**Aguirre, VS, Bojsen-Hansen, M, Slumstrup, D, Casagrande, L, Kawata, D, Ciuca, I, Handberg, R, Lund, MN, Mosumgaard, JR, Huber, D, Johnson, JA, Pinsonneault, MH, Serenelli, AM, Stello, D, Tayar, J, Bird, JC, Cassisi, S, Hon, M, Martig, M, Nissen, PE, Rix, HW, Schönrich, R, Sahlholdt, C, Trick, WH and Yu, J**

**Confirming chemical clocks: asteroseismic age dissection of the Milky Way disk(s)**

Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 16 January 2018 (MN LATEX style file v2.2)

# Confirming chemical clocks: asteroseismic age dissection of the Milky Way disk(s)

V. Silva Aguirre[1], M. Bojsen-Hansen[1], D. Slumstrup[1], L. Casagrande[2], D. Kawata[3], I. Ciucă[3], R. Handberg[1], M. N. Lund[4,1], J. R. Mosumgaard[1], D. Huber[5,6,1], J. A. Johnson[7,8], M. H. Pinsonneault[7], A. M. Serenelli[9], D. Stello[10,6,1], J. Tayar[7], J. C. Bird[11], S. Cassisi [12], M. Hon[10], M. Martig[13], P. E. Nissen[1], H. W. Rix[14], R. Schönrich[15], C. Sahlholdt[1], W. H. Trick[14], and J. Yu[6,1]

[1]*Stellar Astrophysics Centre, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus C, Denmark*
[2]*Research School of Astronomy & Astrophysics, Mount Stromlo Observatory, The Australian National University, ACT 2611, Australia*
[3]*Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT, UK*
[4]*School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK*
[5]*Institute for Astronomy, University of Hawaii, 2680 Wood- lawn Drive, Honolulu, HI 96822, USA*
[6]*Sydney Institute for Astronomy (SIfA), School of Physics, University of Sydney, NSW 2006, Australia*
[7]*Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA*
[8]*Center for Cosmology and AstroParticle Physics, 191 West Woodruff Ave, Ohio State University, Columbus, OH 43210*
[9]*Instituto de Ciencias del Espacio (ICE-CSIC/IEEC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Cerdanyola del Valles, Spain*
[10]*School of Physics, University of New South Wales, NSW 2052, Australia*
[11]*Department of Physics and Astronomy, Vanderbilt University, 6301 Stevenson Circle, Nashville, TN 37235, USA*
[12]*INAF - Astronomical Observatory of Teramo, via M. Maggini, sn. Teramo, Italy*
[13]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[14]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[15]*Rudolf-Peierls Centre for Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK*



16 January 2018

**ABSTRACT**

Investigations of the origin and evolution of the Milky Way disk have long relied on chemical and kinematic identification of its components to reconstruct our Galactic past. Difficulties in determining precise stellar ages have restricted most studies to small samples, normally confined to the solar neighbourhood. Here we break this impasse with the help of asteroseismic inference and perform a chronology of the evolution of the disk throughout the age of the Galaxy. We chemically dissect the Milky Way disk population using a sample of red giant stars spanning out to 2 kpc in the solar annulus observed by the *Kepler* satellite, with the added dimension of asteroseismic ages. Our results reveal a clear difference in age between the low- and high-$\alpha$ populations, which also show distinct velocity dispersions in the $V$ and $W$ components. We find no tight correlation between age and metallicity nor [$\alpha$/Fe] for the high-$\alpha$ disk stars. Our results indicate that this component formed over a period of more than 2 Gyr with a wide range of [M/H] and [$\alpha$/Fe] independent of time. Our findings show that the kinematic properties of young $\alpha$-rich stars are consistent with the rest of the high-$\alpha$ population and different from the low-$\alpha$ stars of similar age, rendering support to their origin being old stars that went through a mass transfer or stellar merger event, making them appear younger, instead of migration of truly young stars formed close to the Galactic bar.

**Key words:** Galaxy: disc — Galaxy: evolution — Galaxy: structure — Asteroseismology — stars: fundamental parameters — stars: kinematics and dynamics

## 1 INTRODUCTION

Spiral galaxies such as ours contain several populations of stars comprising their bulge, disk, and halo, all with different chemical and kinematic properties capturing unique epochs of formation and the different processes that led to their specific characteristics. The disk is the defining stellar component of the Milky Way, and understanding its formation has been identified as one of the most important goals of galaxy formation theory (e.g., Freeman & Bland-Hawthorn 2002; Rix & Bovy 2013; Bland-Hawthorn & Gerhard 2016).

The Milky Way disk has been geometrically separated into a thin and a thick component that dominate at different heights, and were identified using stellar counts more than 30 years ago (Gilmore & Reid 1983; Jurić et al. 2008). It is assumed that the formation history and timescale of these populations are different, and therefore stars from each geometric component of the disk should be associated with a particular chemical, kinematic, and age signature. One of the aims of Galactic archaeology is constructing the evolution history of the disk from these fossil records, which requires an accurate characterisation of the properties of stars belonging to each component.

Due to the difficulties in determining ages for faint field stars based purely on spectroscopic or photometric information, most studies of the Milky Way disk have focused on identifying different populations in the solar neighbourhood using chemistry and kinematics (e.g., Fuhrmann 1998; Soubiran, Bienaymé & Siebert 2003; Bensby, Feltzing & Lundström 2003; Navarro et al. 2011; Ramírez, Allende Prieto & Lambert 2013; Rojas-Arriagada et al. 2016; Allende Prieto, Kawata & Cropper 2016, just to name a few). It is expected that the geometric thin and thick disks can be separated using space velocities as the thick component should be kinematically hotter than the thin disk population, and given their different timescales of formation they should present different trends in $\alpha$-abundance patterns that would make them also identifiable in chemical space. However, neither the chemical nor the kinematic criterion seem to clearly separate the disk components because of the substantial overlap between the phase-space distribution of stars belonging to the high- and low-$\alpha$ abundance sequences (e.g., Feltzing & Bensby 2008; Schönrich & Binney 2009b; Adibekyan et al. 2011; Bensby, Feltzing & Oey 2014). Indeed, the radial and vertical structure of chemically selected disk components varies intrinsically with abundance, especially among low-$\alpha$ stars (Bovy, Rix & Hogg 2012; Bovy et al. 2016). High-resolution spectroscopic surveys suggest a bimodal distribution of $\alpha$-abundances in disk stars (e.g., Anders et al. 2014; Hayden et al. 2015), although the very existence of a chemically distinct disk has been challenged based on chemo-dynamical models and low-resolution spectroscopic observations (Schönrich & Binney 2009b; Loebman et al. 2011).

The reason for the mixing and contamination between chemical and kinematic samples is probably related to stars in the solar neighbourhood being born at different galactocentric radii and migrating to their current position by dynamical processes such as radial migration (e.g., Sellwood & Binney 2002; Roškar et al. 2008; Schönrich & Binney 2009a; Minchev, Chiappini & Martig 2013; Grand, Kawata & Cropper 2015). Viable formation scenarios for the Milky Way disk include inside-out and upside-down formation (e.g., Matteucci & Francois 1989; Brook et al. 2006; Bird et al. 2013; Ness et al. 2016; Schönrich & McMillan 2017), although there are suggestions that the old component of the disk formed outside-in instead (Robin et al. 2014). A natural step forward to gain further insight about these processes would be to also identify stellar populations within the thin and thick disk based on the chronology of formation events, effectively dissecting the disk by adding the age dimension.

Using a very local ($\sim$ 25 pc) but volume complete sample, Fuhrmann (1998, 2011) suggested that a chemical dissection of disk would also result in a clean age separation of the components: the high-$\alpha$ sequence should be occupied by stars older than about 10 Gyr while the low-$\alpha$ population is expected to be composed by stars younger than about 8 Gyr. A transition between the $\alpha$-abundances sequences should occur at 10 Gyr, and the majority of the thick disk stars are expected to have an age of about 12 Gyr. This age gap still needs to be confirmed, and the advent of asteroseismology as a tool for Galactic archaeology promises to test this paradigm by determining precise masses and ages for thousands of stars in distant regions of the Milky Way (e.g., Miglio et al. 2013b; Casagrande et al. 2014, 2016; Anders et al. 2017; Rodrigues et al. 2017).

Among the efforts combining ground-based follow-up observations and oscillations data, the APOKASC catalogue (Pinsonneault et al. 2014) comprises a sample of almost 2000 red giant stars observed by APOGEE (Majewski et al. 2017) that have asteroseismic detections in their frequency power spectrum obtained with the *Kepler* satellite. Martig et al. (2016) showed that the carbon and nitrogen abundances in these giants, altered in their photospheres during the first dredge-up, could be calibrated using the masses derived from asteroseismology to provide a spectroscopic measurement of (implied) ages with a precision of $\sim$ 0.2 dex and an accuracy below 0.1 dex. Coupled to data driven approaches such as *The Cannon* (Ness et al. 2015) they allow us to map the age distribution of thousands of red giants spanning up to $\sim$ 5 kpc (Ness et al. 2016; Ho et al. 2017) and have showed that the high-$\alpha$ stars are predominately old (and much older than low-$\alpha$ stars) throughout the disk.

While methods like *The Cannon* are capable of extending the results of asteroseismology to stars that have no oscillations data, the precision in their derived ages is significantly lower ($\sim$ 0.2 dex) than that obtained by asteroseismic analysis of red giants (e.g., Casagrande et al. 2016; Anders et al. 2017). In this paper we re-examine the APOKASC catalogue using the most up-to-date asteroseismic observations and introduce the age dimension in chemical studies of the Galactic disk. The combination of asteroseismic inference and spectroscopic information allows us to determine precise stellar properties as well as kinematic and chemical patterns of stars, and to chemically dissect the Milky Way disk including age information in a sample spanning distances up to 3 kpc from the Sun.

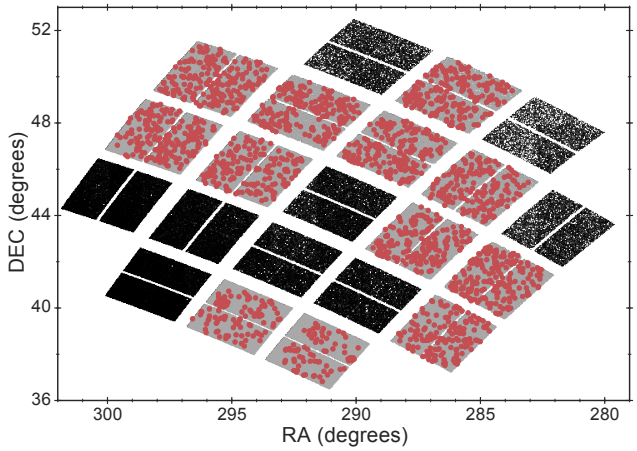


**Figure 1.** Position in the sky of the *Kepler* field of view. Stars with asteroseismic detections in the APOKASC14 sample are shown in red, while grey (black) symbols depict all stars from the 2MASS catalogue falling in the same (other) CCD of the *Kepler* spacecraft.

