Galactic Doppelgängers: The Chemical Similarity Among Field Stars and Among Stars with a Common Birth Origin

M. Ness1,2,3, H-W. Rix1, David W. Hogg1,3,4,5, A. R. Casey6, J. Holtzman7, M. Fouesneau1, G. Zasowski8,9, D. Geisler10, M. Shetrone11, D. Minniti12,13,14, Peter M. Frinchaboy15, and Alexandre Roman-Lopes16

1 Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany; melissa.ness@columbia.edu
2 Department of Astronomy, Columbia University, Pupin Physics Laboratories, New York, NY 10027, USA
3 Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA
4 Department of Physics & Astronomy, Texas Christian University, TCU Box 298840, Fort Worth, TX 76129, USA
5 Department of Physics, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
6 Department of Astronomy, New Mexico State University, Las Cruces, NM 88003, USA
7 Department of Physics & Astronomy, University of Utah, 115 S. 1400 E., Salt Lake City, UT 84112, USA
8 Department of Physics & Astronomy, Universidad de Chile, A. J. Cisternas, 1200 North, La Serena, Chile
9 Space Telescope Science Institute, 3700 San Martin Drive, Baltimore MD 21218, USA
10 Departamento de Astronomía, Casilla 160-C, Universidad de Concepción, Chile
11 University of Texas at Austin, McDonald Observatory, USA
12 Instituto Milenio de Astrofísica, Santiago, Chile
13 Departamento de Física, Facultad de Ciencias Exactas, Universidad Andres Bello, Av. Fernandez Concha 700, Las Condes, Santiago, Chile
14 Vatican Observatory, V00120 Vatican City State, Italy
15 Department of Physics & Astronomy, Texas Christian University, TCU Box 298840, Fort Worth, TX 76129, USA
16 Department of Physics & Astronomy, Universidad de La Serena, A. J. Cisternas, 1200 North, La Serena, Chile

Received 2017 January 24; revised 2017 November 23; accepted 2017 November 24; published 2018 February 5

Abstract

We explore to what extent stars within Galactic disk open clusters resemble each other in the high-dimensional space of their photospheric element abundances and contrast this with pairs of field stars. Our analysis is based on abundances for 20 elements, homogeneously derived from APOGEE spectra (with carefully quantified uncertainties of typically 0.03 dex). We consider 90 red giant stars in seven open clusters and find that most stars within a cluster have abundances in most elements that are indistinguishable even at 0.03 dex precision exist: ~0.3% of all field star pairs and ~1.0% of field star pairs at the same (solar) metallicity [Fe/H] = 0 ± 0.2. Most of these pairs are presumably not birth siblings from the same cluster, but rather doppelgängers. Our analysis implies that “chemical tagging” in the strict sense, identifying birth siblings for typical disk stars through their abundance similarity alone, will not work with such data. However, our approach shows that abundances have extremely valuable information for probabilistic chemo-orbital modeling, and combined with velocities, we have identified new cluster members from the field.

Key words: Galaxy: abundances – Galaxy: disk – Galaxy: formation – methods: statistical – open clusters and associations: individual (abundances) – stars: abundances

1. Introduction

There now exists an abundance of spectroscopic data for stars across the Milky Way from which numbers of element abundances have been, or can be, measured. These data are being delivered by current surveys including Gaia-ESO (Gilmore et al. 2012), APOGEE (Majewski et al. 2017), RAVE (Casey et al. 2017; Kunder et al. 2017), and GALAH (Freeman 2012; De Silva et al. 2015), and there are numerous future spectroscopic surveys being planned, including 4MOST (de Jong & Consortium 2015), MOONS (Cirasuolo et al. 2012), and WEAVE (Dalton et al. 2014). Large volumes of data, combined with new techniques to optimally exploit the data and deliver high-precision stellar abundances, e.g., The Cannon (Ness et al. 2015; Casey et al. 2016), enable us to examine the distribution in abundance space of the stars in the Milky Way disk and halo. This information can then, in principle, be used to constrain the assembly history of our Galaxy. One way to pursue this is to identify stars of common birth sites via their abundance signatures, a process called chemical tagging (Freeman & Bland-Hawthorn 2002). To do this, we need to first know what the intrinsic spread, if any, is in the abundance measurements of stars that are known to be born together.

Stars in open clusters are believed to be born from single star-forming aggregates and therefore these stars should be chemically homogeneous (e.g., De Silva et al. 2007; de Silva et al. 2009; Bland-Hawthorn et al. 2010). The chemical similarity of stars born together has been used to identify members of moving groups or co-natal groups in the Milky Way, including using abundances alone (e.g., Majewski et al. 2012; Hogg et al. 2016) as well as chemically exceptional groups of stars (e.g., Martell et al. 2016; Schiavon et al. 2017). Many of the individual clusters have been studied in numerous studies, including using APOGEE data (e.g., Frinchaboy et al. 2013; Cunha et al. 2015; Souto et al. 2016), and the measured abundance differences within clusters have been demonstrated to be relatively small, on the order of the measurement errors themselves (Reddy et al. 2012; Mitschang et al. 2014; Reddy et al. 2015; Lambert & Reddy 2016). Statistically there is a much greater abundance similarity among stars within a cluster compared to stars from different clusters (e.g., Mitschang
et al. 2013; De Silva et al. 2015). However, although the question of the magnitude of open cluster abundance dispersion has been investigated previously, there has been very little assessment of the true intrinsic dispersion of clusters in their many elements (see also Mitschang et al. 2013).

The abundance dispersion that we measure at face value in groups of stars that are born together in clusters (or potentially also in associations) is a combination of the intrinsic dispersion of the cluster and the measurement uncertainties. Quantifying this intrinsic dispersion for clusters is indispensable not only for the prospects of chemical tagging—this is also critical for understanding the chemical-abundance distribution and dimensionality of the Milky Way disk and key to being able to determine when stars are likely to be born together.

A high level of abundance homogeneity (to within ~0.03 dex) has been demonstrated for three open clusters in APOGEE, M67, NGC 6819, and NGC 2410, from the stellar spectra directly (Bovy 2016). After removing temperature trends, the cluster spectra form one-dimensional sequences. This novel data-driven approach was motivated by the difficulties in obtaining consistent high-precision stellar abundance measurements, and such an approach is optimal to achieve high-precision limits on the intrinsic cluster abundance dispersions, and so assess the homogeneity of a single birth site itself. It is only this data-driven approach, which did not provide actual abundance measurements, that has managed to place tight constraints on the intrinsic cluster dispersion (Bovy 2016). However, the limitation in such a method is that it does not return absolute abundance measurements, and as such, the open cluster measurements cannot be directly compared with those of the field, which is our aim here.

With our modifications to The Cannon and a carefully selected high-fidelity training set, we achieve high precision and report absolute abundance measurements for 90 open cluster stars. From these consistent, high-precision abundance measurements across ranges of temperature and gravity, we make an assessment of the multi-element abundance dispersion (or limits on it) given our thorough characterization of our measurement uncertainties. Our results constitute the largest homogeneous and consistent analysis of the intrinsic abundance dispersion within open clusters.

Comparing these high-precision abundance measurements in APOGEE among stars in an open cluster and among disk field stars can shape our expectation for the chemical similarity of stars that were born together—and those that were presumably not (but are still members of the Milky Way’s dominant disk population). We should expect that—with abundance data of this quality—most intracluster pairs of stars look very similar or indistinguishable, while most pairs of field stars have discernible abundance differences. Importantly, these data can allow us to determine the fraction of field stars that are not related by birth origin, yet are as chemically similar as those born in the same molecular cloud. In their abundances, these pairs of stars look perfectly alike (given the 20 APOGEE abundances), but are in most cases not birth siblings from the same cluster; we dub those stars doppelgängers. The rate of doppelgängers obviously plays a decisive role in the efficacy of chemical tagging.

In our analysis, there are two key ingredients for delivering our high-precision measurements from APOGEE data: (i) We modified our spectral modeling, The Cannon, to correct for spurious abundance variations across fiber number, caused by the varying line-spread function (LSF) across the APOGEE detectors; and in The Cannon’s training step we used abundances that have been corrected for LSF variations. (ii) We selected a high-fidelity training set using labels from APOGEE’s stellar parameter and abundances pipeline ASPCAP (García Pérez et al. 2016): 5000 high-signal-to-noise stars that made it sensible to train The Cannon on a total of 23 labels: three stellar parameters, T eff, log g, and [Fe/H]; 19 [X/Fe] individual abundances from DR13 (Holtzman et al. 2015, Johnson et al. 2016, submitted); plus the mean LSF of the star. In Section 2, we present our training set, and in Section 3, we discuss our modifications to The Cannon and our training set in order to cope with the problem of mean abundance changing as a function of LSF. In Section 4, we show the precision we achieve for our abundances from the cross-validation of the training set and report the abundances we obtain for the individual clusters. We use our uncertainties to determine the intrinsic spread in each element using our cluster model. In Section 5, we examine the intracluster and field pair abundance similarity distributions and from this we demonstrate that the abundance similarity among pairs of stars within clusters, defined via a χ 2 -distance in a 20 element abundance space, is far greater than that among random pairs of field stars. We also determine the doppelgänger rate among Galactic disk stars (for data sets of this quality): we find this rate to be very small (~1%), yet large enough to matter greatly for chemical tagging. In Section 6, we present as an aside new cluster members that we have identified from the field using our approach of examining stellar element abundance similarities between pairs. In Section 7, we discuss the implications of our results.

