmyer R. A., Wright D. J., 2015, *Int. J. Astrobiol.*, 14, 173
- Buder S. et al., 2018, *MNRAS*, 478, 4513
- Capitanio L., Lallement R., Vergely J. L., Elyajouri M., Monreal-Ibero A., 2017, *A&A*, 606, A65

Casagrande L., VandenBerg D. A., 2018, *MNRAS*, 479, L102  
 Close L. M., Richer H. B., Crabtree D. R., 1990, *AJ*, 100, 1968  
 Coronado J., Sepúlveda M. P., Gould A., Chanamé J., 2018, *MNRAS*, 480, 4302  
 De Silva G. M. et al., 2015, *MNRAS*, 449, 2604  
 Duchêne G., Kraus A., 2013, *ARA&A*, 51, 269  
 Duquennoy A., Mayor M., 1991, *A&A*, 248, 485  
 Eggleton P., 2006, *Evolutionary Processes in Binary and Multiple Stars*. Cambridge Astrophysics, Cambridge University Press  
 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  
 Evans D. F., 2018, *RNAAS*, 2, 20  
 Evans D. W. et al., 2018, *A&A*, 616, A4  
 Fernandez M. A. et al., 2017, *PASP*, 129, 084201  
 Ferraro F. R., Carretta E., Fusi Pecci F., Zamboni A., 1997, *A&A*, 327, 598  
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306  
 Gaia Collaboration, 2016, *A&A*, 595, A1  
 Gaia Collaboration, 2018, *A&A*, 616, A1  
 Garnavich P. M., 1988, *ApJ*, 335, L47  
 Gould A., Bahcall J. N., Maoz D., Yanny B., 1995, *ApJ*, 441, 200  
 Henden A. A., Levine S., Terrell D., Welch D. L., 2015, *AAS Meeting Abstracts #225*, p. 336.16  
 Hurley J., Tout C. A., 1998, *MNRAS*, 300, 977  
 Jiménez-Esteban F. M., Solano E., Rodrigo C., 2019, *AJ*, 157, 78  
 Kamdar H., Conroy C., Ting Y.-S., Bonaca A., Johnson B., Cargile P., 2019a, preprint (arXiv:1902.10719)  
 Kamdar H., Conroy C., Ting Y.-S., Bonaca A., Smith M., Brown A. G. A., 2019b, preprint (arXiv:1904.02159)  
 Kos J. et al., 2017, *MNRAS*, 464, 1259  
 Kraus A. L., Hillenbrand L. A., 2009, *ApJ*, 703, 1511  
 Kunder A. et al., 2017, *AJ*, 153, 75  
 Lance G. N., Williams W. T., 1967, *Aust. Comput. J.*, 1, 15  
 Lindegren L., 2018, 'Re-normalising the Astrometric Chi-square in Gaia DR2', Gaia Technical Note: GAIA-C3-TN-LU-LL-124-0  
 Marigo P. et al., 2017, *ApJ*, 835, 77  
 Matijević G. et al., 2010, *AJ*, 140, 184  
 Matijević G. et al., 2011, *AJ*, 141, 200  
 Merle T. et al., 2017, *A&A*, 608, A95  
 Milone A. P. et al., 2016, *MNRAS*, 455, 3009  
 Moe M., Di Stefano R., 2017, *ApJS*, 230, 15  
 Ness M., Hogg D. W., Rix H.-W., Ho A. Y. Q., Zasowski G., 2015, *ApJ*, 808, 16  
 Nordström B. et al., 2004, *A&A*, 418, 989  
 Oh S., Price-Whelan A. M., Hogg D. W., Morton T. D., Spergel D. N., 2017, *AJ*, 153, 257  
 Pourbaix D. et al., 2004, *A&A*, 424, 727  
 Quist C. F., Lindegren L., 2000, *A&A*, 361, 770  
 Raghavan D. et al., 2010, *ApJS*, 190, 1  
 Rebassa-Mansergas A., Gänsicke B. T., Rodríguez-Gil P., Schreiber M. R., Koester D., 2007, *MNRAS*, 382, 1377  
 Rebassa-Mansergas A., Nebot Gómez-Morán A., Schreiber M. R., Gänsicke B. T., Schwöpe A., Gallardo J., Koester D., 2012, *MNRAS*, 419, 806  
 Rebassa-Mansergas A., Ren J. J., Parsons S. G., Gänsicke B. T., Schreiber M. R., García-Berro E., Liu X.-W., Koester D., 2016, *MNRAS*, 458, 3808  
 Ren J., Luo A., Li Y., Wei P., Zhao J., Zhao Y., Song Y., Zhao G., 2013, *AJ*, 146, 82  
 Schlafly E. F., Finkbeiner D. P., 2011, *ApJ*, 737, 103  
 Sharma S., Bland-Hawthorn J., Johnston K. V., Binney J., 2011, *ApJ*, 730, 3  
 Sharma S. et al., 2018, *MNRAS*, 473, 2004  
 Shaya E. J., Olling R. P., 2011, *ApJS*, 192, 2  
 Sheinis A. et al., 2015, *J. Astron. Telesc. Instrum. Sys.*, 1, 035002  
 Skopal A., 2005, *A&A*, 440, 995  
 Skrutskie M. F. et al., 2006, *AJ*, 131, 1163  
 Tokovinin A., 2008, *MNRAS*, 389, 925