## 2 THE APOKASC SAMPLE

### 2.1 Photometric, spectroscopic, and asteroseismic data

The set of stars considered in this study are the 1989 red giants comprising the first combined APOGEE and *Kepler* catalogue, in which the *Kepler* spacecraft detected oscillations during its nominal mission. The position of the sample in the *Kepler* field of view is shown in Fig. 1. From the *Kepler* Input Catalogue (KIC Brown et al. 2011; Huber et al. 2014) we retrieved observations for all our targets in the *griz* filters, which we converted to standard Sloan magnitudes using the transformation derived by Pinsonneault et al. (2012). Infrared observations in the $JHK_S$ bands were extracted from the 2MASS Point Source Catalogue (Skrutskie et al. 2006).

The original APOKASC compilation presented by Pinsonneault et al. (2014) contained asteroseismic information based on observations from the nominal *Kepler* mission spanning quarters Q0 to Q8 (two years of data). In total, *Kepler* acquired 17 quarters of data after four years of observations which are now available for analysis. Thus, we used the updated global seismic parameters $\langle\Delta\nu\rangle$ (average large frequency separation) and $\nu_{\rm max}$ (frequency of maximum power) from Yu et al. (2017, in preparation). Among the many breakthroughs of space-based asteroseismology is the possibility of discriminating between stars in their first ascent up the red giant branch and those that have already ignited helium in their core based solely on their observed pulsation spectra (Bedding et al. 2011; Mosser et al. 2012). We adopt the revised evolutionary classifications from Hon, Stello & Yu (2017)[1], Vrard, Mosser & Samadi (2016), and Stello et al. (2013).

We use spectroscopic parameters from the 13th data release (DR13) of the Sloan Sky Digital Survey (SDSS, SDSS et al. 2016), including metallicity, $\alpha$-abundances, effective temperatures, and heliocentric radial velocities. The uncertainties in metallicity reported in this compilation correspond to the internal precision and are of the order of $\sim$0.03 dex. In an effort to better estimate systematic uncertainties, we added in quadrature the median difference between APOGEE results in clusters and the standard literature values as reported in Table 3 of Tayar et al. (2017) (corresponding to 0.09 dex). Atmospheric properties were available for 1984 of the targets, and we further neglected 5 targets with metallicities consistent with [Fe/H] $< -1$, as the reliability of asteroseismic determination of stellar masses in the metal-poor regime could be slightly biased towards too high values (see e.g., Epstein et al. 2014; Miglio et al. 2016, for a discussion).

### 2.2 Determination of stellar properties

We determined stellar properties for the APOKASC sample combining the photometric, spectroscopic, and asteroseismic observables using the BAyesian STellar Algorithm (`BASTA`, Silva Aguirre et al. 2015, 2017). The procedure for deriving all quantities is divided into two steps, which we describe in the following.

In the first step we determine all physical properties of the stars in the sample such as mass, radius, luminosity, and age, using the so-called *grid-based* method. In its essence, this approach compares all observed quantities with predictions from theoretical models of stellar evolution. Thus, the only requirement to apply the *grid-based* method is a set of tracks or isochrones containing all observed quantities and covering the necessary parameter space. We use a set of BaSTI isochrones (Pietrinferni et al. 2004) including the effects of overshoot in the main sequence and semiconvection in the core He-burning phase that have been extended in metallicity coverage especially for asteroseismic studies (see Silva Aguirre et al. 2013, for a description).

The two atmospheric properties we fit are the effective temperature $T_{\rm eff}$ and bulk metallicity [M/H]. The latter is determined from the DR13 measurements of [Fe/H] and [$\alpha$/Fe] following the prescription of Salaris, Chieffi & Straniero (1993). We complement this information with the global asteroseismic parameters $\langle\Delta\nu\rangle$ and $\nu_{\rm max}$, that must first be determined in our set of isochrones. To do this we consider the asteroseismic scaling relations (Ulrich 1986; Brown et al. 1991), which can be written as:

$$\left(\frac{\langle\Delta\nu\rangle}{\langle\Delta\nu_\odot\rangle}\right)^2 \simeq \frac{\bar{\rho}}{\bar{\rho}_\odot}\,, \tag{1}$$

$$\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}} \simeq \frac{M}{\rm M_\odot}\left(\frac{R}{\rm R_\odot}\right)^{-2}\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{-1/2}\,. \tag{2}$$

Here $\langle\Delta\nu_\odot\rangle$ and $\nu_{\rm max,\odot}$ are the solar reference values and they depend on the asteroseismic pipeline used to extract the pulsation information from the *Kepler* light-curves. For our adopted set of asteroseismic observables they correspond to $\langle\Delta\nu_\odot\rangle = 135.1$ $\mu$Hz and $\nu_{\rm max,\odot} = 3090$ $\mu$Hz (Huber et al. 2011).

Equations 1 and 2 can be used to estimate the theoretical values of $\langle\Delta\nu\rangle$ and $\nu_{\rm max}$ for any point along an evolu-

[1] Here we adopt updated results from the technique by Hon, Stello & Yu (2017) based on a new and improved neural network model (Hon et al. in preparation)

tionary track or isochrone. Testing the accuracy of the asteroseismic scaling relations is currently a very active field of research, and there is strong evidence of a metallicity, effective temperature, and evolutionary phase-dependent offset in the large frequency separation relation (cf. Eq. 1, see White et al. 2011; Sharma et al. 2016; Guggenberger et al. 2016). In `BASTA` we apply the correction by Serenelli (2017, in preparation) to this equation, which has been shown to reproduce a number of classical age determinations (e.g., turn-off fitting, eclipsing binaries, white dwarf cooling curve) in the open clusters M67 (Stello et al. 2016) and NGC 6819 (Casagrande et al. 2016). Several tests carried out on the scaling relations suggest they are accurate to a level of a few percent (see e.g., Huber et al. 2012; Silva Aguirre et al. 2012; White et al. 2013; Miglio et al. 2013a, 2016; Gaulme et al. 2016; Huber et al. 2017).

When the evolutionary phase of each target is known from the asteroseismic analysis of the power spectrum we impose it in the analysis as a Bayesian prior, otherwise all possible stages of evolution are taken into account in `BASTA` when constructing the probability density function for determining the stellar properties. This results in unclassified targets normally having larger statistical uncertainties than RGB or clump stars (see section 2.4 below), but they represent a small fraction of our sample and therefore do not significantly affect the conclusions of our study.

The current version of `BASTA` includes a feature to compute distances based on bolometric corrections from observed magnitudes in different bandpasses and an extinction map, following the procedure outlined by Rodrigues et al. (2014). Thus, the second step in our *grid-based* method consists of using the surface gravity determined by `BASTA` and combining it with the reddening value from the KIC catalogue, the $grizJHK_S$ magnitudes, and the spectroscopic $T_{\rm eff}$ and [Fe/H] to compute the bolometric correction using the synthetic photometry of Casagrande & Vandenberg (2014). We then determine the distance modulus in each of the seven bandpasses and use the median value to extract the bolometric magnitude as described by Torres (2010), and obtain the distance to the star. From the derived distance and coordinates of the target, we recompute the extinction value using the Green et al. (2015) reddening map. This updated interstellar extinction is combined again with the surface gravity, photometric and spectroscopic measurements to obtain a new bolometric correction, and therefore a new distance estimate. The process is repeated until extinction changes by less than 0.001 mag between iterations and thus convergence is reached (normally only two iterations are needed). Our results are in excellent agreement with those from other asteroseismic pipelines used to determine distances such as the Bellaterra Stellar Properties Pipeline (Serenelli et al. 2013) and the PARAM code (Da Silva et al. 2006; Rodrigues et al. 2014). A comparison between the estimates from these pipelines can be found in Fig. 2 of Huber et al. (2017).

### 2.3 Sample completeness

We describe in the following the pruning procedure applied to this sample according to data availability and quality. For the 1979 stars with stellar properties determined, we collected proper motions from the first data release of Gaia (Lindegren et al. 2016; Gaia Collaboration et al. 2016) or the UCAC-4 catalogue (Zacharias et al. 2013) according to availability. The latter were further pruned following the quality check procedure described by Anders et al. (2014) (see their section 3.2), yielding a total of 1838 stars with reliable astrometric solutions. Using our asteroseismically determined distances we followed Johnson & Soderblom (1987); Bensby, Feltzing & Lundström (2003) to determine the (right-handed) Galactic space-velocity components $UVW$ in the Local Standard of Rest as defined by Schönrich, Binney & Dehnen (2010). We performed Gaussian fits to determine the standard deviation in each velocity component of the full sample, and as a further data-quality control we selected only the targets with uncertainties in their individual velocities below half of the computed standard deviation, resulting in a sample of 1593 stars with precise kinematic information available. Finally, we removed 3 stars for which our derived stellar properties resulted in fractional age uncertainties larger than 1.

Our sample of 1590 stars selected in this manner is a subset of the more than 15,000 oscillating red giants detected by *Kepler*, and the fact that this set and not a different one is available for our study is the result of a combination of criteria largely based on plate availability for observations with the APOGEE telescope. Thus, there is no reason to believe *a priori* that the stars comprising the APOKASC sample (or our pruned sub-selection of it) are representative of the physical and kinematic characteristics of the true underlying population of red giants in that direction of the sky. Any inference directly drawn from this sample regarding e.g. its age or composition distribution carries the risk of being biased unless we properly correct for the selection function. To gauge this effect, we follow Casagrande et al. (2016) and take a two-step approach where we first correct for the selection of oscillating giants with available APOGEE spectra, and after that for the target selection effects of the *Kepler* spacecraft as a function of distance (see section 3 below).