Our work makes the first quantification of the doppelgänger rate, which directly tests the viability of strict chemical tagging. Our analysis also demonstrates the diagnostic power in the high-dimensional abundance space: as part of our analysis, we identify new cluster members by combining the chemical information using our pair similarity measure and radial velocity data. The high-precision measurements that we derive are therefore important for the broader assessment of disk (and cluster) formation. In general, such high-precision measurements as we derive, with their carefully characterized uncertainties, are absolutely necessary to make optimum use of the large data sets to characterize the chemical diversity, distribution and, conditioned on expectations of the properties of open clusters, the chemical dimensionality and variability of the Milky Way disk.

2. APOGEE Data

For our analysis, we use two aspects of the APOGEE data: we used the APOGEE spectra (SDSS Collaboration et al. 2017) and stellar parameter and abundances from the SDSS-IV public data release DR13 (Wilson et al. 2012; Zasowski et al. 2013; Holtzman et al. 2015; Nidever et al. 2015; García Pérez et al. 2016). We performed our own signal-to-noise independent continuum normalization on the ascapStar files, similarly to Ness et al. (2015), by fitting a second-order polynomial to pixels identified with weak parameter dependences. For all DR13 spectra, we selected around 5% of the pixels using the criteria of 0.985 > flux > 1.025 and the spectral model coefficients ∥θ T eff , θ log g , θ [Fe/H] ∥ < (0.005, 0.005, 0.005). This normalization was applied consistently to the APOGEE
spectra in both the training and the test step of The Cannon (Ness et al. 2015).

2.1. Training Data

We constructed a training set of data (including their reference labels and their spectra) that consists of 5000 stars with signal-to-noise ratio (S/N) > 200 that span the chemical space of the stars of the Milky Way disk and broadly encompass the label space of the cluster members in the test step. For our reference labels, we used the so-called ASPCAP DR13-corrected labels with small additional corrections for the LSF dependence (see Section 3). Before starting to build our spectral model in the training step, we eliminated stars with problematic abundance measurements from ASPCAP and with the ASPCAP BAD flag checked. Our reference labels span the following range in stellar parameters:

\[
3650 < T_{\text{eff}} < 5760 \text{ K} \\
0.45 < \log g < 3.95 \text{ dex} \\
-1.7 < [\text{Fe/H}] < 0.36 \text{ dex}.
\]

2.2. Test Data

Our test data are the spectra of 97 stars in seven open clusters, where our membership is taken from the crossover between those identified by Mészáros et al. (2013) and those in the APOGEE calibration file table.\(^\text{17}\) For NGC 2420, we took an additional three stars from those identified in Mészáros et al. (2013) that were studied by Souto et al. (2016). The properties of the seven open clusters are summarized in Table 1.

### Table 1

| Cluster   | \(l\)  | \(b\)  | Distance (kpc) | \(E(B − V)\) | \(\text{log}(\text{Age})\) | [Fe/H] |
|-----------|--------|--------|----------------|-------------|-----------------|--------|
| NGC 7789  | 115.532| −5.385 | 2.34           | 0.22        | 9.235           | −0.04  |
| NGC 6819  | 73.978 | 8.481  | 2.36           | 0.24        | 9.174           | 0.03   |
| NGC 6791  | 69.959 | 10.904 | 4.10           | 0.12        | 9.643           | 0.21   |
| NGC 2420  | 198.107| 19.634 | 3.09           | 0.03        | 9.048           | −0.19  |
| NGC 2158  | 186.634| 1.781  | 5.07           | 0.36        | 9.023           | −0.23  |
| NGC 188   | 122.843| 22.384 | 2.05           | 0.08        | 9.632           | 0.05   |
| M67 (NGC 2682) | 215.696| 31.896 | 0.91           | 0.06        | 9.409           | 0.0    |

Note. The cluster data are from the WEBDA Open Cluster Database, http://www.univie.ac.at/webda/, and the [Fe/H] data are from our measurements (see Table 2). Using the DR13 data, we established that the measured abundances depend on LSF width, shown in Figure 1. This figure shows the mean DR13 abundances of our training set of stars, as a function of the FHWM of the LSF for the stars, measured from the apStarLSF files provided by APOGEE. To map the overall trends of abundances across the disk, these trends are not significant and will not affect global measured abundance trends and inferences, for example, with respect to the abundance trends based on the measurements across \((R, z)\), which are on average inclusive of all fiber numbers, independent of position on the sky (except for targeted co-natal groups like streams and open and globular clusters). However, these effects may be extremely important when trying to make the most precise abundances comparison among pairs or groups of stars. Therefore, it is necessary to remove these trends before assessing abundance homogeneity between stars. Such effects may increase the abundance differences, but they may also make stars observed with the same fiber appear spuriously similar in their abundances (Section 5). If one takes the DR13 ASPCAP abundances unaltered, the abundance dispersion of clusters whose stars were observed mostly with low fiber numbers (which equates to the lowest FHWM of the LSF), e.g., NGC 2158, NGC 6791, and M67, appears higher than those with a high fiber number. This is particularly manifest for “problematic” elements such as Cu, the mean abundance of which changes very quickly with LSF or fiber number. This is equivalent to the lowest fiber numbers (Figure 1).

To correct for the LSF variation, we need to add two processing steps to The Cannon: the first affects the reference labels, the second The Cannon’s spectral model. First, we correct the systematic biases in the reference labels shown in Figure 1 by fitting a fifth-order polynomial to the mean trend of each element and adjusting each element by this fit value (given the FHWM of the LSF of the spectrum under consideration). For most elements, this correction is very small or negligible. But for elements like Cu, it can be as high as 0.2 dex for the lowest LSF value (see Figure 1).

Second, we include the FHWM of the LSF as a label in The Cannon, similarly to the stellar parameters and abundances, and we include the FHWM of the LSF as a data point for our model for each star, which is treated in The Cannon’s model in exactly the same way as the flux. We can consider the LSF’s width as a data point, as we know mean LSF width from the apStarLSF files produced for each spectrum. To determine

---

\(^\text{17}\) cal_dr13.fits available at sdss.org.
the LSF, we take the average value of the standard deviation of the Gaussian fit across every pixel of the apStarLSF array. This approach to handling LSF variations is an approximation to the ideal approach, which would be to treat the LSF width as deterministically known at both training and test time, and treat it as a convolution and deconvolution operation. That, however, would significantly complexify both the training and test steps, without manifest practical advantage.

We demonstrate in Figure 2 that this procedure removes the systematic trend of the median [P/Fe] (one of the more affected elements) with the width of the LSF, which is present prior to our correction.

3.0.1. Cross-validation of This Revised Data Analysis

Given the slight modifications to The Cannon and the particular reference set, we initiate the analysis with an assessment of the label precision. We do this with a cross-validation test on the training set, similarly to Ness et al. (2015), Ho et al. (2017), and Casey et al. (2016). We perform a set of 10 leave-10%-out cross-validations, which are illustrated in Figure 3. The stars shown in this figure with label estimates from The Cannon had all been excluded from the training set. The model was constructed using the remaining 90% of the training set in each case. The x-axis shows the input training labels, and the y-axis shows The Cannon’s best-fit labels. We are able to recover the labels to high precision: the stellar parameter precisions are \( \lesssim 45 \) K in \( T_{\text{eff}} \), \( \lesssim 0.1 \) dex in \( \log g \), and \( \lesssim 0.02 \) dex in [Fe/H]. For individual abundances, the precision is 0.02 to 0.12 dex, depending on the element.

Our model fit of The Cannon to the test data is excellent in a \( \chi^2 \)-sense, and we show a typical star, from the cluster stars as a function of the FWHM of the LSF. The dashed red line shows the ASPCAP result and the black line The Cannon’s result. The error bars are the sampling errors on the measurements. The Cannon’s results have the systematic trend of abundance with the FWHM of the LSF removed.

Figure 1. Normalized median measurements of the ASPCAP DR13-corrected labels as a function of measured LSF for the 5000 stars used for the training set: the data are normalized by subtracting the overall median measurement from the data, so the data are pivoted around a value = 0 to highlight the offsets on the same scale. The FWHM of the median measurement is shown in the gray shaded region. The largest difference in the median and the dispersion around the median is written at the top of each panel. We correct for these offsets to produce LSF-corrected labels that we use for training The Cannon. For many elements, this correction is negligible; for some, like P or Cu, it is significant.

Figure 2. Median [P/Fe] measurement, which is one of the more affected elements by the varying LSF across the detector (see Figure 1), of the cluster stars as a function of the FWHM of the LSF. The dashed red line shows the ASPCAP result and the black line The Cannon’s result. The error bars are the sampling errors on the measurements. The Cannon’s results have the systematic trend of abundance with the FWHM of the LSF removed.
restrict the pixels that The Cannon uses to deliver the abundance information; the cross-validation demonstrates that The Cannon “learns” where the information about each element is derived from in the spectra, such that the input labels are successfully reproduced at test time.