Tokovinin A., 2014, *AJ*, 147, 87  
 Tokovinin A., 2018, *ApJS*, 235, 6  
 Tokovinin A. A., 2000, *A&A*, 360, 997  
 Traven G. et al., 2017, *ApJS*, 228, 24  
 Troup N. W. et al., 2016, *AJ*, 151, 85  
 Widmark A., Leistedt B., Hogg D. W., 2018, *ApJ*, 857, 114  
 Willmer C. N. A., 2018, *ApJS*, 236, 47  
 Wittenmyer R. A. et al., 2018, *AJ*, 155, 84  
 Wright E. L. et al., 2010, *AJ*, 140, 1868  
 Zwitter T. et al., 2018, *MNRAS*, 481, 645  
 Žerjal M. et al., 2019, *MNRAS*, 484, 4591

## SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

### solar\_like\_multiples.csv

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.

## APPENDIX A: TABLE DESCRIPTION AND SUMMARY

In Table A1 we provide a list of metadata available for every object detected using the methodology described in this paper. The complete table of detected objects and their metadata is available in electronic form at the CDS. An excerpt of the table, containing a sub-set of columns, for definitive and probable triple candidates is given in Table A2.

**Table A1.** List and description of the fields in the published catalogue of analysed objects.

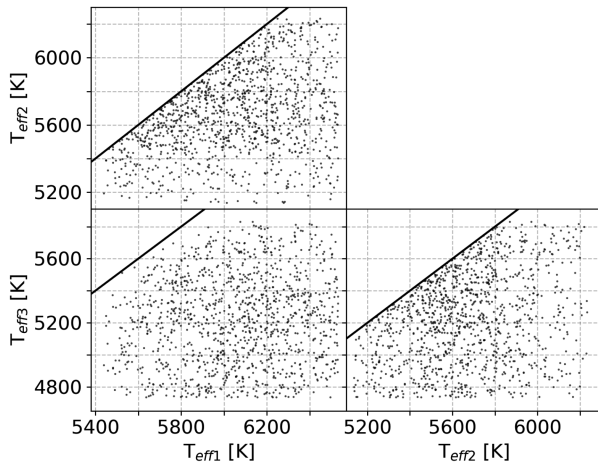
Field	Unit	Description
source_id		<i>Gaia</i> DR2 source identifier
subject_id		Unique internal per-observation star ID
ra	deg	Right ascension from <i>Gaia</i> DR2
dec	deg	Declination from <i>Gaia</i> DR2
ruwe		Value of renormalized astrometric $\chi^2$
m_sim_p		Photometric $\chi^2$ for original parameters
m_sim_f		Spectroscopic $\chi^2$ for original parameters
s1_teff1	K	$T_{\text{eff}1}$ in a fitted single system
s1_feh		[Fe/H] of a fitted single system
s1_sim_p		Photometric $\chi^2$ of a fitted single system
s1_sim_f		Spectroscopic $\chi^2$ of a fitted single system
s2_teff1	K	$T_{\text{eff}1}$ in a fitted binary system
s2_teff2	K	$T_{\text{eff}2}$ in a fitted binary system
s2_feh		[Fe/H] of a fitted binary system
s2_sim_p		Photometric $\chi^2$ of a fitted binary system
s2_sim_f		Spectroscopic $\chi^2$ of a fitted binary system
s3_teff1	K	$T_{\text{eff}1}$ in a fitted triple system
s3_teff2	K	$T_{\text{eff}2}$ in a fitted triple system
s3_teff3	K	$T_{\text{eff}3}$ in a fitted triple system
s3_feh		[Fe/H] of a fitted triple system
s3_sim_p		Photometric $\chi^2$ of a fitted triple system
s3_sim_f		Spectroscopic $\chi^2$ of a fitted triple system
n_stars_p		Best-fitting photometric configuration
n_stars_f		Best-fitting spectroscopic configuration
class		Final configuration classification
flag		Result quality flags

**Table A2.** Subset of results for definitive and probable solar-like triple candidates detected by our selection and fitting procedure. The complete table is given as a supplementary material to this paper in the form of a textual CSV file. It is also available in electronic form at the CDS portal.