To assess the potential biases of the selected sample of 1590 targets, we need to compare their intrinsic properties to an unbiased set of stars in the same field of view. Fortunately, the *Kepler* field has available 2MASS photometry that we can use to quantify these effects. The main panel in Fig. 2 shows a colour-magnitude diagram of all stars in the 2MASS catalogue falling within the same CCDs of *Kepler*'s field of view as our targets (c.f Fig. 1). We repeated the procedure outlined in Casagrande et al. (2016) to find the magnitude and colour ranges where our sample of stars and the underlying 2MASS population are indistinguishable from each other. The resulting cumulative distributions are shown in the right panels of Fig. 2, where we have removed stars and applied the K-Sample Anderson-Darling test until the null hypothesis that both samples are drawn from the same parent distribution cannot be rejected with a significance level greater than 1% for both $J-K_{\rm S}$ and $K_{\rm S}$. This statistical test is similar to the commonly applied Kolmogorov-Smirnoff test but much more sensitive at the edges of the distribution. The procedure leaves us with a final set of 1197 stars that are fully representative in colour and magnitude of the underlying stellar population in the *Kepler* line of sight (see Table 1). The updated evolutionary classifications for this sample used in this paper yield a total of 422 stars in the RGB and 751 in the clump phase,

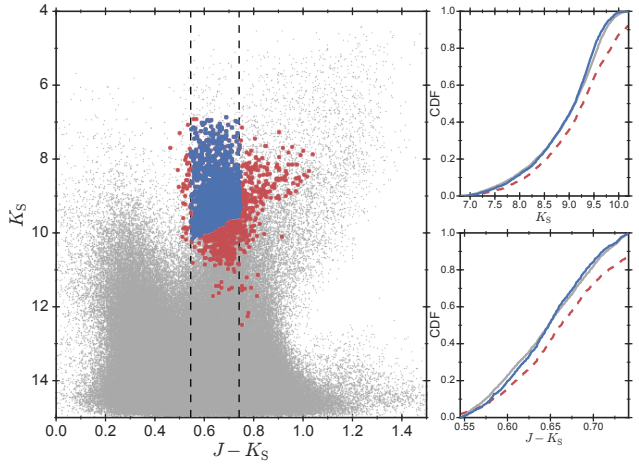


**Figure 2.** Selection function for a photometrically complete sample. *Left*: colour-magnitude diagram of all stars from 2MASS (grey) falling within the same CCD's as the APOKASC sample (shown in red). The stars fulfilling the photometric completeness criteria are shown in blue. *Right*: cumulative distributions in magnitude and colour of the 2MASS sample (grey solid), the APOKASC stars (red dashed), and the pruned complete sample (blue solid). See text for details.

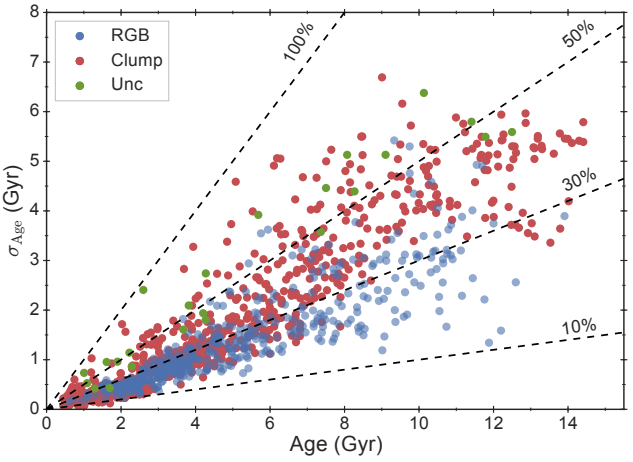


**Figure 3.** Final age uncertainties obtained for the colour-magnitude complete sample, colour-coded according to their asteroseismic evolutionary classification. Dashed lines mark the 10%, 30%, 50%, and 100% fractional uncertainty levels.

**Table 1.** Number of stars considered in this study according to data availability and quality control. See text for details.

| APOKASC Catalogue | |
|---|---|
| Total number of stars | 1989 |
| `BASTA` stellar properties | 1979 |
| Available proper motions | 1838 |
| Galactic velocities cut | 1593 |
| Seismic uncertainties cut | 1590 |
| Colour-magnitude complete sample | 1197 |
| • RGB | 422 |
| • Clump | 751 |
| • Unclassified | 24 |

while it is not possible to assign an evolutionary status to the remaining 24 giants.

### 2.4 Final uncertainties

The procedure outlined in section 2.2 results in a set of stellar properties and their associated errors that takes into account the observational uncertainties but does not yet consider systematics based on the input physics of our stellar models. With this in mind, we derived a second set of properties for our sample with `BASTA` using the same input parameters but a grid of stellar models including the effects of mass-loss. The latter was implemented following the Reimers prescription (Reimers 1977) with an efficiency of $\eta = 0.4$, representing the extreme of the commonly adopted values around $\eta = 0.1 - 0.2$ supported by various observations (e.g., Miglio et al. 2012; Origlia et al. 2014; Miglio et al. 2016). This allows us to determine a conservative estimate of the age uncertainty introduced to our results by mass-loss.

To determine the final uncertainties in the stellar properties we first consider results from our standard grid of models that includes overshooting, semiconvection, and no mass-loss. In `BASTA` we obtain the posterior probability density function and we add in quadrature the 16 and 84 percentiles to determine the statistical uncertainty in our derived stellar properties. We add in quadrature to these uncertainties a systematic component calculated as half the difference between the standard values and those obtained with the grid with very efficient mass-loss. After including this effect our resulting stellar properties have median uncertainties of the order of 2.7% (radius), 3.3% (distance), 7.0% (mass), and 28.5% (age), in line with those obtained by e.g., Casagrande et al. (2014, 2016); Anders et al. (2017); Rodrigues et al. (2017) using similar approaches.

The age uncertainties for our sample are plotted in Fig. 3 where it can be seen that, as age increases, clump stars have larger fractional uncertainties than RGB targets of the same age. This is the result of including mass-loss in our calculations, which in the Reimers prescription is most efficient for low-mass stars and towards the tip of the RGB. The unclassified stars in the sample have in general larger uncertainties than their clump and RGB age counterparts as a consequence of the flat Bayesian prior used for their evolutionary state (as mentioned in section 2.2).

Figure 4 depicts the resulting uncertainties in the Galactic velocity components for our complete sample of stars. The kernel density estimates (KDE) show that the combination of small uncertainties in our derived distances and the applied kinematic pruning criteria (c.f section 2.3) results in **precise** space velocities. We note that the uncertainties in the $V$-velocity component are smaller than those in $U$ and $W$ because the *Kepler* field lies approximately in the direction of the $V$-velocity component; therefore its uncertainty is mainly determined by the precision in the radial velocity measurements from spectroscopy. On the other hand, the

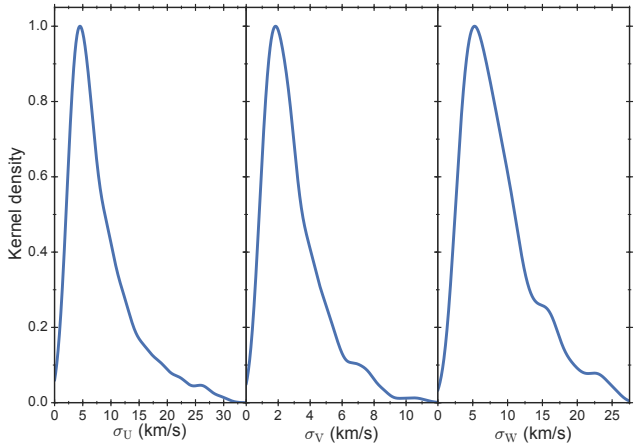


**Figure 4.** Uncertainties in the Galactic velocities derived for the sample. The kernel density estimates have been computed using a Gaussian kernel and scott's rule for determining the bandwidth.

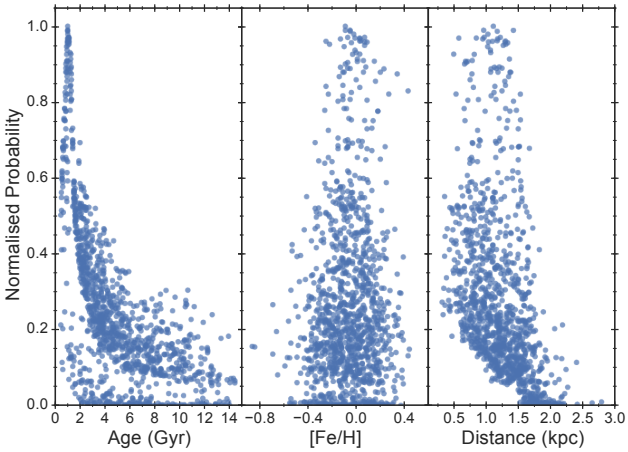


**Figure 5.** Normalised probability of stars being observed by the *Kepler* satellite at different ages, metallicities, and distances.

absolute errors in $U$ and $W$ are mostly dominated by the uncertainties in proper motions.

## 3 CORRECTING FOR TARGET SELECTION EFFECTS

Our colour and magnitude complete sample of giants comprises metallicities between $-1.0 <$ [Fe/H] $< 0.5$, ages from $\sim$ 0.4 to $\sim$ 14 Gyr, gravities within $3.3 < \log g < 2.0$, and distances from $\sim$ 300 to $\sim$ 3000 pc. There are many thousands of red giant stars in the *Kepler* field of view with stellar properties falling within these ranges but only 1197 of the ones in the subsample studied here have detections of oscillations reported. As a result, our oscillating targets represent only a fraction of the stars that could have pulsations detected by *Kepler* at a given combination of age, metallicity, and distance, and therefore it is possible that each star in our photometrically complete sample is either over-or-underrepresented.

To take into account the stellar population biases due to our particular target selection applied in section 2.3, we follow the approach presented by Casagrande et al. (2016), who corrected for the *Kepler* target selection bias generating a multidimensional synthetic data cube. The underlying idea is that at a given distance, there is a ratio between the number of stars where oscillations are detected and the total number of stars where oscillations could have been detected if all stars would have been selected by the *Kepler* science team for observations. This ratio varies as a function of distance because the observed magnitude of stars changes, making them fall in or out of our derived colour-magnitude completeness region defined in section 2.3. Thus, a high value of the ratio represents a high fraction of stars being selected for observations with *Kepler* were oscillations can be detected.

We created a multi-dimensional data cube in age and metallicity using the set of BaSTI isochrones with overshoot and no mass-loss, and randomly populated it by generating synthetic stellar samples from a Salpeter initial mass function. We assigned apparent magnitudes to each synthetic star by running over the distance dimension. Once the probabilities for each synthetic target are calculated as the ratio between the number of stars satisfying our completeness criteria and the total number of stars at a given distance, the probabilities for our actual targets are determined using interpolation along the 3-axis of the data cube. The most underrepresented stars have probabilities consistent with zero, which we then determine by taking the median of 10,000 Monte Carlo realisations of the interpolation assuming a gaussian distribution of their uncertainties in age, metallicity, and distance.

Figure 5 shows the probabilities resulting from our target selection procedure as a function of age, metallicity, and distance. Stars with lower probability are disfavoured by the *Kepler* target selection, while those with high probability values are the most likely to have oscillations detected. The resulting trends can be understood in terms of the stellar properties governing the detectability of pulsations, in particular the oscillation amplitudes that scale proportionally to the stellar luminosity (Kjeldsen & Bedding 1995). For example, it is clear from Fig. 5 that young stars are favoured over old stars. The reason is that at a given metallicity and distance more massive (younger) stars are intrinsically brighter than their less massive (older) counterparts, resulting in an overrepresentation of young stars in our sample. Similarly, stars with the lowest probabilities are scattered across the age and metallicity range but restricted to distanced beyond $\sim$ 1.5 kpc, as detecting the intrinsic variations in brightness characterising the oscillations is harder for the furthest away targets.