4. Results

We now present the results from these data in three steps. First, for an analysis of abundance dispersions, or differences, it is not only precise measurements that matter, but also a good understanding of the uncertainties: over- or underestimates of the individual abundance uncertainties would lead to under- or overestimates of the cluster-intrinsic abundance dispersions or of the abundance differences among pairs of stars. Second, we proceed to quantify the abundance dispersions in each of the elements in each of our clusters. Third, we show how to check for any systematic abundance trends. The typical 1—σ standard deviation measurements for the elements around their mean values range from 0.01—0.1. The cyan points in the background are the labels for all cluster stars used in the training of The Cannon. The DR13 input abundance labels that we adopted for training are already corrected for systematic variations with temperature (Holtzman et al. 2015), except for C and N, which are known to have astrophysical variations along the giant branch (and then additionally corrected as part of our procedure for LSF variation as described in Section 3). For the open clusters, we note that the coolest stars do seem to have the highest measurements of [N/Fe] within a cluster. We find that there are no trends in [Fe/H] along the giant branch for member stars of the clusters, as has been reported for globular clusters.

4.1. Revised APOGEE Abundances and Abundance Uncertainties for Members of the Open Clusters

Based on the goodness-of-fit \( \chi^2 \) metric determined above for each star, we excluded seven stars from the cluster NGC 6791, with \( \chi^2 > 22,000 \). For our analysis, we do not exclude cluster stars with relatively low S/N measurements (S/N < 100), but use the S/N-dependent error to estimate the scaled precision at low S/N. We determine the S/N dependence of the precision using a calibration set of data available as part of the public release of DR13, which includes both co-added spectra and individual visit spectra for a set of \( \approx 1000 \) open clusters, globular clusters, and calibration stars. A sample of elements is shown in Figure 5, where each data point is determined by measuring the variance of the difference between the measured abundance from high S/N spectra (comprised of combined individual observations) and the individual spectra with low S/N, for each of the 1000 stars in the calibration set. The figure shows the dependence across S/Ns of 10 to 200, where the precision flattens above an S/N of about 150.

Our cluster stars have reported S/Ns from 60 to 1000. We then adopt abundance uncertainties for each star that are the quadrature sum of the signal-to-noise scaled uncertainties estimated from the cross-validation and the formal errors returned by the The Cannon’s test step optimizer.

Figures 6–12 show the individual cluster abundances with their corresponding uncertainties. The points are colored by \( T_{\text{eff}} \) to check for any systematic abundance trends. The typical 1—σ standard deviation measurements for the elements around their mean values range from 0.01—0.1. The cyan points in the background are the labels for all cluster stars used in the training of The Cannon. The DR13 input abundance labels that we adopt for training are already corrected for systematic variations with temperature (Holtzman et al. 2015), except for C and N, which are known to have astrophysical variations along the giant branch (and then additionally corrected as part of our procedure for LSF variation as described in Section 3). For the open clusters, we note that the coolest stars do seem to have the highest measurements of [N/Fe] within a cluster. We find that there are no trends in [Fe/H] along the giant branch for member stars of the clusters, as has been reported for globular clusters. There appears to be a small trend with \( T_{\text{eff}} \) for some elements for the cluster members (e.g., [S/Fe] in Figures 6–12), which we do not correct for. However, in calculating our pair metric from field stars (Section 4.3), we take only stars with very similar \( T_{\text{eff}} \) and log g values to mitigate any residual systematic differences due to these parameters, either astrophysical or otherwise.
Figure 4. Example of one star in the cluster NGC 6791 with $S/N = 260$, with the model in cyan and the data in black, showing that the model is a good fit to the data. The top panel shows a 300 Å wavelength range section from the full 2000 Å H-band APOGEE region. The second panel shows the model scatter across this region. The lower panels show 30 Å ranges of spectra zoomed in around lines of individual elements, highlighting the goodness of fit to the individual element absorption features.

Figure 5. $S/N$-dependent performance for a sample of the elements, showing the measured rms difference of individual visit measurements on the y-axis. All elements are with respect to Fe.
We report the measured abundance mean and variance for each of our 90 open cluster stars in Table 2, along with the corresponding ASPCAP results (where available) and 2MASS IDs. Our results compare very well to the ASPCAP results, but using The Cannon, we obtain substantially higher precision and so can report measurement uncertainties that are 20%–50% lower than ASPCAP in some cases. Here we are concerned with precision and not really with accuracy; there is a

Figure 6. M67 stars with a median S/N = 438, colored by effective temperature. The cyan points show the training data. The two values at the top of each subplot are the mean and standard deviation of the measurements, where all elements are with respect to Fe except [Fe/H].

Figure 7. Same as Figure 6, but for NGC 2420 stars with a median S/N = 311.
discussion in Holtzman et al. (2015) regarding the scale of the open clusters with respect to the literature. Overall, we note that there is a large variation in the reported individual element measurements from high-resolution spectroscopy (e.g., see Table 3 of Souto et al. 2016 for a literature comparison of one of the open clusters, NGC 2420). We point out that our study is for open clusters only. We note that in globular clusters there are well-documented anticorrelations, for example, Na–O and Mg–Al (e.g., Mészáros et al. 2015). Processes such as atomic diffusion and turbulent mixing may also affect the chemical...
abundances in globular clusters, although this is unlikely to be significant along the giant branch (e.g., Korn et al. 2007). This investigation of any such signatures that have currently not yet been reported in the more metal-rich open cluster population is a separate issue.

We note again that our analysis here requires a set of data with precise (not necessarily accurate) abundance measurements, with well-understood associated errors and with critical underlying systematics such as the dependence of the abundance measurements on the FWHM of the LSF removed. We therefore do not

Figure 10. Same as Figure 6, but for NGC 6791 stars with a median S/N = 159.

Figure 11. Same as Figure 6, but for NGC 6819 stars with a median S/N = 303.
need to be concerned with the abundance ground truth of these measurements with respect to the literature.

4.1.1. Rescaling some of the Abundance Uncertainties

The quadrature sum of the formal abundance uncertainty for each star and the cross-validation error (typically 10× larger; see Figure 3) represents the most conservative uncertainty estimate for our measurements, essentially an upper limit on the precision; if these uncertainties were too large, the inferred intrinsic dispersion would be too low.

We check this using the best sampled cluster, NGC 6819, with 28 stars, assuming it is chemically homogeneous; the distribution of measured abundance estimates must not be “narrower” than the abundance uncertainties. Figure 13 shows the abundance distribution for all elements for NGC 6819. The data are shown in the binned histograms, and the cross-validation errors are shown in the red Gaussians. The median formal errors for the stars are shown in the narrow blue Gaussian distributions. Under the assumption that the cluster does not have any intrinsic dispersion in any element, the cross-validation errors appear to be a relatively accurate representation of the measurement uncertainties. However, the uncertainties for some elements, for example, both [Na/Fe] and [V/Fe], look to be overestimated from this figure, comparing the wide red Gaussian distributions to the narrower width of the binned data.

The χ² distribution of pairs of stars within a cluster asks how likely the abundance measurements of a pair of stars would be if the true abundances were identical. We can therefore use the χ² metric, for all pairs of stars within a cluster, to check whether our uncertainties are over- or underestimate. The χ² value is given by the following equation:

\[ \chi^2_{nn'} = \frac{(x_n - x_{n'})^2}{\sigma_{ni}^2 + \sigma_{ni'}^2}, \]

where the indices \( n \) and \( n' \) denote the two stars, \( i \) the elements, and \( x_{ni} \) the measurements with uncertainty \( \sigma_{ni} \).

Assuming that the true photospheric abundances of the cluster stars were identical—a problematic assumption for main sequence stars but a quite good assumption for giants; see, e.g., Dotter et al. (2017)—we expect a distribution of \( \chi^2_{nn} \), that has a mean of 1 and a median of 0.45 for every element. Figure 14 shows the distribution of the \( \chi^2_{nn} \) values for all pairs of stars in our fiducial cluster, NGC 6819, for each of our 20 elements. This figure shows that the median and mean values of the distributions for a number of elements have a mean \( \chi^2 < 1 \) and a median \( \chi^2 < 0.45 \), implying that the measurement uncertainties adopted so far must be overestimates. For those elements, we decrease the uncertainties by the factors that make the mean or median of \( \chi^2 \) match the theoretical expectations.

4.2. Individual Abundance Dispersion within Clusters?

Determining whether or not open clusters have a measurable intrinsic abundance spread is critical for the pursuit of chemical tagging and understanding the formation of these systems (Bovy 2016; Liu et al. 2016a). Liu et al. (2016a) use a differential analysis of high signal to noise, high-resolution spectra to claim that there are inhomogeneities in neutron capture elements within the open cluster M67. Furthermore, Liu et al. (2016b), for example, find that the Hyades open cluster is inhomogeneous at the 0.02 dex level for most of the 19 elements that they measure. Certainly there are physical mechanisms that could account for abundance differences between stars within clusters (e.g., Zahn 1992; Liu et al. 2016b). A large fraction of stars reside in binary systems, and this may affect surface abundance compositions due to rotational evolution and internal mixing as well as mass transfer during the red giant phase. Furthermore, the accretion...
Table 2  
Measured Mean Abundances from The Cannon and from ASPCAP for the Cluster Stars