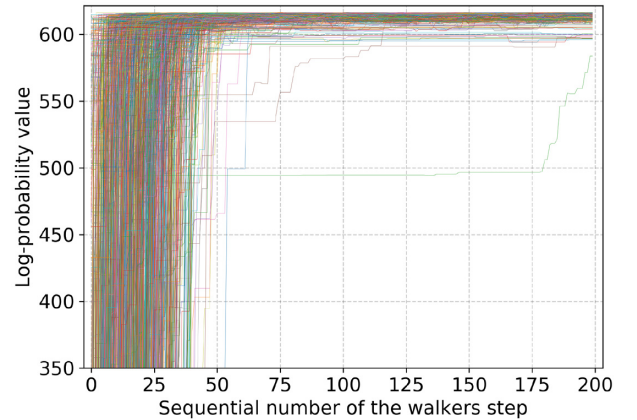
Source_id	RUWE	s2_teff1	s2_teff2	s2_feh	s3_teff1	s3_teff2	s3_teff3	s3_feh	Class	Flag
6157059919188478720	11.1	6002	6000	0.24	5825	5787	5780	0.03	3	1
6777339306532222080	8.0	5875	5819	0.10	5642	5636	5631	-0.06	3	1
6412815502155127808	1.1	5880	5316	-0.02	5922	4703	4701	0.03	>2	4
6564302331580491904	1.1	5991	4701	0.13	5945	4702	4700	0.02	>2	0
5386113598793714304	1.5	5882	5808	0.07	5878	5438	5313	-0.03	3	5
2534579880633620992	1.2	5986	5876	0.15	5750	5739	5714	-0.06	3	6
5484353352124904064	1.1	5876	5802	0.14	5844	5525	5425	-0.00	>2	0
5399712362903401984	1.3	5968	4703	0.08	5922	4703	4701	0.01	>2	4
3688523450018482432	1.1	5820	5688	-0.015	5998	4717	4702	0.09	>2	2
6220408320279116032	7.6	5923	5883	0.11	5724	5677	5669	-0.08	3	1
6198738457927949440	0.9	6052	4798	0.15	5986	4704	4701	0.10	>2	4
5822222074090352384	4.7	5651	5640	-0.04	5866	4702	4701	0.01	>2	1
6355462192511955456	1.0	5700	5692	-0.11	5471	5449	5379	-0.28	>2	4
4678229218054642176	0.9	6028	4791	0.09	5968	4702	4700	0.09	>2	4
4423111085550775680	1.6	5829	5774	-0.02	5645	5633	5398	-0.11	>2	1
6261736621608044416	4.7	5927	5890	0.07	6154	4702	4701	0.20	>2	1
5947219160131475712	1.1	6071	5843	0.22	6007	5724	5265	0.10	3	4
6661888382295654656	1.5	5725	5678	-0.07	5956	4706	4701	0.04	>2	1
6362136502970853120	1.8	5844	5826	0.013	5666	5658	5540	-0.18	>2	1
6645693508028329728	1.3	5936	4715	0.06	5833	4703	4701	-0.02	>2	4

## APPENDIX B: ADDITIONAL FIGURES

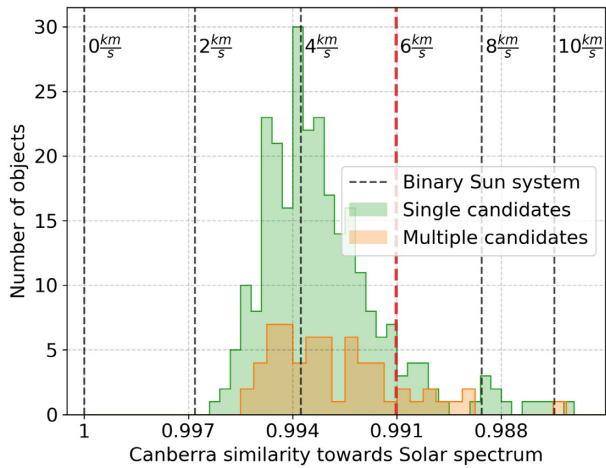
In order to increase the readability and transparency of the text, additional and repeated plots are supplied as appendices to the main text.