The procedure outlined above allows us to gauge the effect of the *Kepler* target selection in our results, and we correct for this effect in all the following results by weighting each star in our fits and distributions by the inverse of the obtained probability and its corresponding uncertainty. In practise this means that we resample our stars a number of times proportional to the logarithm of their inverse individual probability. Before closing this section, we note that different approaches for determining this correction such as population synthesis and galaxy modelling were explored by Casagrande et al. (2016), who found consistent results across

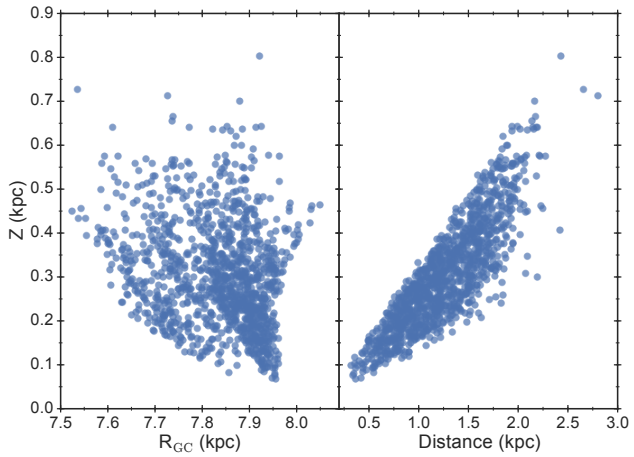


**Figure 6.** Position in the Galaxy of the complete sample determined assuming a solar position of $R_{GC,\odot} = 8$ kpc. *Left:* Galactocentric radii as a function of height above the plane. *Right:* distance from the Sun. See text for details.

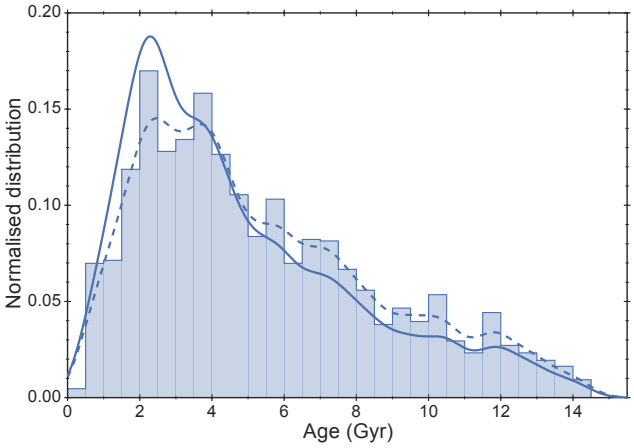


**Figure 7.** Age histogram of the colour-magnitude complete sample (shaded region). Dashed curve depicts the unweighted kernel density estimate of the distribution calculated with a Gaussian kernel, while the solid curve takes into account the age uncertainties as weights. All distributions are normalised such as the integral over the range is unity. See text for details.

all methods. Similarly we have tested the impact of changes in the initial mass function by adopting a uniform distribution in masses instead of a Salpeter IMF when populating the isochrones, and found no significant difference in the distribution of probabilities (see appendix B). Our choice of the data cube for correcting target selection effects is guided by having the least model-dependent approach, but we emphasise that adopting a different methodology does not have a significant impact in our derived results.

## 4 SEPARATING DISK COMPONENTS

After the data pruning, sample completeness, and target selection effects procedures described in sections 2 and 3, we have a set of 1197 red giants with precise physical, chemical, and kinematic properties which is representative of the population of giants in the *Kepler* line-of-sight. We now turn our attention into the analysis of the properties of this sample and its implications for the formation and evolution of the Galactic disk.

The positions in the Galaxy of the stars in our complete sample are shown in Fig. 6. Due to the location of the *Kepler* field at Galactic longitude of $l \simeq 74^\circ$ and the pencil-beam shape nature of the survey, all targets are roughly located at the same galactocentric radii as the Sun $R_{GC,\odot} \simeq 8$ kpc, giving us a fully representative population of red giants in the solar annulus. Our targets also span distances up to $\sim 2$ kpc and almost 1 kpc above the plane, probing much further than commonly studied samples in e.g.,. the solar neighbourhood which are volume-complete to 25 pc (Fuhrmann 2011) or $\sim 40$ pc (Nordström et al. 2004; Casagrande et al. 2011).

Figure 7 shows the age histogram of the sample constructed with equally spaced bins of 0.5 Gyr. Over plotted with a dashed curve is the Gaussian kernel density estimate of the distribution without considering the individual uncertainties in age, which neatly reproduces the main features of the underlying data distribution. When the individual fractional uncertainties in age are included in the computation of the KDE, the bimodal nature of the distribution favours a single highest peak at $\sim 2$ Gyr and shows additional structure at larger ages. The optimal bandwidth in each KDE have been determined using the Sheather-Jones method that minimises the asymptotic mean integrated squared error (see e.g., Sheather & Jones 1991; Venables & Ripley 2002, and references therein).

To investigate if a classical chemical selection of the disk components results in distinct age distributions, we select stars in the low- and high-$\alpha$ sequences following Adibekyan et al. (2011). We divide our sample into six metallicity bins in the range $-1.0 < [M/H] < 0.50$, and after identifying the minima in the histograms of [$\alpha$/Fe] distribution at each metallicity bin we draw the separation line for the components by joining the value of the minima at each bin. The result is two populations of 1030 (low-$\alpha$) and 167 (high-$\alpha$) members each, corresponding to fractions of $\sim 86\%$ and $\sim 14\%$, respectively.

The kinematic and chemical properties of the sample selected in this manner are shown in Fig. 8, where we also mark the position of the young $\alpha$-rich (Y$\alpha$R) stars in our sample. This population was originally identified by Chiappini et al. (2015) and Martig et al. (2015) as a peculiar class of stars with high $\alpha$-abundances and young ages, features not predicted by standard chemical evolution models of the Galaxy. In particular, Martig et al. (2015) identified 14 stars in the APOKASC sample with [$\alpha$/Fe] $> 0.13$ dex and a maximum possible age younger than 6 Gyr. Only 8 of these 14 targets are contained in our colour and magnitude complete sample, and our `BASTA` results confirm their apparent young nature with derived ages below 4 Gyr (consistent with the original definition of this population by Martig et al. (2015)). However, the DR13 values used in our analysis are slightly different from the DR12 spectroscopic results considered by Martig et al. (2015), and two of these 8 stars have $\alpha$-abundances slightly below the threshold of 0.13 dex (at the [$\alpha$/Fe] $= 0.09$ and [$\alpha$/Fe] $= 0.10$ level). Considering the quoted [$\alpha$/Fe] uncertainty of $\sim 0.03$ dex in

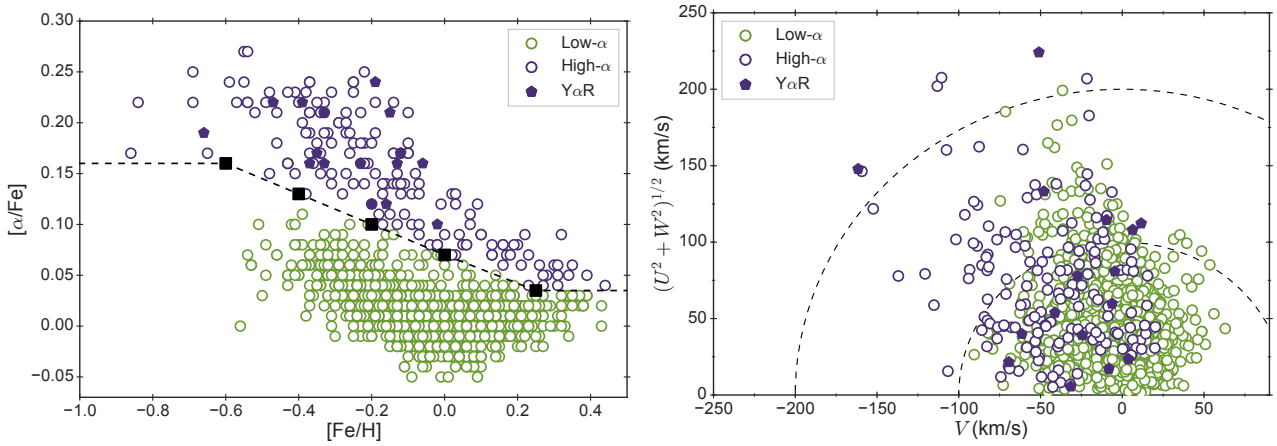


**Figure 8.** Disk components selected based on chemistry. *Left*: $\alpha$-abundances as a function of metallicity. Dashed lines and black squares show the separation line between low-$\alpha$ and high-$\alpha$ disk. *Right*: Toomre diagram for the chemically selected components of the disk. See text for details.

DR13, we define the young $\alpha$-rich population in our sample as those stars with [$\alpha$/Fe] >= 0.10 dex and age below 5 Gyr, obtaining a total of 16 stars fulfilling these criteria in our colour-magnitude complete sample[2].

In a similar result as found by Adibekyan et al. (2012); Bensby, Feltzing & Oey (2014), Fig. 8 shows that our chemically selected high-$\alpha$ and low-$\alpha$ sequences overlap drastically in kinematic space. Analogously, a separation based on their Galactic velocities leads to the opposite behaviour of overlapping in the chemical plane. We verified this by also selecting disk components based only on kinematic information following Soubiran, Bienaymé & Siebert (2003) and using a Gaussian Mixture Model (GMM). The rationale behind this approach is to separate the population in a finite number of Gaussian distributions based on a set of observables (which we chose to be the Galactic velocity components $U$, $V$, and $W$), and then select the number of populations according to the Bayesian Information Criterion (BIC). The model favours two Gaussian components with 1036 and 161 members respectively (fractions $\sim$ 87% and $\sim$ 13%), which overlap in the [M/H] versus [$\alpha$/Fe] plane as expected but separate relatively well in the Toomre diagram. Nevertheless, given the probabilistic nature of the GMM some stars assigned to one Gaussian component have similar kinematics to the other Gaussian sample, making the separation between stellar populations not as clear as it is in chemical space. For this reason, we continue our analysis based on the chemical separation and explore the age dimension as an additional piece of information that can fully disentangle the components of the disk.

### 4.1 The age dimension

In Fig. 9 we show the age distributions of the sample selected according to the chemistry. The resulting KDE show that the low-$\alpha$ sequence peaks at $\sim$ 2 Gyr while the high-$\alpha$ disk does it at $\sim$ 11 Gyr, confirming with asteroseismology that the chemical criterion dissects the disk into two distinct populations in terms of age (as shown by e.g., Fuhrmann 1998, 2011, in the solar neighbourhood). We see what appears to be contamination of old stars in the low-$\alpha$ component and young stars in the high-$\alpha$ sequence, which we further explore in the following.