| Element | NGC 7789 (5 stars) | NGC 6819 (28 stars) | NGC 6791 (16 stars) | NGC 188 (3 stars) |
|---------|---------------------|----------------------|----------------------|-------------------|
|         | ASPCAP              | The Cannon           | ASPCAP               | The Cannon        |
| Fe      | 0.05 ± 0.05         | 0.04 ± 0.02          | 0.02 ± 0.02          | 0.02 ± 0.01       |
| C       | 0.08 ± 0.02         | 0.07 ± 0.02          | 0.06 ± 0.02          | 0.06 ± 0.02       |
| N       | 0.37 ± 0.03         | 0.36 ± 0.03          | 0.33 ± 0.02          | 0.32 ± 0.03       |
| O       | 0.01 ± 0.02         | 0.02 ± 0.02          | 0.02 ± 0.02          | 0.03 ± 0.03       |
| Na      | −0.03 ± 0.0         | 0.01 ± 0.05          | 0.02 ± 0.00          | 0.06 ± 0.03       |
| Mg      | −0.02 ± 0.01        | −0.01 ± 0.01         | 0.00 ± 0.00          | −0.01 ± 0.01      |
| Al      | −0.03 ± 0.01        | −0.05 ± 0.05         | −0.02 ± 0.02         | −0.01 ± 0.02      |
| Si      | −0.02 ± 0.03        | −0.03 ± 0.01         | −0.01 ± 0.04         | −0.03 ± 0.05      |
| S       | 0.02 ± 0.03         | 0.04 ± 0.08          | 0.04 ± 0.06          | −0.03 ± 0.03      |
| K       | 0.11 ± 0.11         | −0.01 ± 0.01         | −0.01 ± 0.03         | −0.03 ± 0.03      |
| Ca      | −0.03 ± 0.01        | −0.01 ± 0.03         | −0.01 ± 0.03         | −0.01 ± 0.03      |
| Ti      | −0.01 ± 0.01        | −0.04 ± 0.02         | −0.03 ± 0.02         | −0.02 ± 0.02      |
| V       | −0.02 ± 0.01        | −0.01 ± 0.02         | 0.01 ± 0.05          | 0.04 ± 0.15       |
| Mn      | −0.02 ± 0.01        | −0.02 ± 0.01         | 0.01 ± 0.01          | 0.04 ± 0.03       |
| Ni      | −0.02 ± 0.01        | 0.01 ± 0.01          | 0.00 ± 0.00          | 0.05 ± 0.02       |
| P       | −0.02 ± 0.03        | 0.00 ± 0.01          | 0.01 ± 0.01          | 0.01 ± 0.01       |
| Cr      | −0.03 ± 0.03        | 0.00 ± 0.03          | 0.02 ± 0.02          | 0.03 ± 0.08       |
| Co      | −0.01 ± 0.06        | 0.03 ± 0.05          | 0.04 ± 0.08          | 0.17 ± 0.09       |
| Cu      | 0.05 ± 0.06         | 0.05 ± 0.02          | 0.17 ± 0.07          | 0.08 ± 0.13       |
| Rb      | 0.06 ± 0.01         | 0.01 ± 0.05          | 0.06 ± 0.04          | −0.01 ± 0.08      |

| Element | NGC 2420 (12 stars) | NGC 2158 (7 stars) | M67 (19 stars) |
|---------|---------------------|---------------------|----------------|
|         | ASPCAP              | The Cannon           | ASPCAP               |
| Fe      | −0.19 ± 0.02        | 0.19 ± 0.03          | 0.02 ± 0.04          |
| C       | −0.06 ± 0.03        | −0.11 ± 0.04         | −0.23 ± 0.04         |
| N       | 0.21 ± 0.04         | 0.25 ± 0.11          | 0.27 ± 0.05          |
| O       | 0.04 ± 0.03         | −0.05 ± 0.11         | 0.04 ± 0.03          |
| Na      | −0.03 ± 0.05        | 0.03 ± 0.08          | 0.01 ± 0.02          |
| Mg      | −0.02 ± 0.01        | −0.03 ± 0.03         | 0.02 ± 0.02          |
| Al      | −0.04 ± 0.04        | 0.01 ± 0.01          | 0.00 ± 0.00          |
| Si      | 0.02 ± 0.03         | 0.02 ± 0.03          | 0.01 ± 0.09          |
| S       | 0.01 ± 0.05         | 0.03 ± 0.03          | 0.01 ± 0.03          |
| K       | 0.07 ± 0.05         | 0.16 ± 0.12          | 0.05 ± 0.03          |
| Ca      | 0.0 ± 0.05          | 0.03 ± 0.04          | 0.02 ± 0.04          |
| Ti      | 0.02 ± 0.02         | −0.01 ± 0.04         | −0.12 ± 0.05         |
| V       | −0.05 ± 0.05        | −0.11 ± 0.11         | −0.14 ± 0.11         |
| Mn      | −0.06 ± 0.01        | −0.07 ± 0.02         | −0.03 ± 0.06         |
| Ni      | −0.02 ± 0.01        | −0.03 ± 0.03         | 0.01 ± 0.01          |
| P       | −0.03 ± 0.09        | −0.01 ± 0.17         | −0.02 ± 0.08         |
| Cr      | −0.05 ± 0.06        | −0.01 ± 0.14         | −0.02 ± 0.04         |
| Co      | −0.15 ± 0.12        | −0.07 ± 0.06         | −0.11 ± 0.06         |

**Note.** All elements are with respect to Fe, [X/Fe], except for Fe which is with respect to [Fe/H]. The values provided in each row are the mean and the 1 − σ dispersion on that mean for the measured values as per Figures 6–12.
or swallowing of planets could dramatically impact chemical homogeneity within clusters (Oh et al. 2017).

With the carefully evaluated measurement uncertainties in hand, we are now in a position to determine these intrinsic spreads. At face value, the dispersion of abundance measurements among stars in a given cluster is comparable to the measurement uncertainties for the individual stars for each element. Qualitatively, this implies that the clusters are nearly homogeneous in their abundances. However, we need to determine formally for each element and for each cluster what the intrinsic dispersion is, given our data.

Calculating the “observed” abundance dispersion of a cluster can be done by simply fitting a Gaussian to the ensemble of abundance estimates (e.g., Ting et al. 2016). But here, we need the “intrinsic” dispersions, which require us to adopt a simple Gaussian model for the intrinsic abundance distribution of the $i$th element within a cluster, which accounts explicitly for the measurement uncertainties:

$$P(x_i) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi (\delta x_i^2 + \sigma_i^2)}} \cdot \exp\left(-\frac{(x_{i,n} - \bar{x}_i)^2}{2(\delta x_i^2 + \sigma_i^2)}\right)$$  

Here, the mean abundance of each element is $\bar{x}_i$ and the intrinsic (error-corrected) abundance dispersion of each element within a cluster is $\sigma_i$. Our observational information is $\{x_{i,n}, \delta x_{i,n}\}$. Of course, the intrinsic dispersion must be positive definite. Then, we are in a position to calculate the probability density function (pdf) for the intrinsic dispersion in each cluster.

We also optimize the right-hand side of Equation (2) over a range of sigma-clipping values of the input data ($x_{i,n}, \delta x_{i,n}$) and take the maximal value. The typical optimal sigma clipping is $\sigma_{\text{clip}} > 3$. This means that only a few stars are excised by sigma clipping and for just a few of the elements. For clusters M67, NGC 2420, NGC 6791, NGC 6819, and NGC 7789, between one to three stars were excised, for only a few elements—as few as one element (in NGC 2420) and as many as only six elements (in NGC 6819).

We should explore the impact of the slightly different choices for the measurement uncertainties, discussed in Section 4.1.1: (i) where we scale this error by the median scaling value calculated for each element for the $\chi^2$ distribution shown in Figure 14, (ii) as (i) but scaling by the mean, and (iii) leaving the uncertainties $\delta x_{i,n}$ to be the quadrature sum of the formal uncertainty from The Cannon and the signal-to-noise dependent cross-validation uncertainty for that element. We find that all three error estimates give almost identical results, and we show here the results for the median scaling. We use the median scaled errors for all of the analysis that follows in the paper.

Basically, we find throughout that the intrinsic abundance mean is very close to the mean of the measurements, as the abundance uncertainties among different stars within a cluster are similar. As the pdf’s for the intrinsic abundance dispersions are by construction restricted to non-negative $\sigma_i$, we characterize them by two numbers, the most likely dispersion and the
median of the pdf. This is shown in Figure 15: in most clusters and for most elements, the intrinsic abundance dispersions are consistent with zero and likely to be nearly zero. The medians of the intrinsic dispersion pdf's (Figure 15) show that for all clusters with \( >10 \) stars (enabling robust dispersion estimates), the intrinsic dispersions are typically \(<0.02\) dex, and \(<0.04\) dex for almost all of the 20 elements. There are only a few stars with spectra in NGC 2158, NGC 7789, and NGC 188, and the measurements for NGC 188 in particular are not robust on the basis of small sample size. NGC 188 has three stars, and one is a clear outlier in some elements which drives the large intrinsic dispersion in this cluster.

Our analysis points toward a small but finite intrinsic dispersion in \([Fe/H]\) for two clusters with robust measurements (NGC 6819 and M67), at the level of \( \sigma = 0.02 \) dex; this is true even if the abundance uncertainties were overestimated by 20\%. The calculated probability of a zero dispersion value compared to the most likely value is only \( P_{pdf}(\sigma_{[Fe/H]=0}) / P_{pdf}(\sigma_{[Fe/H]_{\text{intrinsic}}}) = 0.07 \). In general, these intrinsic abundance dispersion limits (or measurements) for clusters are roughly consistent with the limits placed on the elements by Bovy (2016). But here they are based on direct abundance determinations. Our finding of this small intrinsic dispersion in \([Fe/H]\) in particular merits further investigation, and we expect to obtain further observations of these cluster members to follow up this result.