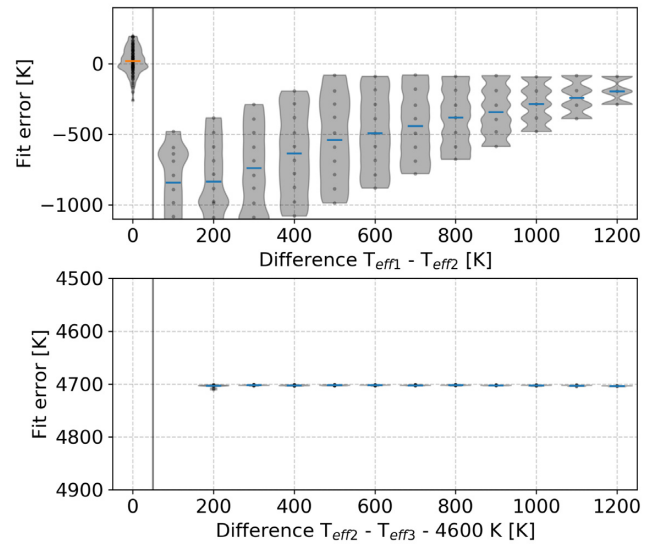
**Figure B1.** Initial distribution of walker parameters considered in the photometric fit. To ensure unique solutions with increasing  $T_{\text{eff}}$ , combinations above the linear line were not considered in the fit.



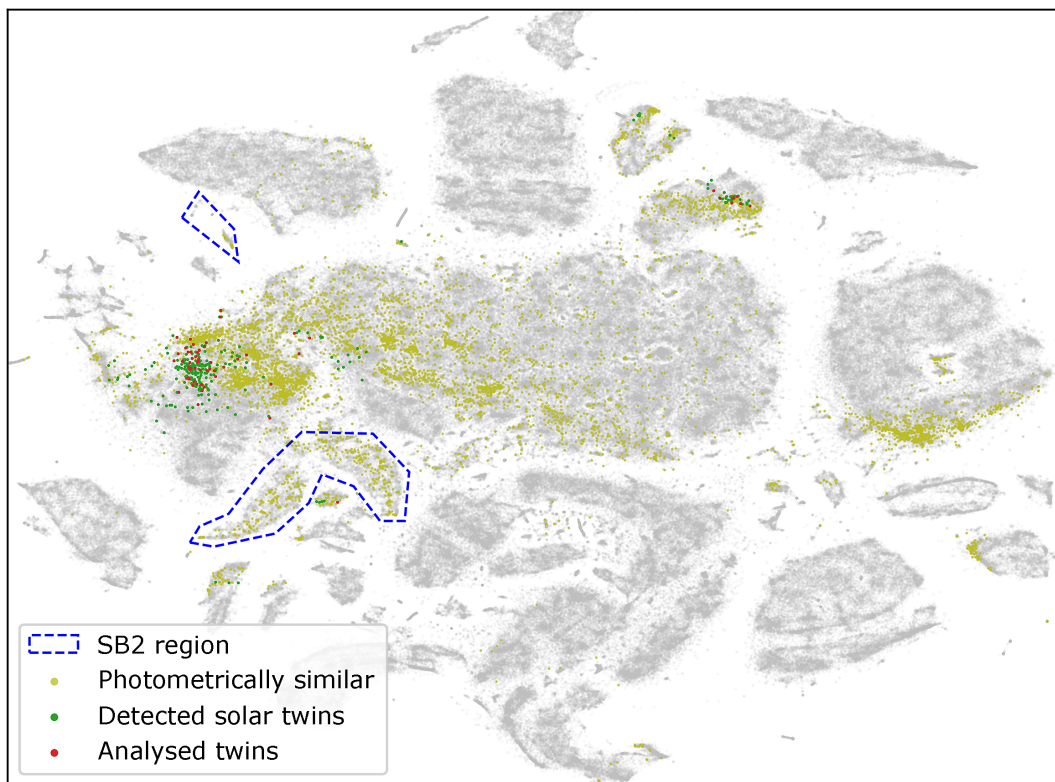
**Figure B2.** Convergence of walkers in the initial photometric fit. The plot shows the log-probability of the posteriors shown in Fig. 6. Walkers converge to the same value of log-probability after 50 steps. Every walker is plotted with a different colour.



**Figure B3.** Same as Fig. 14 but for the case of a triple system where only one component out of three has a radial velocity shift in comparison to the other two.



**Figure B4.** Similar to a plot in Fig. 15, but representing a case when we try to describe a synthetic binary system with  $T_{\text{eff}3} = 0$  using a triple star model.



**Figure B5.** Visual representation of similarities between spectra using dimensionality reduction analysis. Clumps in this 2D projection represent morphologically similar spectra, whose features separate them from the rest of the data set. Blue dashed polygons indicate overdensities, where SB2 spectra are located (projection and regions are taken from fig. 13 in Buder et al. 2018). Each grey dot represents one spectrum of the GALAH survey. Green coloured dots represent objects considered in this paper, of which objects marked with red were analysed for higher-order multiplicity. Yellowish dots represent objects whose *Gaia* magnitudes fall inside the range of determined triple-star candidates.

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.