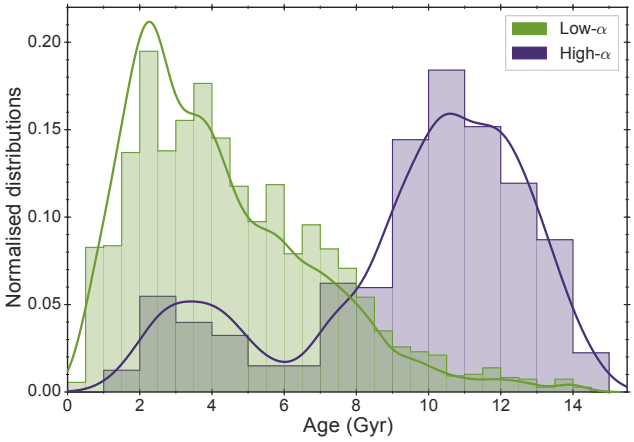


**Figure 9.** Age distributions for the low and high-$\alpha$ disk components. The histograms are constructed using equally spaced bins of 0.5 Gyr (low-$\alpha$) and 1 Gyr (high-$\alpha$), while the solid lines represent the Gaussian KDE computed with the individual fractional uncertainties as weights and the Sheather-Jones method to determine bandwidths.

The left panel in Fig. 10 shows the relation between age and [$\alpha$/Fe], where it can be seen that the young $\alpha$-rich stars account for the majority of the targets in the peak at $\sim$ 3 Gyr seen in the high-$\alpha$ disk distribution (c.f. Fig. 9). Understanding the origin of these stars has been the subject of a number of recent studies and they have been attributed to migrators from the Galactic bar (Chiappini et al. 2015) as well as evolved blue stragglers (Martig et al. 2015; Chiappini et al. 2015; Yong et al. 2016; Jofre et al. 2016). In the former case it is believed that these stars formed in reservoirs of al-

[2] We have verified that a more strict cut in the selection of these young $\alpha$-rich stars does not affect our conclusions in sections 4.1 and 4.2 regarding their origin.

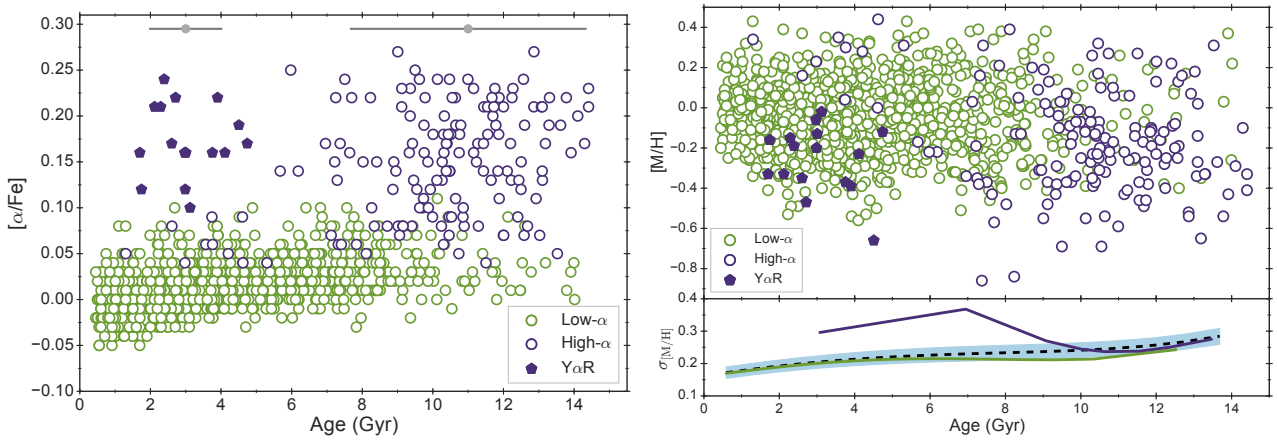


**Figure 10.** Relations between age and chemistry for the sample of giants. *Left*: $\alpha$-abundances as a function of age for the chemically separated components of the disk. Grey symbols indicate median age uncertainties at 3 Gyr and 11 Gyr for the high-$\alpha$ sample. *Right*: Age-metallicity relation of the low-$\alpha$, and high-$\alpha$ disk populations. The bottom panel shows the metallicity dispersion determined from a bootstrap fit to the full sample (black dashed line) and its corresponding uncertainty (shaded region), as well as the high-$\alpha$ (solid violet) and low-$\alpha$ (solid green) components. See text for details.

most inert gas close to the end of the Galactic bar, while the latter scenario proposes that the young $\alpha$-rich stars are the product of mass transfer or stellar merger events. We further explore these scenarios using the kinematics and dynamics of Y$\alpha$R stars later in this section and in section 4.2.

The bulk of the high-$\alpha$ stars are populating the region from $\sim$ 8 to $\sim$ 14 Gyr and show no tight correlation between age and [$\alpha$/Fe]. In addition, old stars are not necessarily $\alpha$-rich; we identify a significant population of stars older than $\sim$ 10 Gyr with [$\alpha$/Fe] $<$ 0.1 (bottom right corner in the left panel of Fig 10). These findings seem in contrast to the results obtained by Haywood et al. (2013), who found a clear correlation between age and [$\alpha$/Fe] for both the low and high-$\alpha$ sequences and a steeper increase in [$\alpha$/Fe] with age for the high-$\alpha$ component. The low-$\alpha$ disk sequence on the other hand shows a gentle increase in [$\alpha$/Fe] with age up until $\sim$ 8 Gyr and remains flat after that, in agreement with the results of Haywood et al. (2013).

It is difficult to pinpoint at the moment the reason for the discrepancy between our results with those presented in Haywood et al. (2013). The authors of that paper considered a sample of 1111 solar neighbourhood dwarfs selected for exoplanet detection studies by Adibekyan et al. (2012), and determined isochrone-based ages for all stars discarding $\sim$ 70% of them due to unreliable results. Their findings are based on a subsample of only 363 stars with meaningful ages, corresponding to bright turn-off dwarfs were no assessment has been made of how representative they are of the underlaying population. Thus, the differences in our results could come from the techniques utilised (asteroseismology versus isochrone fitting), the larger number of stars in the high-$\alpha$ sequence analysed here, or the ensured completeness in our sample and appropriate correction for the target selection effects. We note, however, that other studies based on isochrones ages for turnoff stars also find larger scatter than Haywood et al. (2013) in the $\alpha$ abundances of the oldest disk stars (see e.g. Bensby, Feltzing & Oey 2014; Bergemann et al. 2014). In the future, we expect to perform comparisons of our results with volume complete local samples to further explore the reasons for these discrepancies.

Since our sample spans distances up to 2 kpc, it is interesting to compare the predictions of chemo-dynamical models of the Milky Way with our derived properties. A diagnostic widely used in studies of the solar neighbourhood is the expected relation between age and metallicity of stars, built up as stars chemically enrich the interstellar medium. Thus, more recently formed stars should have a higher abundance of metals than those born at earlier epochs (e.g., Feltzing, Holmberg & Hurley 2001, and references therein). The lack of a tight relation between these parameters is attributed to e.g., the efficiency of dynamical processes capable of erasing these signatures by moving stars from their birth radii to different orbits during their lifetime (see by e.g., Sellwood & Binney 2002; Roškar et al. 2008; Schönrich & Binney 2009a).

The right panel of Fig. 10 shows the age-metallicity relation of our chemically dissected sample. Contrary to what was reported by Haywood et al. (2013) we see no evidence of a significantly tighter relation between age and metallicity for the high-$\alpha$ population than for the lower-$\alpha$ sequence, but rather a broad distribution in chemical composition at all ages. Moreover, it is clear that the old (above $\sim$ 10 Gyr) low-$\alpha$ population shows a spread in metallicity from $-0.4 <$ [Fe/H] $< +0.4$, in disagreement with the results presented by Haywood et al. (2013) where the old low-$\alpha$ stars were believed to be contaminants coming from the low-metallicity end of the low-$\alpha$ sequence with higher Galactic rotation, and hence considered to come from the outer disk. Part of this discrepancy could be explained by the age uncertainties derived for the high-$\alpha$ population in our sample (see grey symbols in Fig. 10), although Haywood et al. (2013) quotes comparable errors of a 1.5 Gyr random component and about 1 Gyr systematic uncertainties for stars older than $\sim$ 9 Gyr. On the other hand sampling biases in the spectroscopic surveys can select against old low-$\alpha$ stars if these are rare (Bergemann et al. 2014), and thus it is not unexpected that they appear in photometrically complete samples such as ours or the Geneva-

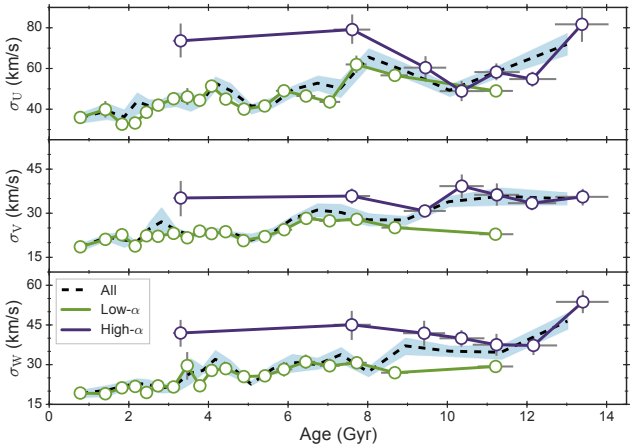


**Figure 11.** Velocity dispersions in the $U$, $V$, and $W$ components as a function of age for all stars on the sample (dashed lines), the low-$\alpha$ (green solid line), and the high-$\alpha$ (purple solid line) sequences, determined using bootstrapping. Shaded region corresponds to bootstrap uncertainties for the full sample. See text for details.

Copenhagen survey (Casagrande et al. 2011). The slope of the age-metallicity relation for the complete sample is at the level of $-0.008 \pm 0.001$ dex Gyr$^{-1}$ while the dispersion[3] in the metallicity shows a clear increase as a function of age (see dashed line in the bottom-right panel of Fig. 10), consistent with the predictions from chemo-dynamical models including radial migration (see e.g., Fig. 4 in Minchev, Chiappini & Martig 2013).

We now investigate variations in the stellar velocity dispersions as a function of age to disentangle kinematic signatures of different formation scenarios for the high- and low-$\alpha$ populations. Figure 11 shows the binned distributions for these two sequences and also for the full sample. The dispersions in the three velocity components show an increase with age for the full sample, and hints of a change in slope (in $V$ and $W$) at $\sim 8-9$ Gyr. This behaviour is indicative of the transition between two populations, which we explore by looking at the velocity dispersions of the individual components chemically selected in the disk.