4.3. Overall Abundance Similarity of Stars within Clusters and in the Field

Based on these data, we can try to set some expectation for how clearly abundances can tell us when stars in the APOGEE sample are, or are not, born together in the same cluster or molecular cloud core. That is, assuming that stars born together in the same cluster or molecular cloud core are chemically homogeneous. We do this by examining the (dis)similarity of stars within clusters and in the field. This is relevant for assessing the much larger sample of APOGEE data and chemical dimensionality and diversity of the Milky Way disk. We note that our sample of open clusters does not span the full range of properties seen in the Galactic disk (i.e., their distribution in \([Fe/H]\), age, and Galactocentric radius). However, these clusters do have an abundance distribution that broadly resembles that of the APOGEE field (disk) red clump sample.

To explore the similarity (over all or many elements) in abundance space for stars within clusters and among field stars, one may be tempted to devise a simple pseudo-Cartesian measure of distance in abundance space, such as \( D_{\text{cluster}}^2 = \sum_{i} (x_{ni} - x_{ni})^2 \), and then compare the distributions \( p(D_{\text{cluster}}^2) \) and \( p(D_{\text{field}}^2) \). However, with imperfect measurements \( x_{ni} \), the use and interpretation of scalar distance measurements in a high-dimensional space is complex and depends strongly on the choices of prior assumptions, whenever the distances become comparable to the measurement uncertainties, as is the case here; we sketch this
issue out in the Appendix A. But it is straightforward to assess, via \( \chi^2 \), how likely the abundance measurements of a pair of stars would be, if the true abundances were identical:

\[
\chi^2_{nn'} = \sum_{i=1}^{I} \frac{(x_{ni} - x_{n'i})^2}{\sigma_{ni}^2 + \sigma_{n'i}^2}.
\]  

The two star indices are \( n \) and \( n' \), and each star \( n \) has a measurement \( x_{ni} \) and each star \( n' \) has a measurement \( x_{n'i} \), for elements \( i = 1 \) to \( I \), with uncertainty \( \sigma_{ni} \) and \( \sigma_{n'i} \), respectively. We adopted the median-rescaled uncertainties from Section 4.1.1 here. For each star, we determine the neighbor most likely to have identical abundances by minimizing \( \chi^2_{nn'} \).

This \( \chi^2_{nn'} \) metric is then calculated for star pairs within the same cluster as well as for the stars pairs drawn from a random sample of 2000 field stars, which have analogously determined abundances (and uncertainties).
Figure 16. Top: $\chi^2$ distribution of abundance differences, $P_{\text{pair}}(\chi^2)$, for pairs with similar log $g$ and $T_{\text{eff}}$ values (see Section 4.3). The black histogram shows $P_{\text{pair}}^{\text{intra}}(\chi^2)$ for all intracluster pairs (600 pairs); the red dashed histogram shows the analogous distribution $P_{\text{pair}}^{\text{field}}(\chi^2)$ for all field pairs (1,018,581 pairs). The intracluster and field pair distributions are clearly very different, but they are not disjoint: 0.3% of field pairs have $\chi^2$ differences as small as the median $\chi^2$ among intracluster pairs: most of these field pairs are presumably not birth siblings, but doppelgängers. The bottom panel shows the analogous distributions $P_{\text{pair}}(\chi^2)$ to Figure 16, but restricted to pairs of the same [Fe/H], and the $\chi^2$ arises from the sum over the 19 $[X/Fe]$ estimates; the distribution is again restricted to pairs with similar log $g$ and $T_{\text{eff}}$ values (see text). The [Fe/H] is set by the (near-identical) metallicity of the two clusters M67 and NGC 6819. The $\chi^2$ distribution of the 377 intracluster pairs is shown by the black histogram. The distribution $P_{\text{pair}}^{\text{field}}(\chi^2)$ from 1054 random field stars with $[\text{Fe/H}] = 0 \pm 0.02$ and similar temperatures and gravities (301,587 pairs) are shown in the red dashed histogram. The intracluster distribution for M67 and NGC 6819 (of 327 pairs) is shown for comparison by the blue dashed-dotted histogram. Even if we select stars of the same [Fe/H], the other element abundances can still discriminate between intracluster and field pairs in many instances: the peaks of the distributions are clearly different with the highest similarity among stars within a cluster and the largest dissimilarity among field stars. However, once again, these distributions are not disjoint: 1.0% of field pairs at solar metallicity have $\chi^2$ differences as small as the median $\chi^2$ among intracluster pairs; these stars are doppelgängers.

The field stars were chosen to have a log $g < 3.9$ (giants) and an S/N > 200 to be comparably high in S/N to the cluster sample. The results for all pairs, i.e., the distributions $P_{\text{pair}}^{\text{cluster}}(\chi^2)$ and $P_{\text{pair}}^{\text{field}}(\chi^2)$, are shown in the top panel of Figure 16. For these pairs, we restricted comparisons to within $\Delta T_{\text{eff}} < 315$ K and $\Delta \log g < 0.7$ to guard against differences in these “nuisance labels” leading to systematic abundance differences; we found that beyond these limits, the $\chi^2$-distance in abundance space is correlated with these parameter differences (see Figure 21 in Appendix C). The $P_{\text{pair}}(\chi^2)$ for pairs from the same cluster are shown in black and pairs of 2000 random field stars are shown in red (dashed line).

Figure 16 demonstrates that in the 20 element abundance space, stars within clusters are, unsurprisingly, chemically far more similar than stars pairs among the field sample. For pairs within a cluster, $P_{\text{pair}}^{\text{cluster}}(\chi^2)$ peaks at ~20, as expected if all stars in a cluster had identical abundances in all elements; however, there is a significant tail of $P_{\text{pair}}^{\text{cluster}}(\chi^2)$ toward substantially higher values of $\chi^2$, implying that some stars differ in some elements even within a cluster. Note that we have included the elements C and N in this comparison, even though their photospheric abundances get altered by dredge-up, to a degree that depends in giants on the mass or age. For those elements, similarity implies a combination of near-identical birth material and age, as expected for open clusters.

For field star pairs, $P_{\text{pair}}^{\text{field}}(\chi^2)$ is far broader, which is unsurprising, as members of a random field star pair will usually differ even in their most elementary abundance, [Fe/H]. The vast majority of the $P_{\text{pair}}^{\text{field}}(\chi^2)$ values lie at $\chi^2$ values far in excess of $\chi^2 \sim 20$. However, there is a small fraction of field pairs (~1%–2%) whose 20 element abundances are indistinguishable, $\chi^2 \lesssim 25$, despite the 0.02–0.03 dex precision that APOGEE affords for individual elements.

An obvious next question to ask is whether the differences between $P_{\text{pair}}^{\text{cluster}}(\chi^2)$ and $P_{\text{pair}}^{\text{field}}(\chi^2)$, shown in Figure 16, are primarily driven by differences in the basic [Fe/H] rather than the high-dimensional [X/Fe]. To do this, we drew up the distributions $P_{\text{pair}}^{\text{cluster}}(\chi^2)$ and $P_{\text{pair}}^{\text{field}}(\chi^2)$, but restricted to solar [Fe/H]. For the intracluster pairs, we consider only two clusters, M67 ([Fe/H] = 0.0) and NGC 6819 ([Fe/H] = 0.03); for this second selection of field pairs, this meant we restricted the sample to red giant stars with $[\text{Fe/H}] = 0.00 \pm 0.02$. A total of 1054 solar-metallicity field giants (again selected with S/N > 200) were used for this comparison. The resulting distributions are shown in the bottom panel of Figure 16. This figure shows that the two distributions (intracluster and field pairs) remain distinctly different: there is valuable discriminating information in the [X/Fe]. However, considering a priori only pairs of the same metallicity naturally increases the overlap between the distributions substantially.

In Figure 16 there are many more pairs (especially intracluster pairs) at very small values of $\chi^2$, far more than expected from the chi-squared distribution with 19 or 18 degrees of freedom. This could have multiple origins, related to our chemical-abundance uncertainty model. Our chemical-space uncertainty analysis is fundamentally empirical; it presumes that all stars are
was calculated above and removed from the determination of the doppelgänger rate, which is as small as the median \( \chi^2 \) among intracluster pairs. An additional constraint for stars that are born together is that these stars must have the same metallicity. Therefore, to assess the implications for the viability of strict chemical tagging, we should consider the abundance similarity of stars at a single metallicity, as shown in the bottom panel of Figure 16.

For the stars at a single (solar) metallicity, 1.0% of field pairs have \( \chi^2 \) differences as small as the median \( \chi^2 \) among intracluster pairs; these stars are doppelgängers.

Excluding the similarity of the field pairs at the same metallicity of [Fe/H] = 0, probable new members of two clusters were discovered, as detailed in Section 6. Although negligible in number relative to all of the pair combinations, these were removed from the determination of the doppelgänger rate, which was calculated above (1.0% at [Fe/H] = 0). The new probable members identified demonstrates that, with the \( \chi^2 \) pair similarity analysis, we have developed an effective technique to identify new cluster members when combined with radial velocity measurements. These members could have also been identified using their overall [Fe/H], radial velocities, spatial proximities, and proper motions.

5. The Fraction of Galactic Doppelgängers

Figure 16 characterizes the distribution of the abundance similarity of stars that are known to be born together, \( p_{\text{pair}}(\chi^2) \), compared to that of random field stars, \( p_{\text{field}}(\chi^2) \).