The low-$\alpha$ sequence follows the behaviour of the full sample and steadily increases its dispersions with age until the transition point, and remains flat thereafter. The high-$\alpha$ sample presents relatively flat dispersions at all ages in the $V$ and $W$ components, and it appears to merge with the low-$\alpha$ sequence in $U$. It is interesting to see that the two chemical populations are clearly distinct in their velocity dispersions in the $V$ and $W$ components, and that the young stars belonging to the high-$\alpha$ sequence (in our definition) are kinematically hot (we remind the reader that not all of them correspond to the Y$\alpha$R population). Nevertheless, this result indicates that the young $\alpha$-rich stars have similar kinematic properties to the rest of the high-$\alpha$ sequence and are therefore likely to be born at the same time as the old high-$\alpha$ stars, rendering support to the idea that their ages are underestimated due to mass transfer or stellar merger events (Yong et al. 2016; Jofre et al. 2016). Our results indicate a $\sim 10\%$ fraction of young $\alpha$-rich stars among the high-$\alpha$ population, in line with the predictions by e.g., Tayar et al. (2015) for mass transfer products (see Jofre et al. 2016, for details).

The indication of a transition point at $\sim 8-9$ Gyr in the velocity dispersions suggest that the chemically low-$\alpha$ disk formed from the initial conditions set by the high-$\alpha$ disk component. Although similar conclusions have been reached by Haywood et al. (2013), the lack of a correlation between age and [$\alpha$/Fe] in the high-$\alpha$ sequence argues against the scenario proposed by those authors of a quiet chemical evolution lasting 4-5 Gyr that formed the high-$\alpha$ population. In contrast, and considering the spread in bulk metallicity for stars older than $\sim 8-9$ Gyr, our results indicate that over a period of more than 2 Gyr stars in the Galactic disk formed with a wide range of [$\alpha$/Fe] and [M/H] independent of time. A variety of formation scenarios could produce this signature, such as chaotic gas-rich mergers occurring with inhomogeneous chemical evolution before the younger low-$\alpha$ population built up (in agreement with $\Lambda$CDM predictions of decreasing merger rates, see e.g., Brook et al. 2004, 2012; Minchev, Chiappini & Martig 2013; Stinson et al. 2013), or thick disk formation in large molecular clumps induced by instabilities in a gas rich disk (e.g., Noguchi 1998; Bournaud, Elmegreen & Elmegreen 2007; Grand et al. 2017). Similar results are predicted from hydrodynamical simulations where the transition point is related to for example, the decline in the star formation rate (Brook et al. 2004) and thus it corresponds to the formation period of the thick disk.

### 4.2 Disk dynamics

To further explore different formation scenarios of the Milky Way disk and the origin of the two chemically defined populations, we compare the dynamical properties of our sample calculated using `galpy`[4] in the `MWPotential2014` configuration (Bovy 2015). Figure 12 shows the maximum vertical height of the stellar orbit as a function of age for both components, including the bootstrap fits to the mean values in both axes. The overall behaviour of the full sample is an increase in $Z_{max}$ with age, consistent with a vertical age gradient of the disk as seen by e.g., Ness et al. (2016); Casagrande et al. (2016). The latter is normally invoked in connection with disk formations scenarios such as upside-down (e.g., Brook et al. 2006; Bird et al. 2013) and pure disk heating, or a combination thereof.

The two chemically selected sequences span a similar range of values in $Z_{max}$ as a function of age, consistent with both populations having similar scale heights and thus not being representative of a geometric decomposition of the disk into a thin and a thick component (as found by e.g., Bovy et al. 2012, 2016). The running mean for each sequence shown in Fig. 12 reveals an average difference in $Z_{max}$ between the populations, with some overlap at the oldest ages.

In Fig. 13 we have plotted the distribution of guiding

[3] All bootstrap fits are constructed from 10,000 realisations of the dataset (with replacement) randomly drawn from a normal distribution of the mean value and uncertainty of the parameter in question, and weighted by the target selection probability of each individual target.

[4] http://github.com/jobovy/galpy

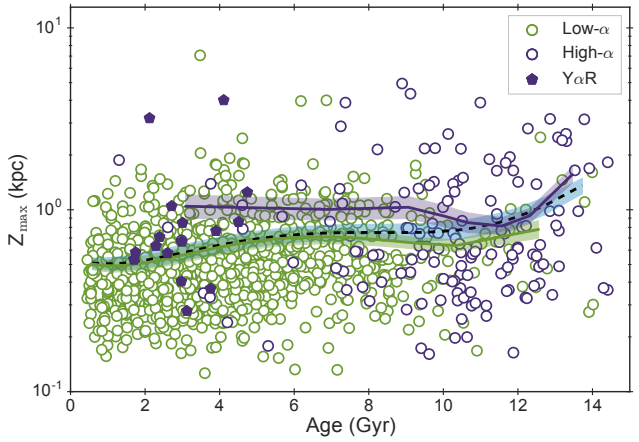


**Figure 12.** Maximum vertical height of the stellar orbit as a function of age for the chemically dissected disk. Solid lines show the bootstrap fit and uncertainties for the mean values of each $\alpha$-selected component, while the dashed line is the same result for the full sample.

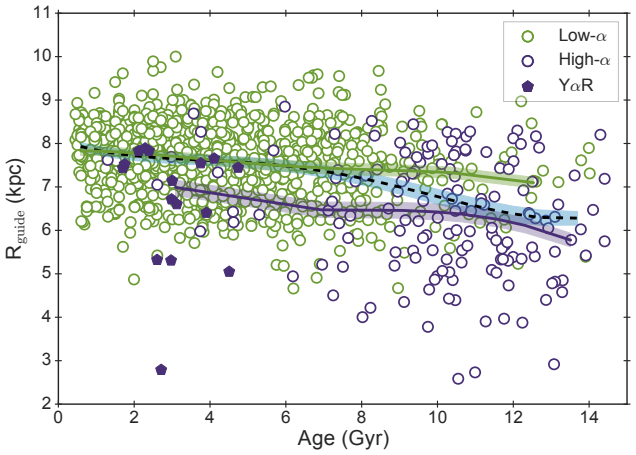


**Figure 13.** Same as Figure 12 for the guiding radius of the orbit.

radii of the orbits, calculated as $R_{guide} = L_z/v_c$ where $L_z$ is the angular momentum and assuming a constant circular rotation speed of $v_c = 220$ km s$^{-1}$. The running mean of $R_{guide}$ shows a gentle decline with age for the full sample while revealing an average lower value for the high-$\alpha$ stars than for the low-$\alpha$ population (see also Reddy et al. 2003; Bensby et al. 2011; Bovy et al. 2012; Cheng et al. 2012; Allende Prieto, Kawata & Cropper 2016). A decrease of the overall age-$R_{guide}$ relation can be explained by a combination of the age-velocity dispersion relation and the age-scale-length relation (see e.g., Mackereth et al. 2017, and references therein). If the decrease is driven by the latter, it would indicate a smaller scale-length for the older population and thus provide direct evidence for an inside-out formation scenario of the Milky Way disk. However, an increasing age-velocity dispersion relation leads to a larger asymmetric drift that can produce a similar trend in the age-$R_{guide}$ plane, because old stars are bound to have more negative velocities and thus come from the inner galaxy. A full chemo-dynamical model of the *Kepler* field is required to further explore this trend and will be the subject of a subsequent study.

Another interesting aspect from Fig. 13 is that all young $\alpha$-rich stars have values of $R_{guide}$ under 8 kpc, suggesting an inner Galaxy origin of this sample as postulated by Chiappini et al. (2015). However, the running mean of $R_{guide}$ as a function of age of the high-$\alpha$ population is also below 8 Kpc, which can be indicative that the young $\alpha$-rich stars are simply a subpopulation of the high-$\alpha$ sequence with similar distribution of orbits (as also proposed by Martig et al. 2015). To test the null hypothesis that both the Y$\alpha$R stars and the rest of the high-$\alpha$ population are drawn from the same parent distribution, we performed 10,000 random realisations with replacement from a normal distribution constructed from the individual mean values and uncertainties in the age-$R_{guide}$ plane for all high-$\alpha$ stars. The K-sample Anderson-Darling results show that the null hypothesis cannot be rejected with a significance level higher than 1.7%, indicating that the guiding radius distribution of the Y$\alpha$R stars is the same as for the rest of the high-$\alpha$ component and therefore are not confined to a specific position in the inner Galaxy (e.g., close to the Galactic Bar).

The argument of an inner-Galaxy origin for the young $\alpha$-rich population was put forward by Chiappini et al. (2015) after finding a larger number of these stars in the CoRoT sample looking towards the Galactic centre than in the sample observing in the anti-centre direction. One possible explanation for the discrepancy in our results could be that our *Kepler* stars are in the solar annulus while those of Chiappini et al. (2015) are distributed across Galactocentric distances from $\sim 4$ to $\sim 14$ kpc. To compare the two samples we retrieved the CoRoT stellar properties published by Anders et al. (2017) (including ages, abundances, position in the Galaxy and guiding radii), and identified the Y$\alpha$R population using the same criteria as described in Section 4. Figure 14 shows the spatial distribution of the CoRoT stars in the centre (LRc01) and anti-centre (LRa01) directions, accompanied by the position of our APOKASC sample.

The study by Anders et al. (2017) adopted proper motions from the UCAC-4 catalogue while our analysis benefited from the publication of the first Gaia data release, resulting in a much higher precision in guiding radius thanks to the improved astrometric properties (as seen in the right panel of Fig. 14). Despite the larger uncertainties in $R_{guide}$ for the CoRoT stars, some general trends can be observed in the spatial distribution of the samples. Even though the CoRoT stars are also mostly confined to $R_{guide} < 8$ kpc, the higher number of Y$\alpha$R found in the inner region could be the result of probing different distances from the Galactic plane: the number of young $\alpha$-rich stars is similar towards the inner and outer part of the Galaxy for $|Z| < 0.3$ kpc (see also Table 1 in Chiappini et al. (2015)). Unfortunately their sample in the anti-centre direction does not reach higher values of $|Z|$ and thus can only probe larger distances from the plane towards the Galactic centre, where they find a larger number of Y$\alpha$R targets. Our sample contains 10 out of 16 young $\alpha$-rich stars at $Z > 0.3$ kpc suggesting that distance to the Galactic plane instead of distance from the Galactic centre could play a role in finding an increasing number of these peculiar stars. This could simply be consequence of the increase in number of stars belonging to the high-$\alpha$

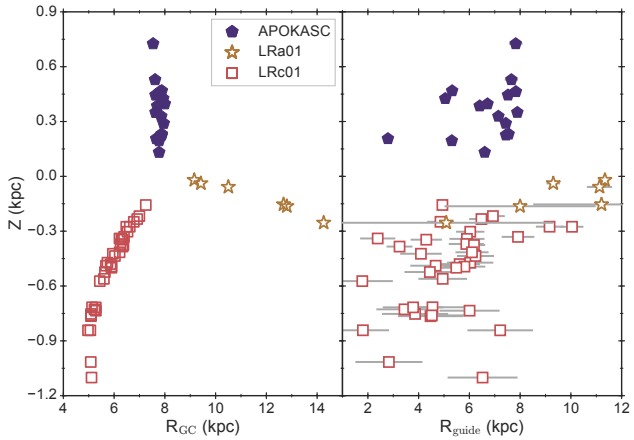


**Figure 14.** Distribution of young $\alpha$-rich stars versus height above the plane in the Galaxy of our sample (labelled APOKASC) and the CoRoT observations (labelled LRa01 and LRc01) extracted from Anders et al. (2017). *Left:* Galactocentric radius. *Right:* guiding radius and uncertainties (in most cases smaller than the symbol size for the APOKASC sample). See text for details.

population as we move away from the plane (see e.g., Hayden et al. 2015), thus increasing the probability of finding stars belonging to the Y$\alpha$R class. Extending the sample of these interesting stars to different Galactic directions, coupled with high precision astrometry from the Gaia mission, would help clarifying the origin of this peculiar class of objects.