From this, we can determine the relative fraction of stars that appear as chemically similar as cluster stars do to one another, yet are not born in the same environment, the so-called doppelgängers.

For the field, 0.3% of random pairs have a \( \chi^2 \) difference that is as small as the median \( \chi^2 \) among intracluster pairs. An additional constraint for stars that are born together is that these stars must have the same metallicity. Therefore, to assess the implications for the viability of strict chemical tagging, we should consider the abundance similarity of stars at a single metallicity, as shown in the bottom panel of Figure 16. For the stars at a single (solar) metallicity, 1.0% of field pairs have \( \chi^2 \) differences as small as the median \( \chi^2 \) among intracluster pairs; these stars are doppelgängers.

6. Identification of New Cluster Members

We examined all “sibling” field pairs that are as chemically similar as cluster pairs to ensure that they are doppelgängers and not true cluster members that are simply not included in our list of analyzed clusters. From this, we found four new potential cluster members (all with \( \chi^2 < 7 \))—one pair associated with M67 and one pair associated with NGC 7789.

We found these pairs only by using their chemical similarity to each other (as similar as that of the intracluster pairs) and then by observing that they had a common velocity and were in proximity to open clusters of the same [Fe/H] and mean radial velocity as the cluster.

The properties of the two newly identified member pairs from the field are summarized in Table 3 (Appendix B). The positions of these stars and their radial velocities along with the cluster center and surrounding field stars are shown colored by radial velocity in Figure 17. The stars associated with M67 have velocities of \( -53.6 \) and \( -54.9 \) km s\(^{-1} \), and the stars we associate with M67 have velocities of \( 34.2 \) and \( 34.1 \) km s\(^{-1} \). These velocity measurements are all within the respective cluster velocity dispersions: the mean velocity and velocity dispersion of M67 is \( 53.6 \pm 0.8 \) km s\(^{-1} \) (Geller et al. 2015). The mean velocity and velocity dispersion of NGC 7789 is \( 54.9 \pm 0.9 \) km s\(^{-1} \) (Gim et al. 1998).

These pairs were excluded from our doppelgänger rate determination in Section 5, as they are potentially actual cluster members. We emphasize that these were not identified as members of the clusters via a test of the similarity of their 19 abundances to the clusters themselves.

These pairs have abundance measurements that are generally consistent with those of the clusters, with some exceptions. These exceptions in particular, in the case of M67, lead to a
high $\chi^2$ (>20) between the potential new members of M67 and the known M67 member stars. Based on a comparison of the abundances of these two stars with the known M67 stars alone, these would not be considered members, under the assumption of near-zero intrinsic abundance spread. These differences however, may be associated with temperature and log $g$ systematics—including those of an astrophysical origin (see the Appendix C). The abundances in all 20 elements of the potential new member stars are shown with the mean values for the M67 and NGC 7789 clusters in Figure 18. The mean $T_{\text{eff}}$ and log $g$ of the M67 cluster stars is $4770 \, \text{K}$ and $2.8 \, \text{dex}$, respectively, and the mean abundances are subsequently pivoted around these mean parameter values. The pair of stars we associate with M67 are hotter and have higher gravities, at $\sim 5200 \, \text{K}$ and $3.7 \, \text{dex}$ (see Table 3 and Figure 20 in Appendices B and C, respectively). This is a temperature difference of $430 \, \text{K}$ and $0.9 \, \text{dex}$, which (see Figure 21 in Appendix C) is a high enough difference in parameters that some abundances are systematically offset in magnitude. This is why the pair metric comparing stars with a restricted difference in $T_{\text{eff}}$ and log $g$ is so powerful—it circumvents these systematics. Although the high $\chi^2$ is particularly driven by a few outlying elements in particular (e.g., [N/Fe] and [Mg/Fe]), even with the largest outliers removed from the $\chi^2$ comparison, the new potential members remain at a larger $\chi^2$ distance from the M67 members compared to the intracluster pair comparison. For the pair associated with NGC 7789, with temperature and log $g$ values around $4900 \, \text{K}$ and $2.7 \, \text{dex}$ (compared to the cluster mean of $4600 \, \text{K}$ and $2.3 \, \text{dex}$), three of the NGC 7789 cluster stars are within $\chi^2 < 19$ of these potential new members. The proper motions, included in Table 3, are consistent with these four stars being open cluster members.

7. Discussion and Conclusion

In this paper, we set out to quantify the similarity among Galactic disk stars with respect to their detailed photospheric element abundances to clarify the prospects of chemical tagging with data of APOGEE’s quality. For such an analysis, we could draw on a unique abundance data set, derived by applying a modified version of The Cannon to APOGEE spectra to ensure maximal abundance precision and LSF. This reanalysis allowed us to recognize and eliminate a likely source of subtle but correlated systematic errors in the abundance estimates. We found this systematic when examining the field APOGEE giant sample, noticing that the most chemically similar pairs of field stars often had small differences in fiber number. The fiber number correlation is presumably just a proxy for the FWHM of the LSF, which is intrinsic to the experimental setup, and we subsequently use this in the paper. Such systematics are very important to remove to use the abundance measurements in concert. The quantification of the abundance uncertainties included S/N-dependent cross-validation and an error rescaling for some elements, exploiting the fact that abundance estimates for member stars of open clusters should on average not differ by less than their uncertainties. This provided a set of APOGEE-derived abundances that is unprecedented in its (pertinent) combination of quality, quantity, and homogeneity. We only considered stars with very similar stellar parameters, giants in a restricted $T_{\text{eff}}$ and log $g$ range; we have 90 spectra across seven open clusters and thousands of “field stars” with the same data, S/N, and data analysis. For all of these stars, we have individual abundances for 20 elements (Fe, C, N, O, Na, Mg, Al, Si, S, K, Ca, Ti, V, Mn, Ni, P, Cr, Co, Cu, and Rb), with a median precision of
0.03 dex. We characterize these abundances by their values \([\text{X}/\text{Fe}]\), except for iron, where we use \([\text{Fe}/\text{H}]\).

On this basis, we undertook an extensive analysis of the chemical (in-)homogeneity of stars within an open cluster, testing the standard assumption that such stellar “birth siblings” were formed at the same time from chemically homogeneous material. This should result in identical element abundances, including those elements (e.g., C and N), whose photospheric abundance has been altered in giants by mass/age-dependent dredge-up. We confirm that the abundances of cluster member stars typically agree with each other within their small (0.03 dex) uncertainties, as had been found before (see Section 1). Moving beyond, we explicitly determined the intrinsic dispersion in these abundances, or derived upper limits for them. We found that (at least in the best sampled clusters), the abundance dispersions are essentially zero, with typical upper limits of \(\sim 0.03\) dex. However, there are exceptions; some elements in some clusters show small but significant dispersion, attributable to a subset of member stars.

To quantify the abundance similarity among stars of known common birth origin (cluster members) to the (dis-)similarity among field stars, we compared the abundances among pairs of stars. It was tempting to define a distance measure in the high-dimensional space of element abundances, but the “curse of dimensionality” makes distance in a 20 dimensional space very dependent on prior assumptions about distance distributions once the distances become comparable to the measurement uncertainties. Therefore, we resorted to a simple \(\chi^2\) statistic, \(p_{\text{pair}}(\chi^2)\), quantifying the likelihood of the 20 observed abundances (differences) for any pair of stars, assuming their intrinsic abundances were identical.

When we construct \(p_{\text{pair}}(\chi^2)\) for intracluster pairs of stars, we find that its median is similar to \(N_{\text{elements}} \approx 20\), again illustrating the level of chemical homogeneity within clusters. Yet, about 20% of intracluster star pairs have \(\chi^2 \geq 40\), indicating significant abundance differences in one or more elements. We find open clusters to be mostly, but not exclusively, homogeneous.

When we construct \(p_{\text{pair}}(\chi^2)\) for pairs of field stars, it looks overall very different. The vast majority of this distribution lies at \(\chi^2 \geq 40\); most pairs of field stars can be clearly recognized as having differing abundances. In particular, 99.7% of the field pairs have \(\chi^2\) in excess of the median \(\chi^2\) for intracluster pairs. Part of why \(p_{\text{pair}}(\chi^2)\) looks so different from \(p_{\text{pair}}(\chi^2)\) is of course that the Galactic (thin) disk has a metallicity spread of about 1 dex, but considering the analogous distributions restricted to solar metallicity shows that the remaining abundances still have great discriminating power. Most commonly, the abundances of field pairs—at the same \([\text{Fe}/\text{H}]\) (e.g., \([\text{Fe}/\text{H}] < 0.02\) are inconsistent with being identical; 99% of field star pairs of the same metallicity appear chemically non-identical on the basis of their other element abundances with APOGEE data.

In turn, however, the above result implies that 0.30% of random pairs among Galactic disk stars have indistinguishable abundances, even with the typical 0.03 dex measurement precision for \(\sim 20\) elements that APOGEE data afford; this has remarkable consequences. Such pairs are much more common than any plausible incidence of common-birth-site pairs (former cluster), \(10^{-4} \ldots 10^{-6}\). Therefore, most of them cannot be birth siblings, but are mere doppelgängers: stars not immediately related by birth, yet looking near-perfectly alike in their abundance patterns. This rate of field star doppelgängers has not been quantified before and has important implications for chemical tagging. Note that this rate estimate applies to stars with typical Galactic disk abundances, \(-0.7 \leq [\text{Fe}/\text{H}] \leq 0.3\) and \([\alpha/\text{Fe}] \sim \text{solar}\), where the abundance space is most densely populated with stars. In the metal-poor old disk or the halo, doppelgängers may be much rarer (Ting et al. 2015), as abundances show greater diversity, and therefore the abundance space is more sparsely populated. It will be interesting to explore whether the inclusion of extensive sets of \(s\)- and \(r\)-process elements (e.g., from Galah) will change this picture qualitatively.

It is also worth noting that the \(\chi^2\) statistic that we use here—while extremely straightforward—is unlikely to be the optimal way to discriminate cluster pairs from field pairs. In particular, there are elements (foremost \([\text{Fe}/\text{H}]\) and \(\alpha\)-elements) in which the abundance variation among field pairs is large compared to our measurement precision. Conversely, there may be elements (when viewed as \([\text{X}/\text{Fe}]\) where even random field pairs typically differ by barely more than the measurement uncertainties (see Figure 19). Those elements are then rather uninformative when it comes to discriminating the two hypotheses and may mostly add variance in the \(\chi^2\) sum. This suggests that a sum of the individual element differences, without weighting and covariance terms, is likely sub-optimal. However, tests on our current set of elements, taking only a subset of the most informative elements for our \(\chi^2\) calculation, indicate there is no, or at least no dramatic, gain expected from removing the least informative elements. A rigorous development of a near-optimal statistic to discriminate birth siblings from field pairs (i.e., minimizing the doppelgänger rate) requires a full characterization of the (error-deconvolved) 20-dimensional abundance distribution of the field (see also Ting et al. 2015). This is beyond the scope of this paper. However, simple experiments to restrict the \(\chi^2\) sum of the abundance differences to subsets of the most discriminating elements showed that the basic picture drawn up here is unlikely to change; the existence of an important doppelgänger population in APOGEE-like data is not a consequence of the sub-optimal \(\chi^2\) statistic.

To summarize the assessment of abundance (\(\chi^2\)) similarities or differences, with 20 abundances at a typical 0.03 dex precision level in hand: most—but not all—pairs of stars in clusters are chemically alike; most—but not all—pairs of field stars are chemically different.

In conclusion, we can now discuss what these findings mean for “chemical tagging.” Such studies of chemical similarity among stars in the context of Milky Way evolution have claimed two main goals. One is an exploration of the Galactic history of chemically homogeneous birth sites of stars, i.e., the history of the (now disrupted) cluster mass function (Ting et al. 2016). This “strict chemical tagging” would require attributing a common birth origin to, say, pairs of stars with considerable confidence, on the basis of abundances alone. Even in optimistic scenarios for the main stellar disk of the Milky Way, only \(10^{-4.5}\) random star pairs would be such birth siblings (our doppelgänger measurement is equivalent to \(10^{-2}\)). The other goal is the empirical derivation of a successively more detailed chemo-orbital distribution of stars in the Galaxy: stars of which abundances are on which orbits. Even considering only \([\text{Fe}/\text{H}]\) and \([\alpha/\text{Fe}]\), it is well known that the Galactic disk structure varies distinctly (Bovy et al. 2012; Hayden et al. 2014). Therefore, more intricate patterns, involving more abundances and ages, could and should
provide better constraints on radial migration or heating. Such “broad chemical tagging” (or chemical labeling) studies take detailed abundances foremost as a lifelong label of stars, independent of orbital phase, without making an explicit reference to a common birth site among pairs or groups of stars.

Our results presented here, in particular the significant incidence of doppelgänger stars in the field, imply that strict chemical tagging is presumably not possible, certainly not for the Milky Way’s main stellar disk components, even with the data quality level presented here. When considering whether a pair of stars was likely born in the same site at the same time, one must consider not only the data likelihoods presented above, but also the prior expectations for being “birth siblings” or doppelgängers, \( P_{\text{sib}} \) and \( P_{\text{dgr}} \), respectively. Then, we have

\[
\frac{p(\text{sib} | \text{data})}{p(\text{dgr} | \text{data})} = \frac{p(\text{data} | \text{sib})}{p(\text{data} | \text{dgr})} \times \frac{p_{\text{prior}}(\text{sib})}{p_{\text{prior}}(\text{dgr})},
\]

where

\[
p(\text{data} | \text{sib})/p(\text{data} | \text{dgr}) \approx p^\text{cluster}(\text{data})/p^\text{field}(\text{data})
\]

from Figure 16, and the ratio of priors is \( p_{\text{prior}}(\text{sib})/p_{\text{prior}}(\text{dgr}) = 10^{-4.4\ldots6} \). As Figure 16 shows, \( p^\text{cluster}(\text{data})/p^\text{field}(\text{data}) \) have basically the same support and that the ratio of priors is so small, it seems hard to imagine to get \( p(\text{sib} | \text{data})/p(\text{dgr} | \text{data}) \geq 1 \).

Although this analysis may dampen the prospects for strict chemical tagging, it is good news for detailed chemical labeling; the analysis shows that the Galactic stellar disk population can be “sliced” into much more detailed abundance-based ways, which will provide new, and probably powerful, ways to constrain the formation of the Milky Way. The C- and N-based age estimates are just one recent example of this approach. Furthermore, combining velocity information (and eventually full dynamics) with the abundance analysis demonstrates promising prospects for the identification of cluster members.

The authors thank Andy Gould for insightful discussions. The authors thank Baitan Tang for a review of the manuscript.

M.N. and H.-W.R. acknowledge funding from the European Research Council under the European Union’s Seventh Framework Programme (FP 7) ERC Advanced Grant Agreement n. [321035]. H.-W.R. acknowledges support of the Miller Institute at UC Berkeley through a visiting professorship during the completion of this work.

D.W.H. was partially supported by the NSF (grants IIS-1124794 and AST-1517237), NASA (grant NNX12AI50G), and the Moore-Sloan Data Science Environment at NYU.

D.G. and D.M. gratefully acknowledge support from the BASAL Centro de Astrofísica y Tecnologías Afines (CATA) grant PFB-06/2007. D.M. is also supported by the Ministry for the Economy, Development, and Tourism, Programa Iniciativa Científica Milenio grant IC120009, awarded to the Millennium Institute of Astrophysics (MAS), and by FONDECYT No. 1130196.

P.M.F gratefully acknowledges support from the National Science Foundation through award AST1311835.

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 http://www.sdss.org.

SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPG Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatario Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autonoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

Appendix A

Distances in High-dimensional Spaces

In the context of chemical tagging, it is tempting to assign a scalar or vector measure of abundance-space distance, \( D \) (or \( \hat{D} \)), to pairs of stars. Such a measure of distance would have to be derived from abundance measurements with uncertainties. In this appendix, we show why the interpretation of such distance measures becomes problematic in high-dimensional spaces, as soon as the true or presumed distances become comparable to the measurement uncertainties. This is why in the analysis presented in this paper, we restrict ourselves to asking how likely the abundance data on a pair of stars are (in the \( \chi^2 \)-sense) if their abundances were identical.

Let us first consider how high-dimensional the space of abundances is; already there is no simple unique answer. Several measures of dimensionality may play a role here: in principle, the dimensionality of abundance space is the length of the periodic system (and then, there are isotopes); in practice, measurements only provide constraints on a subset of elements, between 1 and 35 elements; finally, there is the question of how many astrophysically non-degenerate dimensions the abundance space has.

In principle, it is straightforward to assign a metric to abundance space, e.g., some single Cartesian measure of distance between two points in \( L \)-dimensional abundance space:

\[
d_{\text{cart}} = \sum_{i=1}^{L} D_i^2, \quad \text{with} \quad D_i = |X_{a,i}/H| - |X_{b,i}/H|.
\]

However, in practice, we would need to make a probabilistic statement about the abundance-space distance between two stars, given a set of abundance measurements for those two stars. To get a pdf for any measure of distance, one would have to specify a prior expectation for that distance measure, which then gets modified by the data likelihood. But if one wants to have a pdf for a simple scalar/1D distance measure in light of independent measurements in a high-dimensional abundance space, it appears difficult to avoid the fact that the resulting \( p_{\text{pdf}}(D_{\text{scalar}}) \) depends very strongly on the priors, if the intrinsic distance becomes comparable to the measurement uncertainties; this is the pertinent regime for chemical tagging.
To illustrate this, we will spell out the case for a one-dimensional abundance measurement specifically, and then discuss how to generalize them. For a pair of stars, we have measurements of their abundances,  and , with their measurement uncertainties and . We presume that their true abundances can be characterized by and , as , with . We then get
\[
p_{\text{pdf}}(D, \bar{x} | x_n, x_{n'}, \sigma_n, \sigma_{n'}) \\
= p_L(x_n, x_{n'} | D, \bar{x}, \sigma_n, \sigma_{n'}) \cdot p_{\text{prior}}(D, \bar{x}).
\]

We can now spell out the data likelihood,
\[
p_{L}(x_n, x_{n'} | D, \bar{x}, \sigma_n, \sigma_{n'}) \\
= \frac{1}{2\pi \sigma_n \sigma_{n'}} \left\{ \exp \left[ -\frac{(x_n - \bar{x} + D/2)^2}{2\sigma_n^2} + \frac{(x_{n'} - \bar{x} - D/2)^2}{2\sigma_{n'}^2} \right] \\
+ \exp \left[ -\frac{(x_n - \bar{x} - D/2)^2}{2\sigma_{n'}^2} + \frac{(x_{n'} - \bar{x} + D/2)^2}{2\sigma_n^2} \right] \right\}.
\]

when one integrates over (with a flat prior in ) and simplifies, this becomes
\[
p_{L}(x_n - x_{n'} | D, \sigma_n, \sigma_{n'}) = \frac{1}{\sqrt{2\pi(\sigma_n^2 + \sigma_{n'}^2)}} \\
\times \exp \left( -\frac{D^2 + (x_n - x_{n'})^2}{2(\sigma_n^2 + \sigma_{n'}^2)} \right) \times \cosh \left( \frac{D(x_n - x_{n'})}{\sigma_n^2 + \sigma_{n'}^2} \right).
\]

For the hypothesis that the abundances are identical, i.e., , this becomes
\[
p_{L,D=0}(x_n - x_{n'} | \sigma_n, \sigma_{n'}) \propto \frac{1}{\sqrt{2\pi(\sigma_n^2 + \sigma_{n'}^2)}} \\
\times \exp \left( -\frac{(x_n - x_{n'})^2}{2(\sigma_n^2 + \sigma_{n'}^2)} \right),
\]

corresponding to the simple expression from Equation (2), describing the likelihood of the data under the hypothesis that .