## 5 CONCLUSIONS

By combining observations from the *Kepler* mission with APOGEE spectroscopy, we demonstrate the power of asteroseismology as a tool for Galactic archeology and determine precise physical, chemical, and kinematic properties in a sample of more than a thousand stars fully representative of the stellar population in the direction of the *Kepler* field. Our main results can be summarised as follows:

- Thanks to asteroseismic analysis, we can confirm with high fidelity that there is a clear age difference between the low- and high-$\alpha$ components, with the low-$\alpha$ sequence peaking at $\sim 2$ Gyr and the high-$\alpha$ one at $\sim 11$ Gyr.
- We observe a clear distinction in the $V$ and $W$ velocity dispersions between both components, suggesting a transition between the formation of both sequences $\sim 8-9$ Gyr ago. This renders support for a formation scenario of the high-$\alpha$ component lasting more than 2 Gyr, setting the initial conditions for the evolution of the low-$\alpha$ population.
- We see no tight correlation in [$\alpha$/Fe] or metallicity with age for the high-$\alpha$ sample. Our findings support a flat age metallicity relation with an increasing metallicity scatter as a function of age, consistent with models of radial migration.
- We recover the population of seemingly young $\alpha$-rich stars found by Chiappini et al. (2015); Martig et al. (2015). Their kinematics are similar to old high-$\alpha$ stars rather than low-$\alpha$ population at similar age (Fig. 11). This indicates that the majority of them are likely to be born at the same time as the old high-$\alpha$ stars, and therefore supports the idea that they are stellar merger remnants (Yong et al. 2016; Jofre et al. 2016). We also find evidence that the young $\alpha$-rich stars are vertically hotter than the low-$\alpha$ population (Fig 12), and that their guiding radii follows the same distribution as the rest of the high-$\alpha$ sequence. This seems to be inconsistent with previous interpretations of the spacial distribution of these young $\alpha$-rich stars as coming from the inner part of the Galaxy, suggesting that they formed in the bar region and migrated outwards. The latter scenario could be the result of an incomplete sampling at distances larger than $|Z| > 0.3$ kpc in the anti-centre direction in previous studies (see Fig 14).
- The distribution of $Z_{max}$ and $R_{guide}$ as a function of age for the full stellar sample analysed could provide evidence of inside-out formation of the Milky Way disk. A complete chemo-dynamical simulation of our sample in the *Kepler* field could help disentangle between this scenario and the effects of the asymmetric drift.

It is quite interesting that we find many old high metallicity and low-$\alpha$ stars, which are not seen in the solar neighbourhood studies (Haywood et al. 2013; Bensby, Feltzing & Oey 2014) but are predicted by standard chemical evolution models (see e.g., Fig 15. in Nidever et al. 2014). This may be due to the different methods for age determination (asteroseismology versus isochrone fitting) or could also be related to the difference in sample selection: we have ensured a representative group of stars to distances up to 3 kpc from us while the solar neighbourhood studies are mostly defined by much stricter data availability. A detailed comparison between ages determined from isochrone fitting and asteroseismic inferences could help shed some light on this topic. Similarly, data at different galactocentric radii and height from the plane would serve for studying how the age-[$\alpha$/Fe] and age-metallicity relations change at different location of the disk. Future analysis based on K2 for Galactic archaeology (Stello et al. 2015) can probe many other directions and thus extend our findings to other regions of the Milky Way, while a complete revision of the solar neighbourhood sample based on asteroseismic data will be possible from the TESS mission (Ricker et al. 2015).

## ACKNOWLEDGEMENTS

We would like to thank the anonymous referee for carefully revising the manuscript and suggesting changes that have improved the quality of the paper. Funding for this Discovery mission is provided by NASA's Science Mission Directorate. The authors acknowledge the dedicated team behind the Kepler and K2 missions, without whom this work would not have been possible. Funding for the Stellar Astrophysics Centre is provided by The Danish National Research Foundation (Grant agreement No. DNRF106). V.S.A. acknowledges support from VILLUM FONDEN (research grant 10118). L.C. is the recipient of an Australian Research Council Future Fellowship (project number FT160100402). D.K. and I.C. acknowledges the support of the UK's Science & Technology Facilities Council (STFC Grant ST/N000811/1 and Doctoral Training Partnerships Grant ST/N504488/1). MNL acknowledges the support of The Danish Council for Independent Research — Natural Science (Grant DFF-4181-00415). D.H. acknowledges

support by the Australian Research Council's Discovery Projects funding scheme (project number DE140101364) and support by the National Aeronautics and Space Administration under Grant NNX14AB92G issued through the Kepler Participating Scientist Program. J.A.J., M.P., and J.T. acknowledge support from NSF Grant AST-1211673. A.S. acknowledges support from grant ESP2015-66134-R (MINECO). D.S. is the recipient of an Australian Research Council Future Fellowship (project number FT1400147). W.H.T. acknowledges funding from the European Research Council under the European Union's Seventh Framework Programme (FP 7) ERC Grant Agreement n. [321035].

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofsica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatrio Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autnoma de Mxico, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

## REFERENCES

Adibekyan V. Z., Santos N. C., Sousa S. G., Israelian G., 2011, Astronomy and Astrophysics, 535, L11
Adibekyan V. Z., Sousa S. G., Santos N. C., Delgado Mena E., González Hernández J. I., Israelian G., Mayor M., Khachatryan G., 2012, Astronomy and Astrophysics, 545, A32
Allende Prieto C., Kawata D., Cropper M., 2016, Astronomy and Astrophysics, 596, A98
Anders F. et al., 2017, Astronomy and Astrophysics, 597, A30
Anders F. et al., 2014, Astronomy and Astrophysics, 564, A115
Bedding T. R. et al., 2011, Nature, 471, 608
Bensby T., Alves-Brito A., Oey M. S., Yong D., Meléndez J., 2011, The Astrophysical Journal Letters, 735, L46
Bensby T., Feltzing S., Lundström I., 2003, Astronomy and Astrophysics, 410, 527
Bensby T., Feltzing S., Oey M. S., 2014, A&A, 562, A71
Bergemann M. et al., 2014, Astronomy and Astrophysics, 565, A89
Bird J. C., Kazantzidis S., Weinberg D. H., Guedes J., Callegari S., Mayer L., Madau P., 2013, The Astrophysical Journal, 773, 43
Bland-Hawthorn J., Gerhard O., 2016, Annual Review of Astronomy and Astrophysics, 54, 529
Bournaud F., Elmegreen B. G., Elmegreen D. M., 2007, The Astrophysical Journal, 670, 237
Bovy J., 2015, The Astrophysical Journal Supplement Series, 216, 29
Bovy J., Rix H.-W., Hogg D. W., 2012, The Astrophysical Journal, 751, 131
Bovy J., Rix H.-W., Liu C., Hogg D. W., Beers T. C., Lee Y. S., 2012, The Astrophysical Journal, 753, 148
Bovy J., Rix H.-W., Schlafly E. F., Nidever D. L., Holtzman J. A., Shetrone M., Beers T. C., 2016, The Astrophysical Journal, 823, 30
Brook C. B., Kawata D., Gibson B. K., Freeman K. C., 2004, The Astrophysical Journal, 612, 894
Brook C. B., Kawata D., Martel H., Gibson B. K., Bailin J., 2006, The Astrophysical Journal, 639, 126
Brook C. B. et al., 2012, Monthly Notices of the Royal Astronomical Society, 426, 690
Brown T. M., Gilliland R. L., Noyes R. W., Ramsey L. W., 1991, ApJ, 368, 599
Brown T. M., Latham D. W., Everett M. E., Esquerdo G. A., 2011, AJ, 142, 112
Casagrande L., Schönrich R., Asplund M., Cassisi S., Ramírez I., Meléndez J., Bensby T., Feltzing S., 2011, A&A, 530, 138
Casagrande L. et al., 2016, Monthly Notices of the Royal Astronomical Society, 455, 987
Casagrande L. et al., 2014, The Astrophysical Journal, 787, 110
Casagrande L., Vandenberg D. A., 2014, Monthly Notices of the Royal Astronomical Society, 444, 392
Cheng J. Y. et al., 2012, ApJ, 746, 149
Chiappini C. et al., 2015, Astronomy and Astrophysics, 576, L12
Da Silva L. et al., 2006, A&A, 458, 609
Epstein C. R. et al., 2014, The Astrophysical Journal, 785, L28
Feltzing S., Bensby T., 2008, Physica Scripta, T133, 014031
Feltzing S., Holmberg J., Hurley J. R., 2001, A&A, 377, 911
Freeman K., Bland-Hawthorn J., 2002, Annual Review of Astronomy and Astrophysics, 40, 487
Fuhrmann K., 1998, Astronomy and Astrophysics, 338, 161
Fuhrmann K., 2011, Monthly Notices of the Royal Astronomical Society, 414, 2893
Gaia Collaboration et al., 2016, Astronomy and Astrophysics, 595, A2
Gaulme P. et al., 2016, The Astrophysical Journal, 832, 121
Gilmore G., Reid N., 1983, Monthly Notices of the Royal Astronomical Society, 202, 1025
Grand R. J. J. et al., 2017, arXiv.org, arXiv:1708.07834
Grand R. J. J., Kawata D., Cropper M., 2015, Monthly Notices of the Royal Astronomical Society, 447, 4018
Green G. M. et al., 2015, The Astrophysical Journal, 810, 1
Guggenberger E., Hekker S., Basu S., Bellinger E., 2016, Monthly Notices of the Royal Astronomical Society, 460, 4277
Hayden M. R. et al., 2015, The Astrophysical Journal, 808, 132
Haywood M., Di Matteo P., Lehnert M. D., Katz D., Gómez A., 2013, Astronomy and Astrophysics, 560, A109
Ho A. Y. Q., Rix H.-W., Ness M. K., Hogg D. W., Liu C., Ting Y.-S., 2017, The Astrophysical Journal, 841, 0
Hon M., Stello D., Yu J., 2017, Monthly Notices of the Royal Astronomical Society, 469, 4578
Huber D. et al., 2011, ApJ, 743, 143
Huber D. et al., 2012, ApJ, 760, 32
Huber D. et al., 2014, The Astrophysical Journal Supplement Series, 211, 2
Huber D. et al., 2017, The Astrophysical Journal, 844, 0
Jofre P. et al., 2016, Astronomy and Astrophysics, 595, A60
Johnson D. R. H., Soderblom D. R., 1987, Astronomical Journal (ISSN 0004-6256), 93, 864
Jurić M. et al., 2008, The Astrophysical Journal, 673, 864