An extension of Equation (5) to the -dimensional case, say , brings us to a well-established conundrum. If we stick to the seemingly natural prior for all , then this implies as a prior for the -distance . The same problem arises for any prior for that is flat for small . Like a distribution, the inferred most likely (scalar) distance would always tend to the squared sum of the measurement uncertainties.

Appendix B
New Cluster Members

Table 3 presents newly identified NGC 7789 and M67 cluster members.

## Mixing of Carbon and Nitrogen along the Giant Branch

Figure 20 shows both the known cluster stars and our new members in the and plane for M67. For the pair comparisons, the stars are restricted to being near in and . However, for the comparison shown in Figure 18, while the pairs are themselves similar, they are not required to be similar to the cluster stars from which the mean abundance values are determined. There are stars that are newly identified members at the hot end of the giant branch, and these stars have a [N/Fe] measurement that falls outside the 1σ mean cluster measurements (calculated from stars at intermediate and values along the giant branch). There are also small discrepancies in a few of the elements in M67 for the hot stars at the base of the giant branch compared to the mean abundance of M67 stars, and this is likely due to small systematic dependences in temperature from the stellar model abundance determinations propagated by The Cannon (see the caption of Figure 18).
Temperature and log g values of M67, including both already determined APOGEE members used for our analysis and the new members identified by measuring abundance similarities between field pairs, colored by their [N/Fe] (left) and [C/Fe] values (right). The [C/Fe] and [N/Fe] ratios in the stellar atmosphere change up the giant branch as there is mixing from the interior of the star. The two newly identified M67 stars, which are outliers in a few of the abundances compared to the mean M67 cluster measurements, most notably for [N/Fe], are at the base of the giant branch and at hotter temperatures and higher gravities than the stars in the cluster from which mean results have been determined. These two stars would not have yet undergone the first dredge-up.

Figure 20.

Field pairs shown with their median \( \chi^2 \) abundance differences as a function of \( \Delta T_{\text{eff}} \) and log g. For these field pairs, we restricted comparisons to within \( \Delta T_{\text{eff}} < 315 \text{ K} \) and \( \Delta \log g < 0.7 \) (shown with vertical lines in the figures) to guard against differences in these “nuisance labels,” leading to systematic abundance differences; we found that beyond these limits the \( \chi^2 \) distance in abundance space is correlated with these parameter differences, as shown above.

Figure 21.

ORCID iDs

H-W. Rix @ https://orcid.org/0000-0003-4996-9069
David W. Hogg @ https://orcid.org/0000-0003-2866-9403
A. R. Casey @ https://orcid.org/0000-0003-0174-0564
J. Holtzman @ https://orcid.org/0000-0002-9771-9622
D. Geisler @ https://orcid.org/0000-0002-3900-8208
M. Shetrone @ https://orcid.org/0000-0003-0509-2656
D. Minniti @ https://orcid.org/0000-0002-7064-099X
Peter M. Frinchaboy @ https://orcid.org/0000-0002-0740-8346
Alexandre Roman-Lopes @ https://orcid.org/0000-0002-1379-4204

de Silva, G. M., Gibson, B. K., Lattanzio, J., & Asplund, M. 2009, A&A, 500, L25
Dotter, A., Conroy, C., Cargile, P., & Asplund, M. 2017, ApJ, 840, 99
Freeman, K., & Bland-Hawthorn, J. 2002, ARA&A, 40, 487
Freeman, K. C. 2012, in ASP Conf. Scr. 458, Galactic Archaeology: Near-Field Cosmology and the Formation of the Milky Way, ed. W. Aoki et al. (San Francisco, CA: ASP), 393
Frinchaboy, P. M., Thompson, B., Jackson, K. M., et al. 2013, ApJL, 777, L1
Gao, X.-H. 2016, RAA, 16, 184
García Pérez, A. E., Allende Prieto, C., Holtzman, J. A., et al. 2016, AJ, 151, 144
Geller, A. M., Latham, D. W., & Mathieu, R. D. 2015, AJ, 150, 97
Gilmore, G., Randich, S., Asplund, M., et al. 2012, Msngr, 147, 25
Gim, M., Hesser, J. E., McClure, R. D., & Stetson, P. B. 1998, PASP, 110, 1172
Hayden, M. R., Holtzman, J. A., Bovy, J., et al. 2014, AJ, 147, 116
Ho, A. Y. Q., Ness, M. K., Hogg, D. W., et al. 2017, ApJ, 836, 5
Hogg, D. W., Casey, A. R., Ness, M., et al. 2016, ApJ, 833, 262
Holtzman, J. A., Shetrone, M., Johnson, J. A., et al. 2015, AJ, 150, 148
Korn, A. J., Grundahl, F., Richard, O., et al. 2007, ApJ, 671, 402
Kunder, A., Kordopatis, G., Steinmetz, M., et al. 2017, AJ, 153, 75
Lambert, D. L., & Reddy, A. B. S. 2016, ApJ, 831, 202
Liu, F., Yong, D., Asplund, M., et al. 2016, ApJ, 825, 146
Martell, S. L., Shetrone, M. D., Lucatello, S., et al. 2016, ApJ, 825, 146
Márquez, I., Fouesneau, M., Rix, H.-W., et al. 2016, MNRAS, 465, 3655
Mészáros, S., Holtzman, J., García Pérez, A. E., et al. 2013, AJ, 146, 133
Mészáros, S., Martell, S. L., Shetrone, M., et al. 2015, AJ, 149, 153

References

Bland-Hawthorn, J., Krumholz, M. R., & Freeman, K. 2010, ApJ, 713, 166
Bovy, J. 2016, ApJ, 817, 49
Bovy, J., Rix, H.-W., Liu, C., et al. 2012, ApJ, 753, 148
Casey, A. R., Hawkins, K., Hogg, D. W., et al. 2017, ApJ, 840, 59
Casey, A. R., Hogg, D. W., Ness, M., et al. 2016, arXiv:1603.03040
Cirasuolo, M., Afonso, J., Bender, R., et al. 2012, Proc. SPIE, 8446, 84460S
Cunha, K., Smith, V. V., Johnson, J. A., et al. 2015, ApJL, 798, L41
Dalton, G., Trager, S., Abrams, D. C., et al. 2014, Proc. SPIE, 9147, 91470L
de Jong, R. S. & Consortium 2015, IAUAGA, 22, 2255843
De Silva, G. M., Freeman, K. C., Asplund, M., et al. 2007, AJ, 133, 1161
De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, MNRAS, 449, 2604
Mitschang, A. W., De Silva, G., Sharma, S., & Zucker, D. B. 2013, MNRAS, 428, 2321
Mitschang, A. W., De Silva, G., Zucker, D. B., et al. 2014, MNRAS, 438, 2753
Ness, M., Hogg, D. W., Rix, H.-W., Ho, A. Y. Q., & Zasowski, G. 2015, ApJ, 808, 16
Nidever, D. L., Holtzman, J. A., Allende Prieto, C., et al. 2015, AJ, 150, 173
Oh, S., Price-Whelan, A. M., Brewer, J. M., et al. 2017, arXiv:1709.05344
Reddy, A. B. S., Giridhar, S., & Lambert, D. L. 2012, MNRAS, 419, 1350
Reddy, A. B. S., Giridhar, S., & Lambert, D. L. 2015, MNRAS, 450, 4301
Roeser, S., Demleitner, M., & Schilbach, E. 2010, AJ, 139, 2440
Schiavon, R. P., Zamora, O., Carrera, R., et al. 2017, MNRAS, 465, 501
SDSS Collaboration, Albareti, F. D., Allende Prieto, C., et al. 2017, ApJS, 233, 25
Souto, D., Cunha, K., Smith, V., et al. 2016, ApJ, 830, 35
Ting, Y.-S., Conroy, C., & Goodman, A. 2015, ApJ, 807, 104
Ting, Y.-S., Conroy, C., & Rix, H.-W. 2016, ApJ, 816, 10
Wilson, J. C., Hearty, F., Skrutskie, M. F., et al. 2012, Proc. SPIE, 8446, 84460H
Zahn, J.-P. 1992, A&A, 265, 115
Zasowski, G., Johnson, J. A., Frinchaboy, P. M., et al. 2013, AJ, 146, 81