Kjeldsen H., Bedding T. R., 1995, A&A, 293, 87
Lindegren L. et al., 2016, Astronomy and Astrophysics, 595, A4
Loebman S. R., Roškar R., Debattista V. P., Ivezić Ž., Quinn T. R., Wadsley J., 2011, The Astrophysical Journal, 737, 8
Mackereth J. T. et al., 2017, Monthly Notices of the Royal Astronomical Society, 471, 3057
Majewski S. R. et al., 2017, The Astronomical Journal, 154, 0
Martig M. et al., 2016, Monthly Notices of the Royal Astronomical Society, 456, 3655
Martig M. et al., 2015, Monthly Notices of the Royal Astronomical Society, 451, 2230
Matteucci F., Francois P., 1989, Monthly Notices of the Royal Astronomical Society (ISSN 0035-8711), 239, 885
Miglio A. et al., 2012, MNRAS, 419, 2077
Miglio A. et al., 2016, Monthly Notices of the Royal Astronomical Society, 461, 760
Miglio A. et al., 2013a, EPJ Web of Conferences, 43, 03004
Miglio A. et al., 2013b, MNRAS, 429, 423
Minchev I., Chiappini C., Martig M., 2013, Astronomy and Astrophysics, 558, A9
Mosser B. et al., 2012, A&A, 540, 143
Navarro J. F., Abadi M. G., Venn K. A., Freeman K. C., Anguiano B., 2011, Monthly Notices of the Royal Astronomical Society, no
Ness M., Hogg D. W., Rix H. W., Ho A. Y. Q., Zasowski G., 2015, The Astrophysical Journal, 808, 16
Ness M., Hogg D. W., Rix H. W., Martig M., Pinsonneault M. H., Ho A. Y. Q., 2016, The Astrophysical Journal, 823, 114
Nidever D. L. et al., 2014, The Astrophysical Journal, 796, 38
Noguchi M., 1998, Nature, 392, 253
Nordström B. et al., 2004, A&A, 418, 989
Origlia L., Ferraro F. R., Fabbri S., Fusi Pecci F., Dalessandro E., Rich R. M., Valenti E., 2014, Astronomy and Astrophysics, 564, A136
Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2004, ApJ, 612, 168
Pinsonneault M. H., An D., Molenda-Żakowicz J., Chaplin W. J., Metcalfe T. S., Bruntt H., 2012, The Astrophysical Journal Supplement, 199, 30
Pinsonneault M. H. et al., 2014, The Astrophysical Journal Supplement Series, 215, 19
Ramírez I., Allende Prieto C., Lambert D. L., 2013, The Astrophysical Journal, 764, 78
Reddy B. E., Tomkin J., Lambert D. L., Allende Prieto C., 2003, MNRAS, 340, 304
Reimers D., 1977, A&A, 57, 395
Ricker G. R. et al., 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003
Rix H.-W., Bovy J., 2013, The Astronomy and Astrophysics Review, 21, 61
Robin A. C., Reylé C., Fliri J., Czekaj M., Robert C. P., Martins A. M. M., 2014, Astronomy and Astrophysics, 569, A13
Rodrigues T. S. et al., 2017, Monthly Notices of the Royal Astronomical Society, stx120
Rodrigues T. S. et al., 2014, Monthly Notices of the Royal Astronomical Society, 445, 2758
Rojas-Arriagada A. et al., 2016, Astronomy and Astrophysics, 586, A39
Roškar R., Debattista V. P., Quinn T. R., Stinson G. S., Wadsley J., 2008, The Astrophysical Journal Letters, 684, L79
Salaris M., Chieffi A., Straniero O., 1993, ApJ, 414, 580
Schönrich R., Binney J., 2009a, MNRAS, 396, 203
Schönrich R., Binney J., 2009b, MNRAS, 399, 1145
Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829
Schönrich R., McMillan P. J., 2017, Monthly Notices of the Royal Astronomical Society, stx093
SDSS C. et al., 2016, arXiv.org, arXiv:1608.02013
Sellwood J. A., Binney J. J., 2002, MNRAS, 336, 785
Serenelli A. M., Bergemann M., Ruchti G., Casagrande L., 2013, MNRAS, 429, 3645
Sharma S., Stello D., Bland-Hawthorn J., Huber D., Bedding T. R., 2016, The Astrophysical Journal, 822, 15
Sheather S. J., Jones M. C., 1991, Journal of the Royal Statistical Society. Series B. Methodological, 53, 683
Silva Aguirre V. et al., 2013, ApJ, 769, 141
Silva Aguirre V. et al., 2012, ApJ, 757, 99
Silva Aguirre V. et al., 2015, Monthly Notices of the Royal Astronomical Society, 452, 2127
Silva Aguirre V. et al., 2017, The Astrophysical Journal, 835, 1
Skrutskie M. F. et al., 2006, The Astronomical Journal, 131, 1163
Soubiran C., Bienaymé O., Siebert A., 2003, Astronomy and Astrophysics, 398, 141
Stello D. et al., 2013, ApJ, 765, L41
Stello D. et al., 2015, The Astrophysical Journal Letters, 809, L3
Stello D. et al., 2016, The Astrophysical Journal, 832, 1
Stinson G. S. et al., 2013, Monthly Notices of the Royal Astronomical Society, 436, 625
Tayar J. et al., 2015, The Astrophysical Journal, 807, 82
Tayar J. et al., 2017, The Astrophysical Journal, 840, 0
Torres G., 2010, AJ, 140, 1158
Ulrich R. K., 1986, ApJ, 306, L37
Venables W. N., Ripley B. D., 2002, Modern Applied Statistics with S, 4th edn., Statistics and Computing. Springer
Vrard M., Mosser B., Samadi R., 2016, Astronomy and Astrophysics, 588, A87
White T. R., Bedding T. R., Stello D., Christensen-Dalsgaard J., Huber D., Kjeldsen H., 2011, ApJ, 743, 161
White T. R. et al., 2013, MNRAS, 433, 1262
Yong D. et al., 2016, Monthly Notices of the Royal Astronomical Society, 459, 487
Zacharias N., Finch C. T., Girard T. M., Henden A., Bartlett J. L., Monet D. G., Zacharias M. I., 2013, The Astronomical Journal, 145, 44

## APPENDIX A: STELLAR PROPERTIES DERIVED WITH BASTA

The stellar properties derived with `BASTA` for the sample of stars analysed in this paper are published in the online version of the article. A description of all fields available is presented in Table A1.

## APPENDIX B: TESTS OF TARGET SELECTION

Our procedure for correcting for target selection effects outlined in section 3 assumes flat priors in age, metallicity and distances. In fact, the elegance of the data-cube approach is to return the probability of observing a star at any given point of the parameter space without invoking any Galactic modelling. The only assumption entering our method is the IMF, which we adopt to be a Salpeter one. We have tested the effect of drastically changing this prior, by assuming a uniform distribution in mass (flat IMF), and generating a new multi-dimensional data cube to determine the probability of detection. Figure B1 shows the resulting normalised probability as a function of age, metallicity, and distance, which can be directly compared to Fig. 5. The overall shape of the distributions is unchanged.

Figure B2 shows the age distribution of the low- and high-$\alpha$ samples corrected for target selection using the probabilities

**Table A1.** Stellar properties determined with `BASTA` for the APOKASC sample.

| Field | Description |
|---|---|
| `KIC` | *Kepler* Input Catalogue Identifier |
| `Mass` | Mass in solar units |
| `Mass_err` | Mass uncertainty in solar units |
| `Rad` | Radius in solar units |
| `Rad_err` | Radius uncertainty in solar units |
| `Grav` | Surface gravity in dex |
| `Grav_err` | Surface gravity uncertainty in dex |
| `Age` | Age in units of Gyr |
| `Age_err` | Age uncertainty in units of Gyr |
| `Lum` | Luminosity in solar units |
| `Lum_err` | Luminosity uncertainty in solar units |
| `Dist` | Distance in pc |
| `Dist_err` | Distance uncertainty in pc |
| `Prob` | Target selection probability (cf section 3) |

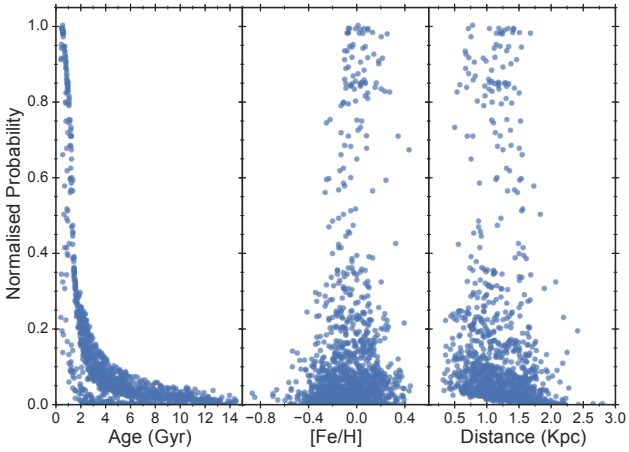


**Figure B1.** Same as Fig 5 but applying a correction for target selection based on a uniform prior in the IMF.

calculated with a uniform mass prior. As it can be see the resulting distributions are almost unchanged compared to those shown in Fig. 9, also peaking at values of $\sim 2$ and $\sim 11$ Gyr for the low- and high-$\alpha$ components.

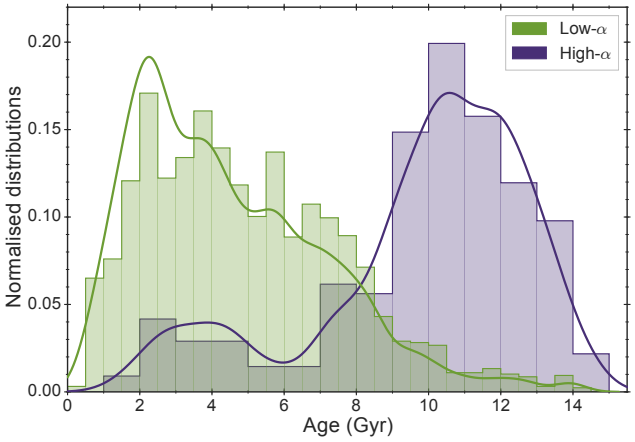


**Figure B2.** Same as Fig 9 but applying a correction for target selection based on a uniform prior in the IMF.